



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería Informática



**TFG del Grado en Ingeniería
Informática**

**Herramienta para medir la
eficiencia en Python**



Presentado por Guillermo Calvo Álvarez
en Universidad de Burgos — 20 de junio
de 2019

Tutor: Jesús Enrique Sierra García



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería Informática



D. nombre tutor, profesor del departamento de nombre departamento, área de nombre área.

Expone:

Que el alumno D. Guillermo Calvo Álvarez, con DNI dni, ha realizado el Trabajo final de Grado en Ingeniería Informática titulado título de TFG.

Y que dicho trabajo ha sido realizado por el alumno bajo la dirección del que suscribe, en virtud de lo cual se autoriza su presentación y defensa.

En Burgos, 20 de junio de 2019

Vº. Bº. del Tutor:

Vº. Bº. del co-tutor:

D. nombre tutor

D. nombre co-tutor

Resumen

La eficiencia, definida como la capacidad de conseguir un objetivo con el mínimo número de recursos posibles, es algo que se busca en todos los ámbitos con mayor o menor prioridad.

En el caso de la informática, es algo muy importante, ya que de ello depende muchas veces que un programa sea viable para la puesta en práctica en un entorno laboral o productivo, sobre todo si hablamos de sistemas de tiempo real.

En este trabajo se abordará el tema de la medición de eficiencia, pero lo que en verdad se pretende es conseguir una herramienta capaz de mostrar los datos necesarios para hacer un análisis de eficiencia de códigos diseñados en lenguaje Python. (Archivos .py) E incluso hacer comparaciones entre estos.

Para ello se investigará sobre los diferentes métodos de medición de eficiencia existentes, la forma de interpretar código en lenguaje Python y, diferentes maneras realizar un análisis y hacer representaciones con la información resultante.

Descriptores

Eficiencia, Python, interpretar, análisis . . .

Abstract

A **brief** presentation of the topic addressed in the project.

Keywords

keywords separated by commas.

Índice general

Índice general	III
Índice de figuras	V
Índice de tablas	VI
Introducción	1
Objetivos del proyecto	3
2.1. Objetivos del Proyecto	3
Conceptos teóricos	5
3.1. Fundamentos Básicos	5
3.2. De ByteCode a una medida de eficiencia	7
3.3. Tipos de Analisis	10
3.4. Estructura interna	11
Técnicas y herramientas	13
Aspectos relevantes del desarrollo del proyecto	17
5.1. Transcurso del desarrollo	17
Trabajos relacionados	23
Conclusiones y Líneas de trabajo futuras	25
7.1. Conclusiones	25
7.2. Trabajos de Futuro	26
7.3. Introduccion	27

7.4. Decisiones Previas	27
7.5. Tipos de Operaciones	28

Índice de figuras

Índice de tablas

Introducción

La eficiencia en el mundo informático, es un termino fundamental y con mucho peso, que condiciona en la gran mayoría de ocasiones todo lo relacionado a este. Cuanta mayor eficiencia menos recursos se gastan para poder realizar un objetivo y estos recursos pueden ser utilizados para otras metas. La búsqueda de una buena eficiencia tiene una gran importancia dentro del ámbito laboral, donde cada recurso ahorrado puede ser fundamental, para el desarrollo de otras tareas. (En sistemas de tiempo real es algo fundamental)

Aunque la la eficiencia es algo que se puede aplicar a casi cualquier campo conocido, en este caso nos vamos a centrar meramente en la eficiencia aplicada a código Python, ya que es el tema principal de este trabajo. Para poder medir la eficiencia de un código hecho en lenguaje Python, primero se tendrá que empezar por determinar y desarrollar que es lo que hace que un código sea mas eficiente que otro y el como conseguir esa información.

A la primera pregunta la respuesta seria, "depende". Ya que según el entorno y el objetivo para el que ha sido creado un programa, lo que determina si un programa esta hecho de forma eficiente puede variar entre diferentes métricas. Para establecer un forma de medir la eficiencia en esta herramienta se ha decidido utilizar las operaciones como forma primaria de obtención de información.

Hay que dejar claro que hay mas formas de medir la eficiencia de un programa como por ejemplo:

- Medir el tiempo de ejecución: Esta es la forma de medirla eficiencia

mas utilizada, esto se debe a la facilidad de llevarla a cabo, pero como gran inconveniente se podría decir que es una manera que devuelve resultados muy poco afinados. Esto es por la gran influencia que tiene algunos factores externos en los resultados que puede devolver. El mayor de estos factores seria el entrono de pruebas donde se ejecuta el código del programa, ya que si hacemos una ejecución en dos ordenadores totalmente diferentes, es muy probable que algoritmos con una cierta complejidad tarde mas en ser ejecutados en un equipo que en otro. Con este método para poder sacar resultados con una gran fiabilidad habría que tener un gran control sobre el entorno, lo cual requiere de una gran preparación previa.

