

---

## PROCESADOR DE MATRICES EN FORMATO XML

---

201801369 – Brayan Hamllelo Estevem Prado Marroquín

### Resumen

El problema consiste en alojar objetos en las bases de datos en sitios distribuidos, de manera que los costos de la transmisión de datos y su proceso se puedan ser minimizados. Un objeto de base de Datos es una entidad de una Base de Datos.

Para lo cual se requerirá el poder simplificar el procesamiento de los datos para poder simplificar los costos de este. Para lo cual utilizando la lógica de la programación Orientada a Objetos y los Tipos de Datos Abstractos (TDA) y observando los datos como una matriz; se podrá realizar la acción solicitada de forma eficiente y a su vez se podrá reportar los cambios en los datos por medio de archivos con formato XML. Así mismo el poder observar el comportamiento de los datos principales de las Bases de Datos por medio de un reporte en formato de imagen (png); dichas acciones se realizaron en el lenguaje de Programación Python.

### Palabras clave

#### **Abstracción**

*Se basa en poder describir un objeto con propiedades y métodos principales sin pensar en detalle. La abstracción separa el comportamiento específico de un objeto, a esta división que se realiza se le conoce como la barrera de abstracción. (Lara, 2017)*

### Summary

The problem is hosting objects in databases in distributed sites, so that the costs of data transmission and its processing can be minimized. A database object is an entity in a database.

This will require the power to simplify data processing to simplify data costs. For which using object-oriented programming logic and Abstract Data Types (ADT) and observing the data as an array; the requested action can be performed efficiently and in turn changes to the data can be reported using xml-formatted files. Also being able to observe the behavior of the main database data by means of a report in image format (png); these actions were performed in the Python Programming language.

### Keywords

#### **Abstraction**

*It is based on being able to describe an object with main properties and methods without thinking in detail. Abstraction separates the specific behavior of an object, this division that is performed is known as the abstraction barrier.*

### **Clases**

*Una clase (en programación) es una agrupación de datos (variables o campos) y de funciones (métodos) que operan sobre esos datos. A estos datos y funciones pertenecientes a una clase se les denominan variables y métodos o funciones miembro. (Anónimo, s.f.)*

### **Nodo**

*es un componente que forma parte de una red. En otras palabras, tanto si se trata de Internet como de Intranet (utilizada en ámbitos cerrados, con acceso limitado a los usuarios autorizados), cada servidor u ordenador constituye un nodo y se encuentra conectado a otro u otros nodos. (Gardey, 2009)*

### **XML**

*proviene de extensible Markup Language ("Lenguaje de Marcas Extensible"). Se trata de un metalenguaje (un lenguaje que se utiliza para decir algo acerca de otro) extensible de etiquetas que fue desarrollado por el World Wide Web Consortium (W3C), una sociedad mercantil internacional que elabora recomendaciones para la World Wide Web. (Gardey, Definicion.de, 2010)*

### **Matriz**

*En general, una matriz es un conjunto ordenado en una estructura de filas y columnas. Los elementos de este conjunto pueden ser objetos matemáticos, clases, tipos de datos primitivos y Tipos de Datos Abstractos.*

### **Classes**

*A class (in programming) is a grouping of data (variables or fields) and functions (methods) that operate on that data. This data and functions belonging to a class are called variables and member methods or functions. (Anonymous, s.f.)*

### **Node**

*is a component that is part of a network. In other words, whether it is the Internet or Intranet (used in closed scopes, with limited access to authorized users), each server or computer is a node and is connected to another or other nodes. (Gardey, 2009)*

### **XML**

*comes from extensible Markup Language. It is a label-extendable metalanguage (a language used to say something about another) that was developed by the World Wide Web Consortium (W3C), an international trading company that develops recommendations for the World Wide Web. (Gardey, Definicion.de, 2010)*

### **Array**

*In general, an array is an ordered set in a row and column structure. The elements in this set can be mathematical objects, classes, primitive data types, and Abstract Data Types.*

## Introducción

Basándose en el Paradigma de Programación Orientado a Objetos, específicamente en los Tipos de Datos Abstractos (TDA) y de forma particular en este caso se trabajará por medio de Listas Enlazadas las cuales adoptarán el papel de recibir los datos que se estarán enviando desde un archivo XML; dichos datos a su vez serán almacenados en nodos que crean dichas listas enlazadas.

