

Memoria Algoritmo A Estrella

Nerea Jiménez González

Yhondri Acosta Novas

Marzo 2020



UNIVERSIDAD
COMPLUTENSE
MADRID

Índice

1. Detalles de la implementación	3
1.1. Lenguaje utilizado	3
1.2. Procedimiento seguido para la implementación	3
1.3. Ampliaciones realizadas	4
2. Código ejecutable	5
2.1. Simulación	5
3. Manual de usuario	7

Índice de figuras

1. Ventana inicial	5
2. Ventana elección de archivos	6
3. Ventana final	7
4. Botón para selección de archivos	8
5. Botón para ejecutar ID3	8

1. Detalles de la implementación

1.1. Lenguaje utilizado

Para la implementación, se utiliza el lenguaje JAVA.

1.2. Procedimiento seguido para la implementación

En un primer lugar, resolvemos el problema que se da en el enunciado de la práctica para poder utilizarlo de base.

Para la lectura del archivo, creamos una clase *Reader*, en la cual implementamos la lectura de los archivos de entrada.

Para la construcción del árbol final, implementamos una clase *Node*, cuyo constructor recibe un *String*, que será el valor seleccionado tras aplicar el **ID3**, y otro nodo para seguir en una futura recursión.

En la clase *ID3* es donde implementamos el algoritmo como tal. Tenemos dos listas, una de *String attributes*, donde almacenamos los atributos leídos, y *data*, que es una de listas de *String* que almacena los valores de los atributos. Tenemos un atributo *string targetAttribute*, el cual va a ser el valor del primer nodo. Por defecto, elegimos el primer atributo, pero este puede cambiar en la primera iteración de la recursión. A partir de la primera recursión, el valor de este atributo no cambia, ya que va a ser el nodo del cual cuelgue el resto del árbol. Guardamos el resto de atributos no elegidos en una lista de string llamada *noTargetAttributes*, y a partir de aquí aplicamos el **ID3**.

Para empezar guardamos en un map llamado *targetValues*, el número de veces que aparece cada valor del atributo *targetAttribute*. Si el tamaño del map es de **1**, significa que hemos llegado al final de ese subárbol. Si por otro lado, está **vacía**, no tenemos información por lo que no seguimos por esa rama, y volvemos hacia atrás.

Si todo va bien, a continuación, calculamos el mérito de cada atributo y nos quedamos con el mejor, es decir, nos quedamos con el de menor mérito. Para ello, iteramos sobre cada uno de los atributos y obtenemos el mérito a partir de los valores. Iteramos sobre cada atributo de la lista *restoAtributos*. De cada atributo iteramos en un bucle interior sobre los valores de ese atributo. Para el cálculo del mérito, usamos varias funciones:

- *getPartitionFromData*: Calcula el número de repeticiones que tiene cada uno de los posibles atributos del atributo dado. Por ejemplo: *target*

tAttribute = Temperatura -¿caluroso(4), templado(6), frío(4).

- *getTargetValuesFromData*: Obtiene los ejemplos a partir del atributo dado. Ejemplo: Temperatura -¿Caluroso(4).
- *getEntropy*: Calculamos la entropía siguiendo la fórmula $infor(p, n) = -p * \log_2 p - n * \log_2 n$

A continuación, calculamos el **infor** que es la suma de los méritos de los valores. Finalmente comparamos y nos quedamos con el menor.

El siguiente paso es un punto de parada. Si no hemos obtenido mérito estamos al final de un subárbol, por lo que obtenemos el valor del nodo y lo devolvemos.

El siguiente paso es preparar el siguiente subárbol, esta llamada recursiva se trata de la ampliación de la práctica, por lo que la explicamos en el siguiente apartado.

Al final de esta función, hemos obtenido el atributo elegido y sus atributos, que es el valor que se retorna en un objeto *Node*. Para construir este objeto, usamos la función auxiliar *getNoRepeatedValues* para obtener los atributos del mejor atributos.

1.3. Ampliaciones realizadas

Para la explicación de la recursividad, vamos a utilizar el atributo **Temperatura** como si fuese el atributo mejor de la iteración.

Una vez hemos elegido el primer nodo, aplicamos la recursividad para calcular los subárboles.