- Complejidad O: Esta forma realmente nos sirve para conocer la complejidad del código, pero también es a veces es tomado como referencia para hacerse una estimación de la eficiencia. Al ser tan solo una estimación no se puede realizar un análisis en profundidad sobre este aspecto.

Pero se quiere implementar la medición con operaciones por ser una manera con soluciones invariables sea cual sea el entono en el que se haga y devuelve datos con una cierta profundidad como para poder hacer todos los análisis que se requieran.

En el caso de este trabajo se ha decidido que para conseguir la información de cuantas operaciones hay en la ejecución de un programa se utilizara un interprete, que como bien su nombre indica, sera el responsable de interpretar el código que forma el programa y tras hacerle unas modificaciones nos guardara la información que deseada.

Objetivos del proyecto

2.1. Objetivos del Proyecto

Objetivo general

Desarrollar un software capaz de dar la información necesaria para realizar un análisis de eficiencia sobre un código de extensión .py.

- Poder hacer análisis de un código python a través del cálculo de operaciones.
- Mostrar esa información de una manera agradable y entendible.

Objetivos funcionales

- El usuario podrá elegir el tipo de análisis que quiera realizar: de un fichero individual o comparando varios ficheros.
- El usuario podrá modificar la información resultante, eligiendo las operaciones que saldrán.
- El usuario puede cambiar la métrica de los análisis, decidiendo como pondera cada tipo de operación
- El usuario podrá realizar varios análisis sin necesidad de salir de la herramienta
- (Exportar la información?)

Objetivos tecnicos

- Obtener el numero de operaciones a través del análisis por parte del interprete
- Diseñar una manera de guardas los niveles de eficiencia de cada tipo de operación.
- Obtener los datos finales con los que realizar la gráficas a través de la ponderación
- Utilizar Tkinter para realizar la interfaz de la herramienta
- Utilizar Matplotlib para generar las gráficas tras los respectivos análisis
- (alguna herramienta mas?)

Conceptos teóricos

Algunos conceptos teóricos de L^AT_EX¹.

3.1. Fundamentos Básicos

La eficiencia

La eficiencia vista desde un ámbito general se definiría como el uso razonado de los recursos para conseguir un objetivo x . Esta definición puede ser aplicada a cualquier ámbito en el que se hable de eficiencia pero en el caso de este proyecto se aplicara esta definición en el ámbito computacional. La definición citada anteriormente menciona la palabra recursos, en el ámbito que estamos tratando, cuando se habla de recursos la palabra que más suele asociarse es tiempo. Esto no quiere decir que sea la única medida de eficiencia, ni mucho menos, también se podría poner como recurso el espacio en memoria, pero el tiempo suele ser el recurso que más condiciona y la gente más en cuenta tiene en las características de los programas.

Aquí llegamos a la razón de existir de este proyecto, cómo medimos lo que tarda un programa en ejecutarse de manera fiable? La manera más extendida para hacer esto, es ejecutar el programa que se desee y esperar a que termine, contado el tiempo pasado desde el comienzo hasta el final de este. Esta manera a pesar de ser sencilla, a la hora de la verdad es muy poco fiable. Para que los resultados obtenidos por este método sean considerados verídicos se tendría que asegurar que el entorno de pruebas tenga unas condiciones idóneas, como por ejemplo que el procesador que del ordenador

¹Créditos a los proyectos de Álvaro López Cantero: Configurador de Presupuestos y Roberto Izquierdo Amo: PLQuiz

donde se ejecuta el programa no este gastando recursos en otros procesos que estén activos en segundo plano. El problema de esto es que conseguir un entorno así muchas veces no es tarea fácil. Y como respuesta a este problema surge la idea principal de este proyecto, medir la eficiencia de un programa de una manera fiable, que no se vea afectado por factores externos a este. Para lograr un método métrico para la eficiencia de un programa sin que su valor se vea afectado por el cuándo y dónde se haga el análisis, se decidió hacerlo a través de las operaciones que se ejecutan en el código del programa. Y una vez sacados ponderarlos según la complejidad de cada uno.

El interprete

Un interprete es un programa que analiza y ejecuta el código de otro programa instrucción a instrucción. Esta es la manera con la que se pueden detectar los diferentes elementos que componen un programa. Al realizar este trabajo en Python hay que comentar antes que el interprete se trata de una maquina de pilas, lo que no deja de ser una maquina virtual que emula el funcionamiento de un ordenador. El funcionamiento principal de este se basa en almacenar y gestionar información que va recogiendo en las diversas pilas que tiene, para de esta manera poder hacer sus operaciones.