Así mismo se hace la salvedad que en dicho (XML) se están enviando lo que son Matrices con dimensiones  $n \times m$  las cuales serán recibidas en las listas que se mencionan anteriormente. Las cuales tendrán la intención de poder convertir dichos datos en una matriz binaria y a su vez creando una matriz patrón, la cual tiene el fin de poder comparar filas y sumar las mismas que cumplan con el patrón de  $x$  fila. Para final con la escritura de un archivo (XML) enviando los datos de las sumas entrantes.

## Desarrollo del tema

Para poder entrar en contexto se deberá de conocer como temas fundamentales los siguientes:

- a) Programación Orientada a objetos
- b) Tipos de Datos Abstractos
- c) Listas Enlazadas
  - a. Listas Enlazadas Simples
  - b. Lista Doblemente Enlazada
  - c. Listas Circulares
- d) Vectores y Matrices
- e) Que es XML
  - a. Lectura y Escritura de archivos XML

Para lo cual se empezará con:

## PROGRAMACION ORIENTADA A OBJETOS:

Para hablar de la Programación Orientada a Objetos se debe de hablar de primero que es un Paradigma, lo cual se puede mencionar que un Paradigma es un modelo o patrón el cual se refiere a la forma que un programador pueda leer e interpretar el código de distintos programadores, dándole solución a los problemas que se le plantean al mismo.

A su vez se puede mencionar que hay distintos Paradigmas de Lenguajes de programación entre ellos:

- Paradigma Estructural
- Paradigma Funcional
- Paradigma Modular
- Paradigma Lógico
- Paradigma Procedimental
- Paradigma Orientado a Eventos
- Paradigma Orientado a Objetos

Mas en nuestro caso se hablará del Paradigma Orientado a Objetos o también llamado Programación Orientada a Objetos (POO).

La programación orientada a objetos se puede mencionar que se refiere a una manera específica de poder organizar el código por medio de pequeños fragmentos del mismo llamadas *clases*, de las cuales se puede obtener *Objetos*, los cuales podrán interactuar con el resto del código para obtener el funcionamiento adecuado de la aplicación (APP) y/o software (Programa).

La intención de este paradigma es el poder programar de una forma especial, y que se observe como objetos de la vida real; con lo cual se puede aprender a resolver los problemas de maneras distintas a las que se conocían anteriormente en la programación Estructural.

Para poder hablar de dicho paradigma se debe de conocer de igual manera los conceptos de clases y objetos los cuales se describirán a continuación:

- a) *Clases: es una agrupación de datos (variables o campos) y de funciones (métodos) que operan sobre esos datos. A estos datos y funciones pertenecientes a una clase se les denominan variables y métodos o funciones miembro. (Anónimo, s.f.)*
- b) *Objetos: consta de un ente perteneciente al POO la cual consta de datos y tareas que puede realizar en el transcurso del programa; los cuales se puede instanciar a partir de una clase.*

### **TIPOS DE DATOS ABSTRACTOS (TDA)**

Para poder hablar de los tipos de datos abstractos debemos de hacer referencia al concepto de *Abstracción o Encapsulamiento*, el cual es un tema fundamental para dicho de tipos de datos.

- *Abstracción: se puede mencionar que la abstracción o separación de la especificación de un objeto o algoritmo de su implementación, en base a la utilización que se le dará en el programa y solo depende de una interfaz explícitamente definida (la especificación) y no de los detalles de su representación física (la implementación) ya que están ocultos.*

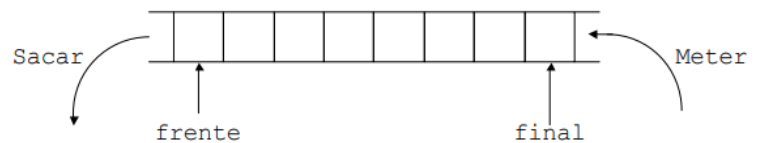
Conociendo el concepto de abstracción se puede mencionar que los Tipos de Datos Abstractos son Tipos de Datos que puede crear el usuario (programador) los cuales se realizan a partir del análisis y la abstracción de los datos y recursos que se manejaran en las mismas, así mismo se puede mencionar varios tipos de TDA'S entre ellas se mencionan:

- Colas
- Pilas
- Listas enlazadas

Se menciona que cada TDA constara de su especificación la cual será independiente de su implementación y a su vez nos proporcionara las operaciones que podrá realizar las cuales nos determinaran su utilización.