Calculamos los nuevos atributos a partir del nodo elegido siguiendo los siguientes pasos:

1. En la variable *finalAttributesMaxValue* guardamos el valor de nodo atributo seleccionado, por ejemplo **Temperatura**
2. A continuación filtramos para quedarnos con los atributos de que aún no han sido seleccionados, siguiendo el ejemplo serían **Humedad, Viento y Jugar** y los guardamos en una lista de strings llamada *newRestoAtributos*
3. Obtenemos los atributos del atributo elegido, de **Temperatura** serían, caluroso, frío, templado y lo guardamos en un set de *Strings* llamado, *noRepeatedValues*

4. Iterando sobre *noRepeatedValues*, dejamos en una lista llamada *partition* los atributos restantes de todos los elementos de *noRepeatedValues* excepto el mejor

Finalmente, devolvemos el valor de `id3(partition, newRestoAtrtributos)`.

2. Código ejecutable

Se adjunta en el archivo formato zip, con nombre **Código**.

2.1. Simulación

Al ejecutar `p2.jar` se abre la siguiente ventana.

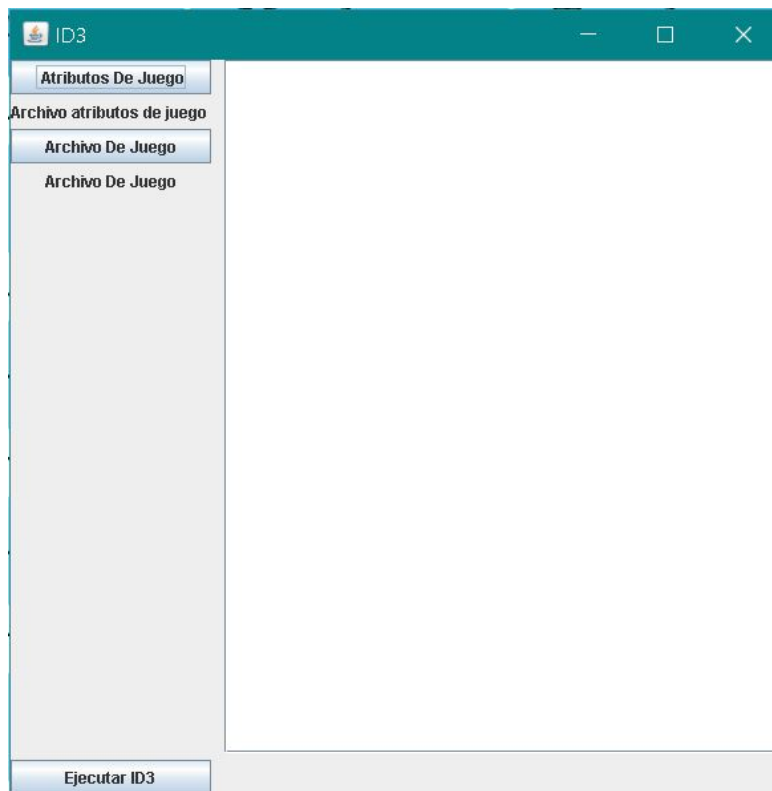


Figura 1: Ventana inicial

Elegimos los archivos en los cuales está la información para poder ejecutar el algoritmo.