El interprete elegido para esta practica se trata de ByteRun, un interprete hecho en código python y para código python. Este fue elegido por ser un interprete que fue creado con el objetivo principal el aprendizaje, lo cual conlleva que su complejidad no sea muy alta y por lo tanto es relativamente fácil implementar nuevas funcionalidades en este. Por otro lado en contra parte de esa baja complejidad hay que mencionar que ByteRun no es un intérprete muy rápido, como por ejemplo podría ser Cpython, pero como en este trabajo el principal requisito que se busca en un interprete es el de poder implementar nuevas funcionalidades en este para conseguir una forma de poder detectar las operaciones ejecutadas y guardarlas. El ByteRun era una opción muy acertada que se adaptaba perfectamente a los requisitos del desarrollo.

Bytecode

. El ByteRun al ser un interprete de Python, no analiza directamente el código de alto nivel, lo que este analiza es un código intermedio llamado Bytecode.

El ByteCode se trata de un código intermedio sacado a través de la traducción de un código de alto nivel al ser tratado por una serie de procesos (lexer, compiler...). Para sacar este código intermedio se ha utilizado un

modulo de Python que actúa de desensamblador y descompone el código de alto nivel. El aspecto de este código se trata de una serie de bytes que a simple vista una persona no podría entender.

Imagen de bytes

Pero el modulo Dis, cuenta con una función que permite ver el ByteCode en una especie de lenguaje ensamblador, lo cual resulta muy útil para entender la estructura del código.

Imagen ByteCode Ensamblador

La forma de entender este tipo de código es la siguiente:

Como ejemplo usaremos la primera fila donde se encuentra la instrucción "LOAD CONTS", la primera columna muestra el número 1 que representa la línea donde está la instrucción en el código de Python, la segunda columna es un índice que indica que la instrucción "LOAD CONTS", en En este caso, está en la posición 0, la tercera columna es el nombre de la instrucción en sí misma, mostrándola con un nombre comprensible para una persona, la cuarta columna indica la posición de ese argumento en la pila y la última columna muestra cuál es el argumento para la instrucción.

3.2. De ByteCode a una medida de eficiencia

Como bien se ha comentado antes, una vez obtenido el Bytecode, el interprete, en nuestro caso el ByteRun, es capaz de ir analizando y ejecutando el bytecode obtenido. Pero el ByteRun no es capaz ir almacenando las operaciones que va traduciendo. En este momento es donde se debe implementar una nueva funcionalidad en el interprete para que sea capaz de hacer esto, para así mas tarde poder hacer un calculo de la eficiencia.

El Diccinario de procesos

. Cuando el ByteRun lee una instrucción y lo primero que tiene que hacer es identificar que tipo de instrucción esta leyendo, hay muchos tipos

diferentes desde un "LOAD", a un "LOOP", o un ".ADD", pero a nosotros no nos interesan todas las instrucciones, solo queremos almacenar aquellas que son consideradas por nosotros como ".operaciones", o lo que es lo mismo, instrucciones que necesiten de algun tipo de calculo por parte del procesador, y que por lo tanto tenga un tiempo de ejecución.

Por suerte para nosotros el interprete ya hace distinciones entre los diferentes tipos de instrucciones y guarda en una lista todos aquellos valores en los cuales requiere hacer una operación.

Aprovechando esto, cada vez que alguno de esas instrucciones es llamada guardamos el nombre que que tiene esta al ser visualizado el bytecode mediante el modulo dis en un diccionario, el cual decidimos llamar diccionario de operaciones.

Este diccionario tiene la siguiente estructura:

Imagen ejemplo

Se trata de un diccionario compuesto por una lista de clave primaria, y los componentes de esta lista son los siguiente:

- Nombre de la operación detectada por la interfaz
- Primeras 3 Iniciales del tipo del primer operador
- Primeras 3 Iniciales del tipo del segundo operador

Esta estructura para las keys del diccionario se basa en el hecho de que un operación necesita de dos valores con los que operar, y dependiendo de el tipo de estos valores, dos mismas operaciones con diferentes valores, pueden llegar a tener nivel de eficiencia mucho mas bajo o alto que el otro.

Según el ByteRun va analizando todo el código cuando detecta una instrucción dentro de la lista de operaciones este llama a una función, que en caso de no tener ninguna entrada de esa operación y sus respectivos operadores la genera y pone de valor un 1. Posteriormente si se va encontrando mas operaciones iguales en vez de crear otra entrada en el diccionario de

operaciones, incrementa en 1 el valor enlazado con la clave que la corresponde.

De esta manera cuando el ByteRun acaba de analizar todas las operaciones, conseguimos tener un diccionario con todas las operaciones que han surgido a lo largo de la ejecución del programa.

