

Apéndice A

A.1. Extensión de la Herramienta IDA

La herramienta Identifier Analyzer (IDA) tiene una característica extra, puede recibir como parámetro de entrada la ruta¹ de un archivo XML (Extensible Markup Language). Este archivo debe contener información asociada a ids, literales y comentarios (propia de una aplicación JAVA); esta información es similar a la que es capturada por el Analizador Sintáctico que IDA posee (Módulo de Extracción de Datos - Ver Capítulo 4). La herramienta IDA se encarga de leer la información provista por el archivo xml y ejecutar directamente los algoritmos de análisis de ids que tiene implementada. Por último, los resultados de la ejecución se escriben en otro archivo XML que será creado en la misma ruta que el archivo leído como entrada (Ver Figura A.1).

La ventaja de que IDA soporte interacción con archivos XML, permite un intercambio de datos estándar con otras aplicaciones, lo que conlleva a tener compatibilidad para compartir información. De esta forma, IDA puede formar parte de un proceso de análisis más extenso que involucre otras herramientas asociadas a la comprensión de sistemas.

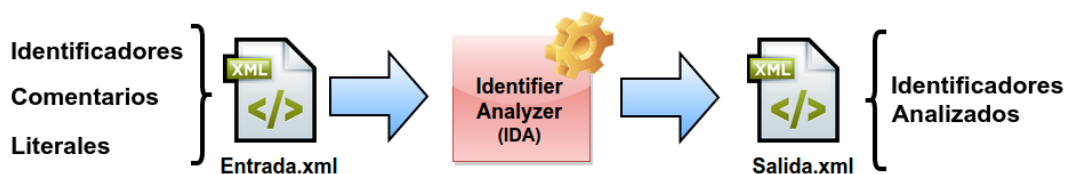


Figura A.1: Arquitectura de la Extensión de IDA

¹La ruta del archivo en un sistema de archivos de un sistema operativo determinado.

Para ingresar en IDA un archivo XML, simplemente se realiza a través del archivo JAR¹ (correspondiente a la herramienta), ejecutando la siguiente orden en la línea de comandos del sistema operativo²:

```
java -jar IDA.java <argumento>
```

En <argumento> se coloca la ruta donde se encuentra ubicado el archivo XML a procesar (un ejemplo en linux es /home/entrada.xml³). Este argumento no es obligatorio, y en caso de no pasarlo, se ejecuta la interfaz normal de IDA que fue descripta en el capítulo 4. La herramienta IDA procesa el archivo XML ingresado ejecutando los algoritmos de análisis de ids que tiene implementados (Greedy, Samurai y Expansión Básica). A continuación, se describe como debe estar estructurado el archivo XML de entrada.

Archivo XML de Entrada

El archivo xml que se ingresa, el comienzo debe marcarse con <entrada> y el fin con </entrada> (Ver Figura A.2), en su interior puede contener las siguientes listas de elementos:

Lista de Identificadores: Lista de ids que van a ser analizados, se enmarca con <lista_ids> y </lista_ids>; cada elemento de esta lista se indica con <id> y </id>; dentro de cada uno de estos elementos se coloca: el nombre del id con <nombre>**nmId**</nombre>, y el número de línea con <linea>**45**</linea> (Ver Figura A.2).

Lista de Frases: Listado de frases (asociadas a comentarios y literales) el inicio y fin de esta lista se indica con <lista_frases> y </lista_frases>; cada elemento de esta lista se enmarca con <frase> y </frase>; dentro de cada uno de estos elementos de la lista se agrega: la frase correspondiente con <texto>**File System**</texto>, y el número de línea con <linea>**25**</linea> (Ver Figura A.2).

¹JAVA Archive.

²Se recomienda utilizar los sistemas operativos Windows o UNIX con *java runtime enviroment* instalado

³No es necesario que se llame entrada, pero si que tenga extensión xml.

Lista de Clases: Este listado se corresponde a las clases que posee el programa, se enmarca con `<lista_clases>` y `</lista_clases>`; cada elemento de este listado se indica con `<clase>` y `</clase>`; en cada uno de estos elementos se agrega: el nombre del método `<nombre>Person</nombre>`, el número de línea donde comienza la clase `<linea_inicio>12</linea_inicio>`, y la línea donde finaliza la clase `<linea_fin>38</linea_fin>` (Ver Figura A.2).