#### **COLAS:**

Colección de elementos homogéneos (del mismo tipo: Tipo Elemento) ordenados cronológicamente (por orden de inserción) y en el que sólo se pueden añadir elementos por un extremo (final) y sacarlos sólo por el otro (frente). Es una estructura FIFO (First In First Out):



*Ilustración I. TDA Cola*

Fuente: Departamento de Lenguajes y Ciencias de la Computación.

#### **PILAS:**

Colección de elementos homogéneos (del mismo tipo: Tipo Elemento) ordenados cronológicamente (por orden de inserción) y en el que sólo se pueden añadir y extraer elementos por el mismo extremo, la cabeza. Es una estructura LIFO (Last In First Out):

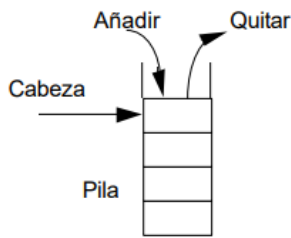


Ilustración II. TDA Pila

Fuente: Departamento de Lenguajes y Ciencias de la Computación.

## LISTAS ENLAZADAS

Colección de elementos homogéneos (del mismo tipo: Tipo Elemento) con una relación LINEAL establecida entre ellos. Pueden estar ordenadas o no con respecto a algún valor de los elementos y se puede acceder a cualquier elemento de la lista.

- Hay que tener en cuenta que la posición de la inserción no se especifica, por lo que dependerá de la implementación. No obstante, cuando la posición de inserción es la misma que la de eliminación, tenemos una subclase de lista denominada pila, y cuando insertamos siempre por un extremo de la lista y eliminamos por el otro tenemos otra subclase de lista denominada cola.
- En el procedimiento Eliminar el argumento elem podría ser una clave, un campo de un registro Tipo Elemento.

### Implementación

La implementación o representación física puede variar:

- Representación secuencial: el orden físico coincide con el orden lógico de la lista.

- Ejemplo: arrays. Para insertar o eliminar elementos podría exigirse que se desplazaran los elementos de modo que no existieran huecos. Por otra parte, tiene la desventaja de tener que dimensionar la estructura global de antemano, problema propio de las estructuras estáticas.



Ilustración III. Ejemplo de Listas

- Representación enlazada: da lugar a la lista enlazada ya estudiada, en la que el orden físico no es necesariamente equivalente al orden lógico, el cual viene determinado por un campo de enlace explícito en cada elemento de la lista. Ventaja: se evitan movimientos de datos al insertar y eliminar elementos de la lista.

Podemos implementar una lista enlazada de elementos mediante variables dinámicas o estáticas (el campo enlace será un puntero o un índice de array):

- Variables de tipo puntero: Esta representación permite dimensionar en tiempo de ejecución (cuando se puede conocer las necesidades de memoria). La implementación de las operaciones es similar a las ya vista sobre la lista enlazada de punteros.



Ilustración IV. Punteros

- Variables estáticas: se utilizan arrays de registros con dos campos: elemento de la lista y campo enlace (de tipo índice del array) que determina el orden lógico de los elementos, indicando la posición del

siguiente elemento de la lista (simulan un puntero). Esta representación es útil cuando el lenguaje de programación no dispone de punteros o cuando tenemos una estimación buena del tamaño total del array.