Ponderación

Pero con solo tener los tipos de operaciones no es suficiente. Para hacer un análisis de la eficiencia es necesario saber el nivel de eficiencia que dispone cada tipo de operación, ya que no todas las operaciones tiene la misma complejidad y el procesador necesita hacer mas cálculos para ejecutarlos. Esto se consigue gracias a el procesador teórico.

Procesadores Teóricos

El procesador teórico tiene la función de darle un nivel de eficiencia a cada tipo de operación. Se le puso este nombre por ser la parte que simula la complejidad que le supondría a un determinado procesador, de esta manera según los valores que se guarden en el procesador se pueden simular diferentes tipos de entornos.

Estos procesadores se tratan de unos ficheros de tipo .csv que guardan una matriz, llamada matriz de traducción, la cual esta hecha de tal tiene una estructura pensada para poder sacar mas los niveles de eficiencia que guarda en cada celda según sea necesario.

La dos primeras celdas de la matriz hacen de indice para las tuplas, la primera columna guarda el tipo de operación mientras que la segunda guarda el tipo del primer operador. Siguiendo esta estructura si tuviéramos por ejemplo una suma de un entero con un decimal, para encontrar el nivel de eficiencia que habría que aplicarle a esta operación en concreto la herramienta busca primero el tipo del operador, en este caso seria `^DDz` seguido a esto busca en la segunda columna deja el valor `int`, una vez encontrados estos dos valores ya tendría localizada la tupla donde se encuentra el valor deseado, y aquí es donde entra en juego el indice de las columnas que se haya en la

primera fila de la matriz, esta guarda los diferentes tipos que puede tomar el segundo operador, en este caso "flo", y con esto puede localizar que celda de la tupla es la que se necesita.

Foto ejemplo

Ponderación

Una vez tenemos tanto el diccionario de operaciones como un Procesador teórico ya solo falta realizar la ponderación para obtener los datos necesarios para hacer los análisis oportunos.

La manera en que se hace esto es bastante simple, consiste en darle un valor a cada tipo de operación (sumatorio) formulario

ejemplo

3.3. Tipos de Analisis

Con los datos obtenidos ya se pueden hacer diferentes tipos de análisis, en el caso de esta herramienta se han implementado dos tipos:

- Análisis individual
- Análisis comparativo

Analisis Individual

En este análisis solo se permite la entrada de un fichero, de formato .py, el cual tras pasar por los pasos citados en los puntos anteriores, muestra una gráfica circular con las diferentes operaciones que el interprete(ByteRun) a detectado al ejecutarlo. Y además de esto deja re calcular el resultado mostrado en el gráfico según las operaciones que queramos mostrar o no y el procesador con el que se quiera hacer la ponderación.

Este análisis esta pensado para optimizar el código del programa, ya que de esta manera puedes detectar que operaciones son las que mas recursos consumen.

Análisis Multiple

En este análisis se deben elegir mas de un fichero, los cuales ejecuta y calcula los ciclos de reloj que tarda en ejecutarse cada uno, pero al mostrar

los resultados varia según el numero de ficheros que se hayan escogido. En caso de ser menos de 10 saldrá una gráfica de barras en la cual cada barra representa el total de ciclos de reloj de cada fichero. Y en caso de que se metan mas de 10 se cambiaría por una gráfica de puntos para facilitar la visibilidad de los resultados. Al igual que en el análisis individual en todo momento se pueden cambiar los resultados de la gráfica cambiando las operaciones que queramos que se tengan en cuenta y el procesador con el que se quiera ponderar las operaciones.

Este análisis esta pensado para compara ficheros que tengan como finalidad hacer lo mismo, y de esta manera saber cual de ellos tiene en que clase de entornos mejor eficiencia.

3.4. Estructura interna

A la hora de programar todo esto, al inicio se planteo hacerlo con a través del modelo, vista, controlador. Pero a medida que se fue avanzando mas en el desarrollo esto se fue difuminando debido al tipo de desarrollo que

Técnicas y herramientas

Metodología Ágil

Subsubsecciones

(No se muy bien que mete aquí) Para mantener un flujo de trabajo en el desarrollo del proyecto se ha decidido implementar la metodología agil XXXX, esta metodología se

Herramienta de documentación

Latex

Es un sistema de composición de textos especializado en los textos técnicos y científicos. Se decidió hacer este trabajo con este sistema para componer la documentación, porque facilita de gran manera la buena estructuración del documento, gracias a su , además de ofrecer una gran cantidad de aspectos tipográficos, que consiguen dar una gran calidad y profesionalidad a los documentos resultantes y al componer Latex texto mediante marcas en un archivo fuente, esto permite previsualizar el documento desde cualquier entorno sin perder el formato, lo resulta extremadamente útil para el desarrollo de documentos como el que estas leyendo. En mi caso he utilizado como editor de texto TexMaker, ya que me gustaba las disposición en la que dispone su interfaz y me facilitaba mucho ([Enlace pagina de TexMaker](#))

(Desarrollar un poco mas)

Herramienta de Gestion

GitHub

Se trata de una plataforma online donde la gente puede almacenar y gestionar los proyectos que estén desarrollando mediante gestión de versiones.