Lista de Métodos: Similar al listado anterior pero para métodos, se enmarca con `<lista_metodos>` y `</lista_metodos>`; cada elemento de este listado se indica con `<metodo>` y `</metodo>`; en cada uno de estos elementos se coloca: el nombre del método `<metodo>getPerson</metodo>`, la línea donde comienza el método `<linea_inicio>20</linea_inicio>`, y la línea donde finaliza la método `<linea_fin>25</linea_fin>` (Ver Figura A.2).

Es importante tener en cuenta, que el dato obligatorio en el archivo xml ingresado son los nombres de los ids, dado que es la principal fuente de información a analizar. Por otro lado, el resto de los datos: Métodos, Clases, Frases, números de líneas (de cualquier elemento), también son importantes, pero solo colaboran con el análisis de los ids y no son indispensables.

Luego de que IDA analiza los ids, el próximo paso es escribir los resultados de cada ejecución en un nuevo archivo XML de salida que se describe en la siguiente sección.

```
<entrada>
  <lista_ids>
    <id>
      <nombre>nmId</nombre>
      <linea>10</linea>
    </id>
    <id>
      <nombre>fs</nombre>
      <linea>12</linea>
    </id>
  </lista_ids>
  <lista_frases>
    <frase>
      <texto>name identifier</texto>
      <linea>9</linea>
    </frase>
    <frase>
      <texto>file system</texto>
      <linea>19</linea>
    </frase>
  </lista_frases>
  <lista_clases>
    <clase>
      <nombre>Person</nombre>
      <linea_inicio>7</linea_inicio>
      <linea_fin>30</linea_fin>
    </clase>
  </lista_clases>
  <lista_metodos>
    <metodo>
      <nombre>getPerson</nombre>
      <linea_inicio>11</linea_inicio>
      <linea_fin>19</linea_fin>
    </metodo>
  </lista_metodos>
</entrada>
```

Figura A.2: Ejemplo de Archivo XML de entrada.

Archivo XML de Salida

El archivo de salida contiene los ids analizados por las distintas técnicas y se crea en la misma ubicación que el archivo XML pasado por entrada (siguiendo con el ejemplo de la sección anterior se creará en `/home/salida.xml`¹). Este archivo xml se le indica el inicio con `<salida>` y el fin con `</salida>` (Ver Figura A.3), en su interior posee la siguiente lista:

Lista de Identificadores Analizados: Es el listado de los ids incluyendo el análisis realizado en cada uno, `<lista_analisis_ids>` señala el comienzo a la lista y `</lista_analisis_ids>` indica el fin; cada elemento de la lista se indica con `<id>` y `</id>`; dentro de cada elemento de la lista se encuentra: el nombre del id analizado enmarcado con `<nombre>nmId</nombre>`, la división greedy del id se ubica entre `<div_greedy>nm-id</div_greedy>`, la división samurai del id entre `<div_samurai>nm-id</div_samurai>`, la expansión desde greedy entre `<exp_greedy>name identifier</exp_greedy>`, y la expansión desde samurai entre `<exp_samurai>name identifier</exp_samurai>` (Ver Figura A.3).

```
<salida>
  <lista_analisis_ids>
    <id>
      <nombre>nmId</nombre>
      <div_greedy>nm-id</div_greedy>
      <div_samurai>nm-id</div_samurai>
      <exp_greedy>name identifier</exp_greedy>
      <exp_samurai>name identifier</exp_samurai>
    </id>
    <id>
      <nombre>fs</nombre>
      <div_greedy>fs</div_greedy>
      <div_samurai>fs</div_samurai>
      <exp_greedy>file system</exp_greedy>
      <exp_samurai>file system</exp_samurai>
    </id>
  </lista_analisis_ids>
</salida>
```

Figura A.3: Ejemplo de Archivo XML de salida.

¹Si ya existe un archivo con el nombre `salida.xml`, el mismo se sobrescribirá.