- a. Ejemplo: Una lista formada por A, B, C, D, E (en este orden) puede representarse mediante el siguiente array de registros:

datos	A		D		B		E	C		
enlace	4		6		7		-1	4		
	0	1	2	3	4	5	6	7	8	9

Ilustración V. composición de Listas

### LISTA DOBLEMENTE ENLAZADA:

Son listas en las que cada nodo, además de contener los datos información propios del nodo, contiene un enlace al nodo anterior y otro al nodo siguiente:

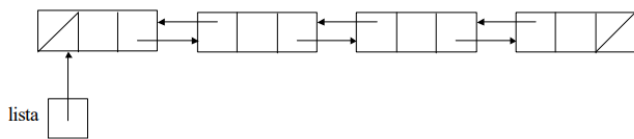


Ilustración VI. Lista Doblemente Enlazada 1

Aparte de que ocupan más memoria (el tamaño de un puntero por cada nodo más), los algoritmos para implementar operaciones para listas dobles son más complicados que para las listas simples porque requieren manejar más punteros. Para insertar un nodo, por ejemplo, habría que cambiar cuatro punteros (o índices si se simulan con un array):

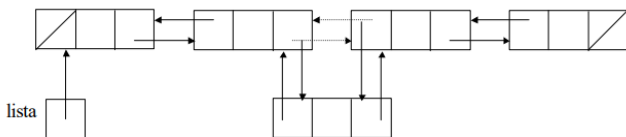


Ilustración VII. Lista Doblemente Enlazada 2

### LISTAS CIRCULARES:

Son listas en las que el último elemento está enlazado con el primero, en lugar de contener el valor NULL o NULO. Evidentemente se implementarán con una representación enlazada (con punteros o variables estáticas). Con punteros sería:

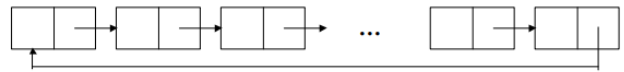


Ilustración VIII. Lista Circular

### UTILIZACION DE LISTAS:

Las listas se utilizan en casi todo tipo de software, especialmente su representación enlazada con punteros. Un ejemplo de aplicación del TAD lista son las tablas hash con encadenamiento de sinónimos. La tabla sería un array de listas de longitud variable: un nodo por cada colisión, lo que evitaría sobredimensionar la tabla y además su posible desborde:

```
/* TIPOS */
typedef ¿? TipoElemento /* un registro p.
ej. */
TipoLista TablaHash[NmEmpl];
```

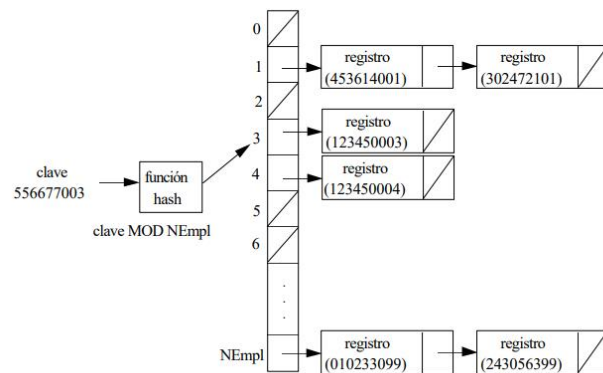


Ilustración IX. Ejemplo de la utilización de Listas

### VECTORES Y MATRICES

En programación, una matriz o vector (llamados en inglés arrays) es una zona de almacenamiento continuo, que contiene una serie de elementos del mismo tipo, los elementos de la matriz. Desde el

punto de vista lógico una matriz se puede ver como un conjunto de elementos ordenados en fila (o filas y columnas si tuviera dos dimensiones).

En principio, se puede considerar que todas las matrices son de una dimensión, la dimensión principal, pero los elementos de dicha fila pueden ser a su vez matrices (un proceso que puede ser recursivo), lo que nos permite hablar de la existencia de matrices multidimensionales, aunque las más fáciles de imaginar son los de una, dos y tres dimensiones.

		Matriz						Vector	
		y							
		0	1	2	3	4	5		
x	0	10	10	10	10	10	10	"Ricardo"	
	1	8	8	7	8	9	10	"Fernando"	
	2	6	7	7	8	9	10	"Cecilia"	
	3	9	10	9	10	9	10	"Martha"	
	4								
	5								
		Calificación[ , ]						Nombre[ ]	

Ilustración X. Vectores y Matrices

## QUE ES XML

XML es un subconjunto de SGML (Estándar Generalised Mark-up Language), simplificado y adaptado a Internet

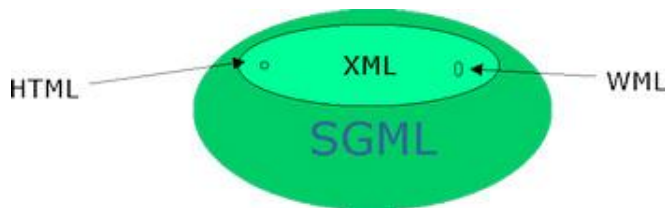


Ilustración XI. SGML

- XML no es, como su nombre puede sugerir, un lenguaje de marcado.