Ha sido la herramienta de gestión principal elegida debido a que es la que nos han enseñado su uso en la Universidad y por lo tanto con la que más practica se tenía de antemano, además de ser gratuita para proyectos de código abierto.

La hemos utilizado para albergar el código del proyecto, y desde ahí poder gestionar el avance del desarrollo

(Desarrollar algo mas)

Herramientas Proyecto

Python

Se a utilizado este lenguaje de programación para realizar el trabajo, esto se debe a las gran cantidad de librerías que dispone este lenguaje para poder operar con una gran variedad de componentes, lo que ya de por si nos ofrece mucha flexibilidad. También hay que destacar que se elgio este lenguaje por ser un tipo de lenguaje interpretado, lo cual hacia que fuese mas facil tratar el tema del interprete. La versión con la que se ha contado a lo largo del proyecto ha sido la 2.7.X. Se intento utilizar un versión 3.7 y 3.6, pero tras tener varios problemas de consistencia entre este y el interprete que elegimos para la realización del proyecto, al final esto nos forzó a tener que probar una versión 2.

(Meter: Version? Licencia? PaginaWeb)

Tkinter

Es una librería de Python para el desarrollo de interfaces gráficas. En principio se barajó la posibilidad de utilizar algún otro tipo de librería para realizar la interfaz como por ejemplo C, pero al tener dificultades de implementar otras librerías la versión de Python 2.7., y no tener Tkinte una gran dificultad de aprendizaje para poder realizar una interfaz básica.

(Meter: Version? Licencia? PaginaWeb)

Matplotlib

Es una librería diseñada para representar datos en gráficos 2D. Es una librería clave en el proyecto para la representación de los gráficos resultantes de los análisis.

(Meter: Version? Licencia? PaginaWeb)

Microsoft Excel

Es una aplicación de hoja de cálculos, utilizada mayoritariamente para tareas financieras y de logística. En el caso de este proyecto se utilizó únicamente para dar valores a las matrices de traducción, ya que la estructura de celdas que tiene esta aplicación facilitaba mucho hacer esta tarea.

(Meter: Version? Licencia? PaginaWeb)

Csv

Modulo de Python que permite abrir, leer y escribir en ficheros .csv. En el caso de este trabajo ha sido utilizado para poder leer los procesadores teóricos y poder sacar los valores deseados para realizar la ponderación.

os

Se trata de un modulo de Python, que permite acceder a funciones que permite leer y escribir archivos y acceder al sistema de archivos. En el caso de este proyecto se ha utilizado para acceder al sistema de archivos y poder recoger las rutas de algunas carpetas de forma automática.

Dis

Un modulo de Python que desensambla el código origen y lo transforma a ByteCode descomponiendo cada parte del código original. Este modulo es vital para el funcionamiento del ByteRun ya que es el encargado de transformar el código de alto nivel a ByteRun

tkk

Esta parte de la memoria tiene como objetivo presentar las técnicas metodológicas y las herramientas de desarrollo que se han utilizado para llevar a cabo el proyecto. Si se han estudiado diferentes alternativas de metodologías, herramientas, bibliotecas se puede hacer un resumen de los aspectos más destacados de cada alternativa, incluyendo comparativas entre las distintas opciones y una justificación de las elecciones realizadas. No se pretende que este apartado se convierta en un capítulo de un libro dedicado a cada una de las alternativas, sino comentar los aspectos más destacados de cada opción, con un repaso somero a los fundamentos esenciales y referencias bibliográficas para que el lector pueda ampliar su conocimiento sobre el tema.

Aspectos relevantes del desarrollo del proyecto

Planteamiento principal (Diferentes concepciones del trabajo según se investigan/desarrollaba)

(Meter que conceptos aprendidos de la uni?)

Aquí hablar en cada apartado de los inconvenientes que han ido surgiendo.

5.1. Transcurso del desarrollo

El planteamiento inicial del proyecto era crear una herramienta para medir la eficiencia del código Python, en base a los ciclos de reloj y el espacio en memoria que este ocupa. Al ser este un planteamiento previo al inicio del desarrollo, este ha ido dando algunos cambios según nuevos requisitos han ido surgiendo o restricciones con las que no habíamos contado han ido apareciendo. Y aquí se intentara contar como han surgido esos cambios inesperados y como han afectado este al desarrollo del proyecto.