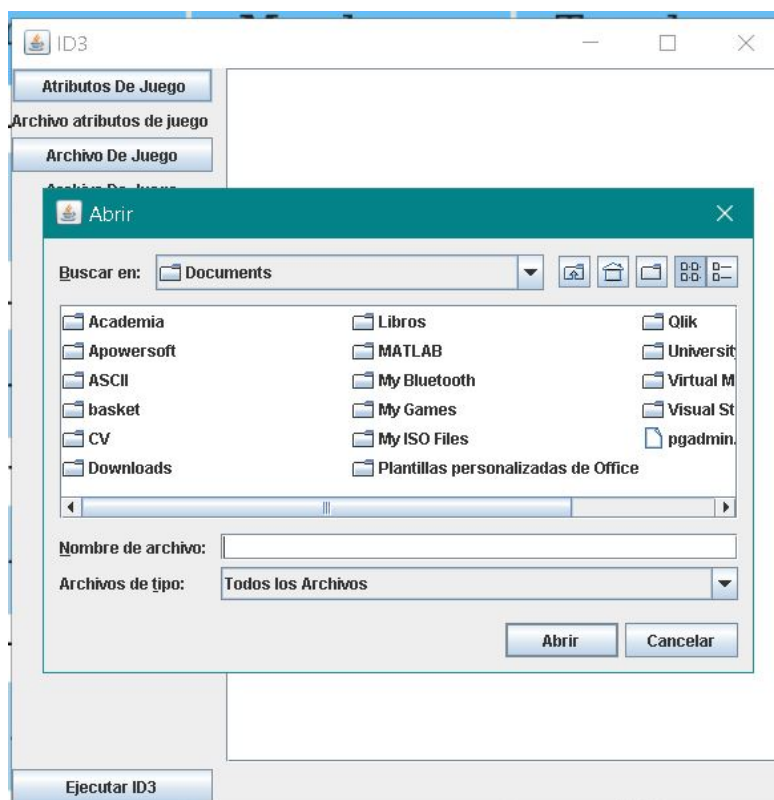


Figura 2: Ventana elección de archivos

Tras elegir ambos archivos, clickeamos el botón de Ejecutar ID3 y en la parte derecha de la ventana podemos ver el resultado tras aplicar el algoritmo. Hemos intentado recrear el árbol que se originaría:

- ->: indica el atributo seleccionado
- = : indica el valor que toma el atributo seleccionado

Si no se llega a seleccionar el valor Jugar, significa que no disponemos de suficiente información para continuar por esa rama.

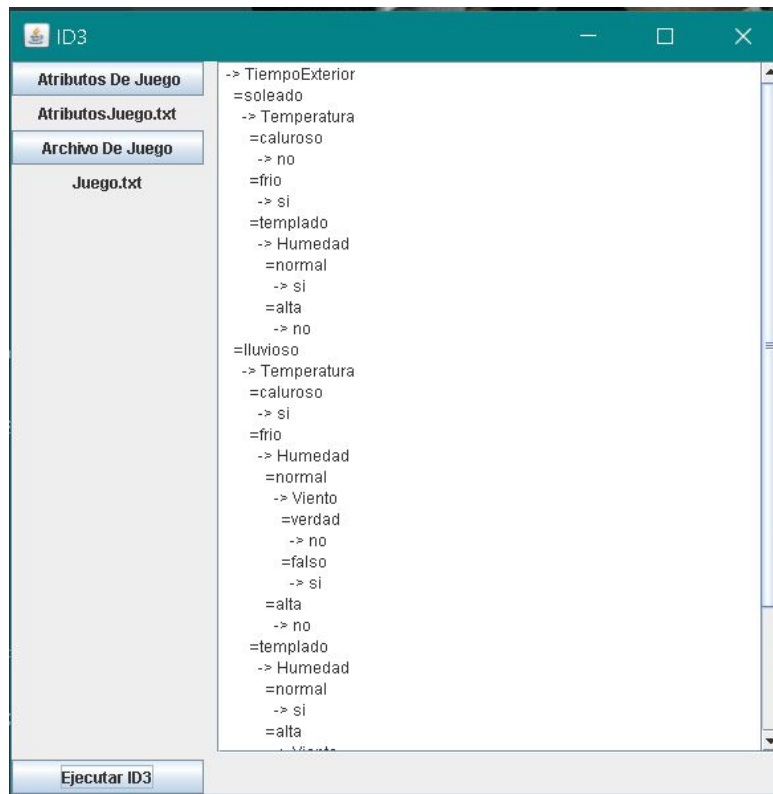


Figura 3: Ventana final

En este ejemplo, el primer atributo seria **TiempoExterior**,

3. Manual de usuario

En el zip adjunto, hay un archivo ejecutable denominado **p2.jar**. Ejecutamos este archivo con doble click.

Una vez aparezca la venta, clickeamos en el botón *Atributos de Juego* para seleccionar el archivo que contiene los atributos del Juego. Después, clickeamos en el segundo botón *Archivo de Juego* para seleccionar el archivo de Juego.

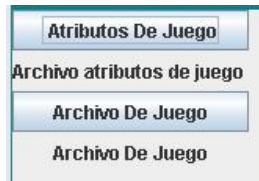


Figura 4: Botón para selección de archivos

Una vez seleccionados ambos archivos, clickeamos el botón de la esquina inferior izquierda *Ejecutar ID3*.



Figura 5: Botón para ejecutar ID3

Tras clickear este botón, en la parte derecha de la ventana nos aparecerán los atributos y sus valores elegidos en cada iteración.

Memoria Práctica 2

Abril 2020

Ult. actualización 26 de abril de 2020

L^AT_EX lic. LPPL & Nerea Jiménez y Yhondri Acosta & CC-ZERO

Esta obra está bajo una licencia Creative Commons “Reconocimiento-NoCommercial-CompartirIgual 3.0 España”.