- XML es un metalenguaje que nos permite definir lenguajes de marcado adecuados a usos determinados.

## CARACTERISTICAS

- XML es un subconjunto de SGML que incorpora las tres características más importantes de este:
  - Extensibilidad
  - Estructura
  - Validación
- Basado en texto.
- Orientado a los contenidos no presentación.
- Las etiquetas se definen para crear los documentos, no tienen un significado preestablecido.
- No es sustituto de HTML.
- No existe un visor genérico de XML.

## ESTRUCTURA DE UN DOCUMENTO XML

Un documento XML está formado por datos de caracteres y marcado, el marcado lo forman las etiquetas:

```

Prologo { <?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>
          <!DOCTYPE persona SYSTEM "persona.dtd"
Cuerpo { <persona>
          <nombre>Luis</nombre>
          <apellidos>Pérez</apellidos>
          </persona>
    
```

Ilustración XII. Estructura de un Archivo XML

## ESTRUCTURA DE ETIQUETAS

```

Nombre del atributo  Contenido del elemento
<autor país = "España"> Jose Ramón </autor>
Nombre de elemento  Valor de atributo  Etiqueta de fin
    
```

Ilustración XIII. Estructura de Etiquetas en XML

## COMPONENTES DE UN DOCUMENTO XML

- En un documento XML existen los siguientes componentes:
  - Elementos: Pieza lógica del marcado, se representa con una cadena de texto(dato) encerrada entre etiquetas. Pueden existir elementos
  - vacíos (<br/>). Los elementos pueden contener atributos.
  - Instrucciones: Ordenes especiales para ser utilizadas por la aplicación que procesa
  - <?xml-stylesheet type= "text/css" href= "estilo.css">
  - Las instrucciones XML. ¿Comienzan por <? ¿Y terminan por?>.
  - Comentarios: Información que no forma parte del documento. ¿Comienzan por <! -- y terminan por -->.
  - Declaraciones de tipo: Especifican información acerca del documento:
  - <!DOCTYPE persona SYSTEM "persona.dtd">
  - Secciones CDATA: Se trata de un conjunto de caracteres que no deben ser interpretados por el procesador:
  - <![CDATA [ Aquí se puede meter cualquier carácter, como <, &, >, ... Sin que sean interpretados como marcación]]>

## Conclusiones

- A. Se presento el análisis de los temas utilizados para el proyecto en función, con el fin de tener una orientación para la ejecución de este.
- B. Se conoció los conceptos básicos de la Programación Orientación a Objetos con el fin de utilizar de forma adecuada las TDA en la operación del proyecto.
- C. La utilización de documentos XML como base para la estructuración de datos es de utilidad ya que a partir de cada etiqueta se podrá obtener diversos datos que se utilizará en el mismo proceso.

## Referencias

- Anonimo. (s.f.). Obtenido de [https://www.ecured.cu/Clases\\_\(programaci%C3%B3n\)#:~:text=Una%20clase%20%28en%20programaci%C3%B3n%29%20es%20una%20agrupaci%C3%B3n%20de,Objetos%20se%20basa%20en%20la%20programaci%C3%B3n%20de%20clases](https://www.ecured.cu/Clases_(programaci%C3%B3n)#:~:text=Una%20clase%20%28en%20programaci%C3%B3n%29%20es%20una%20agrupaci%C3%B3n%20de,Objetos%20se%20basa%20en%20la%20programaci%C3%B3n%20de%20clases).
- Gardey, J. P. (2009). *definicion.de*. Obtenido de <https://definicion.de/nodo/>
- Gardey, J. P. (2010). *Definicion.de*. Obtenido de <https://definicion.de/xml/>
- Lara, D. (6 de julio de 2017). *Laraveles*. Obtenido de <https://laraveles.com/series/poo/la-abstraccion-programacion-orientada-objetos/>