Lo primero por donde se partió el proyecto fue el intérprete. Se buscó un intérprete capaz de satisfacer las necesidades que se plantearon. En esta etapa surgieron varias propuestas, pero tras algunas puestas en común, se llegó a la conclusión de que el intérprete ByteRun era la mejor idea para hacer un primer intento.

ByteRun al final fue el intérprete elegido porque nos proporcionaba una serie de requisitos que encajaban a la perfección con lo que teníamos planeado previamente. He aquí alguno de esos requisitos:

- Estaba desarrollado en Python: Al ser Python un lenguaje conocido, que ha sido impartido en la Universidad, era preferible que estuviese hecho en un lenguaje que ya tuviésemos conocimientos previos. Además que un intérprete de Python hecho en Python nos parecía que tenía un cierto atractivo.
- Los autores de este interprete contaban con una buena documentación sobre el funcionamiento de esa herramienta, lo cual nos venía como anillo al dedo, para así entender mejor su funcionamiento y poder hacer mejor el upgrade a este.
- Es un intérprete hecho con la finalidad del aprendizaje, esto hace que no sea el más rápido, pero esto también conlleva el hecho de que no sea muy complejo. Y como el objetivo del proyecto no es tener la herramienta más rápida, si no que funcione y seamos capaces de implementar los cambios necesarios, esto para nosotros era una gran ventaja.

Una vez se tenía claro que el ByteRun era el intérprete que haría de base para desarrollar la herramienta, lo primero era entender su funcionamiento y hacer pruebas con diferentes ficheros para comprobar que de verdad funciona.

En este punto nos encontramos con los primeros imprevistos, los autores del intérprete ya contaban con unos tests hechos, pero para poder ejecutar estos test utilizaban una `****` para Python llamada tox. Tox en algunos tipos de versiones de Python no funciona adecuadamente, tras informarnos sobre esta, decidimos implementarlo en Python 2.7. la última de las versiones 2.x de Python.

Una vez solucionado esto se probaron diferentes tests que estaban ya creados, los cuales funcionaban correctamente. Pero para poder medir los límites del interprete se generaron otra serie de tests con los cuales se pudo demostrar alguna de las limitaciones del interprete.

(Entre alguna de ellas están la implementación de algunas bibliotecas que el intérprete no reconoce?¿?) Seguido de las pruebas se pasó a plantear la modificación del intérprete. Al principio como la intención del proyecto era poder medir la eficiencia a través de las operaciones y de las direcciones de memoria, nos pusimos a investigar en que parte del intérprete conseguiríamos la información necesaria para esto, y la conclusión a la que llegamos no era la esperada. Esto se debió a que al ser el ByteRun un intérprete de Pila, este no guarda la información que va procesando directamente en direcciones de memoria, si no que las guarda en su sistema de pilas interno.

Por culpa de esto nos vimos obligados a cambiar el planteamiento principal y centrarnos en la eficiencia a través de las operaciones y diferentes formas sacar resultados a partir de esto. Una vez centrados en sacar las operaciones del interprete, como esperábamos, el interprete para saber como actuar según el tipo de instrucción que fuera a ejecutar, este distinguía entre diferentes tipos de instrucción, y para las que nosotros llamaríamos operaciones, tenía un par de listas con los nombres de estas según el ByteCode. Con esta diferenciación, nosotros tendríamos que guardar una instrucción cuando esta se encuentra en alguna de estas listas.

El proceso de guardado de matrices paso por varios procesos para al final conseguir al a nuestro criterio es el método mas optimo de entre todos ellos. En la primera idea implementada solo se crearon unas variables por cada tipo de operación disponible, esto a parte de ser poco eficiente, limitaba mucho el tipo de operaciones a analizar, ya que no se tenían en cuenta los tipos de los operadores.

La segunda idea implementada fue la de la matriz de operaciones: se decidió guardar las operaciones en una matriz con la misma estructura que posteriormente se aplicaría a la matriz de traducción.

(Imagen matriz) Reaprobar la del artículo

Esta idea ya daba una mayor variedad de operaciones, ya que si tenía en cuenta los tipos de los operadores, pero por otra parte era incluso menos eficiente que cuando solo había variables. La matriz hacia tener un gran numero de posiciones en la matriz sin ningún tipo de uso, ya que la matriz tenía unas 350 posiciones de las cuales por fichero se solían llenar entre 6 y 12. Esto equivalía un porcentaje muy grande de desuso, por lo que se decidió cambiar este método por otro.

El método que ya se decidió como definitivo, como ya se explica en el apartado de xxxxxx, se trata de un diccionario que guarda un contador por cada operación diferente, teniendo en cuenta los operandos. Y de esta manera nos ahorramos las celdas vacías que nos surgían en la matriz.

(Foto diccionario)

Tras implementar esto empezamos a plantearnos los distintos tipos de análisis que se harían con los datos obtenidos y la conclusión a la que llegamos fue el de hacer un análisis individual en el que solo mostrásemos los porcentajes de eficiencia de un fichero dado y otro en el pudiésemos comparar la eficacia entre varios ficheros.

Primero empezamos por el análisis individual, y en este caso como el controlador y la vista están unidas, se podría decir que era un análisis en paralelo, según se hacia esa parte del controlador, se implementaban los parámetros de la interfaz.

La interfaz también paso por una serie de versiones: Al principio la interfaz

tan solo se componía de una ventana donde se añadían elementos y a veces se superponían, literalmente era un caos.

Foto si es posible

Para evitar que los elementos colisionasen entre ellos , se paso a otro método donde cada vez que se se activasen algunos botones en particular, se añadiese una nueva ventana, esto evitaba el problema de las colisiones, pero por otra parte la final de los análisis era muy molesto, porque podías terminar con 4 ventanas diferentes abiertas, lo que al final era muy lioso.

Foto si es posible

Explicar funcionamiento de la interfaz

Alguna foto

Para evitar los problemas tanto de la primera versión de la interfaz como de la segunda, a través de investigar un poco mas en el funcionamiento de la biblioteca Tkinter y sus diferentes funcionalidades, se consiguió realizar una interfaz que utilizase solo una ventana, pero el contenido de esta se fuese cambiando según interactuemos con esta. A decir verdad la interfaz por dentro crea un numero de marcos, por así llamarlos, donde ya están todos los elementos que se tiene que mostrar en pantalla, cada marco representa un tipo distinto de ventana que puede surgir según se desarrolle la interacción con la herramienta, lo único que solo se muestran por pantalla aquel marco que corresponda con la funcionalidad que se este mostrando en ese momento.

Fotos

En concreto esta herramienta cuenta con un total de 6 (A revisar) Marcos que son los siguientes:

- - - - -

Es una manera muy útil de hacer una interfaz mucho mas fluida y sencilla para la vista. La única pega que se le podría poner a este método, es que al iniciar la aplicación tiene que generar los marcos para ir rotando entre ellos.

Esto ha condicionado más de una vez algunas complementaciones hechas en la interfaz.

Foto?

Para acabar con el desarrollo de las funcionalidades principales se crearon los llamados procesadores teóricos a través de la matriz de traducción. Como ya se ha explicado en puntos anteriores los procesadores con tienen una matriz que contiene valores que dan un nivel de eficiencia a los distintos tipos de operaciones. Para guardar la matriz se decidió utilizar el formato csv. Hay varios motivos detrás de esta elección:

1 La simplicidad del formato hace posible poder configurarlo con un gran numero de aplicaciones, aunque en el caso de este proyecto se opto por Microsoft Excel.

2 Hay una biblioteca en Python de la que ya teníamos algo de experiencia, que permitía leer y escribir en los tipos de archivos csv. (nombre de la biblioteca)

Afinales del desarrollo del proyecto se trato de implementar una funcionalidad mas, pero al final se decidió desechar esta, debido a que implicaba muchos mas recursos de los pensados en un planteamiento previo y para dejarlo inconcluso preferimos abortar.

Articulo

Una vez terminado el desarrollo de las funcionalidades principales se le dieron unas pinceladas a algunos aspectos de este como proporciones de algunos elementos en la interfaz, cambios en las nomenclaturas de variables, etc. Y mientras esto ocurría empezamos a desarrollar un articulo sobre el propio proyecto, para exponerlo en el congreso europeo, EuroSim. Pensamos que seria una oportunidad muy buena de ver si el trabajo echo hasta el momento era visto atractivo y útil por otras personas (Del sector.. explicar mas a fondo)

Trabajos relacionados

Este apartado sería parecido a un estado del arte de una tesis o tesina. En un trabajo final grado no parece obligada su presencia, aunque se puede dejar a juicio del tutor el incluir un pequeño resumen comentado de los trabajos y proyectos ya realizados en el campo del proyecto en curso.

Conclusiones y Líneas de trabajo futuras

Para terminar se contarán las conclusiones obtenidas tal el desarrollo de todo este proyecto y las posibilidades que deja abierta para desarrollar más cosas en el futuro:

7.1. Conclusiones

- Se consiguieron realizar los objetivos principales propuestos para este proyecto. Entre estos estaría la implementación de una forma de medir la eficiencia de una manera flexible, con el propósito de que sea útil en más de .
- La labor de comprensión realizada para entender un código hecho por otras personas, en este caso concreto para entender el funcionamiento del ByteRun, fue más duradero de lo esperado ya que a pesar de no ser un interprete muy complejo tenía una serie de conceptos sobre las pilas y el Bytecode, al final fue una buena manera de aprender algo más sobre los diferentes tipos de interpretes que había y en concreto saber más sobre los interpretes de pila.
- Otro gran reto de esta aplicación fue la interfaz gráfica, ya que casi no había tenido experiencia creando una. Tanto elegir las herramientas necesarias para hacerla como el posterior aprendizaje sobre el funcionamiento de Tkinter, es otra de las cosas que me llevo aprendidas tras este proyecto. Aunque no todas las ideas que se plantearon para este proyecto salieron finalmente a la luz, ha sido muy satisfactorio y

enriquecedor, ver como ideas que vas planteando se van plasmando poco a poco en algo que vas invirtiendo horas para finalmente formar aquello que estabas intentando lograr. Personalmente también me ha servido para aprender sobre la cuestión que hay que llevar en un proyecto de estas características y como reaccionar ante los diferentes obstáculos que uno se va encontrando en el desarrollo de este.

7.2. Trabajos de Futuro

Este proyecto tiene un fuerte potencial para seguir siendo desarrollado, ya que a partir de este punto puede evolucionar hacia diferentes frentes:

En la herramienta se han implementado aquellos análisis que se han creído mas interesantes en una primera instancia que puede ser usado en una ámbito mas general. Pero se pueden implementar mas análisis dependiendo del objetivo que se tenga detrás, por poner un ejemplo, un análisis que estuvo a punto de ser implementado también en este proyecto fue, el análisis de operaciones individuales, en el cual tras analizar el código de un fichero mostraría la progresión que ha tenido la ejecución una operación en concreto. Esta seria una buena forma de en que parte de la ejecución un código se aglomeran mas algunos tipos de operaciones. Y este es solo un ejemplo de los muchos tipos de análisis mas que se pueden implementar.

Y en lo que mas potencial creemos que puede derivar en un futuro esta herramienta seria el poder medir la eficiencia de códigos de casi cualquier tipo, esto seria relativamente sencillo tras haber desarrollado esta herramienta, ya que por la forma interna en la que esta construida . La interfaz esta separada del resto del programa y gracias a esto se podrían añadir interpretes que pudieran analizar tipos de lenguaje de programación, y al resto de la herramienta habría que hacerle unos cambios mínimos para que permitiese la entrada de ficheros con otro tipo de formato que no sea solamente .py. Esto haría que la herramienta fuese útil para casi cualquier tipo de desarrollo.

Como últimos planteamiento también se podría mirar la migración de version, de la version 2.7 de Python a alguna de las versiones 3. Algunos elementos de la interfaz podrían se mejorables como el evitar la superposición de los nombres de los tipos de operaciones detectados en el análisis individual cuando se trata de un fichero con muchos tipos.

7.3. Introduccion

El objetivo de este TFG es el de desarrollar una herramienta capaz de analizar la eficiencia de un código hecho en lenguaje Python e indicar posibles mejoras en este.

7.4. Decisiones Previas

Dejare indicado en esta parte del documento aquellos detalles que he decidido tomar antes de la realizacion de TFG.

Herramientas

Para crear la documentación y memoria del TFG, utilizare la distribución de latex MikTeX 2.9.6 con el editor de texto TexMaker. Para desarrollar el código de la herramienta utilizare el entorno de desarrollo de anaconda (jupyter/spider)

Metodología de trabajo

Para seguir un flujo de trabajo constante y gestionar el trabajo, he decidido utilizar el método kanban, ya que me parece un método bastante sencillo de llevar a cabo y es el que creo que mejor se ajusta a el tipo de desarrollo que utilizare en este trabajo.

En este caso he destaco dos figuras que influyen en el flujo de trabajo y son: El desarrollador: en este caso soy yo el que ocupa este papel, se trata de la persona que va realizando las tareas que estan en el estado "to do" se encarga de pasarla al estado "in progress" al final en "done".

El Planificador: Esta papel lo lleva a cabo mi tutor?¿?¿?¿?, el cual se encarga de ir seleccionando cuales son las tareas que deben ponerse con estado "to do", y asesoramiento de cual seria la mejor forma de afrontar estas tareas.

—— Comentar las reuniones semanales —————

Información previa

Utilice la siguiente información sobre el bytecode:

<https://pymotw.com/2/dis/>

<http://www.aosabook.org/en/500L/a-python-interpreter-written-in-python.html>

<http://akaptur.com/blog/2013/08/14/python-bytecode-fun-with-dis/>

7.5. Tipos de Operaciones

1. Suma
2. Resta
3. Multiplicación
4. División (/,//)
5. Potencia
6. Modulo