

Laboratorio de Software

Práctica nº 5

Temas

- Arreglos de primitivos y objetos.
- Clases de la API para manejar colecciones de objetos
- Genéricos

1.- Se quiere mantener un conjunto de alumnos ordenados por número de legajo, correspondientes a una materia en una Facultad. Defina una clase **Alumno** en el paquete **practica5**, que contenga información del alumno, entre otras cosas el legajo, apellido y nombre y DNI.

- Defina una clase llamada **Materia** en el mismo paquete que mantenga información sobre la materia y la nómina de alumnos. Use colecciones genéricas.
- Defina en la clase **Materia** 2 métodos: **agregaAlumno(Alumno a)** que agregue a alumnos a la nómina de alumnos existentes y **getAlumnos()** que retorne el conjunto de alumnos inscriptos en la materia.
- Pruebe las clases definidas creando agregando al menos 10 alumnos en una **Materia**. Recorra la colección e imprima sus elementos.
- Su colección está ordenada, ¿qué requisito tuvo que cumplir para poder mantenerla ordenada?
- Si quisiera que el criterio de orden se mantenga sobre el apellido y nombre ¿cómo lo haría?
- Analice una solución que permita mantener una nómina no sólo de alumnos sino también de empleados, clientes, etc. ¿Cómo definiría la colección?

2.- Considere la siguiente clase:

```
public class Veterinaria<E> {
    private E animal;

    public void setAnimal(E x) {
        animal = x;
    }
    public E getAnimal() {
        return animal;
    }
}

public class Animal{
}

public class Gato extends Animal {
}

public class Perro extends Animal {
}
```

a) Indicar cual es el resultado de las siguientes operaciones:

- i) `Veterinaria <Animal> vet = new Veterinaria <Gato>();`
- ii) `Veterinaria <Gato> vet = new Veterinaria <Animal>();`
- iii) `Veterinaria <?> vet = new Veterinaria<Gato>();`
`vet.setAnimal(new Gato());`
- iv) `Veterinaria vet = new Veterinaria ();`
`vet.setAnimal(new Perro());`
- v) `Veterinaria vet = new Veterinaria <?>();`
- vi) `Veterinaria <? extends Animal> vet = new Veterinaria<Gato>();`

3.- Se desea recuperar colecciones de palabras similares y para ello se deben implementar un motor de comparación basado en una técnica particular de **string matching**. Esta técnica define una función de similitud entre dos cadenas que indicará cuán similares son dichas cadenas.

Las técnicas utilizadas por estos tipos de motores de comparación están basadas en *tokenización* (descomposición en unidades básicas), y como resultado del proceso obtendremos un conjunto de *tokens*, o lo que es lo mismo, de Q-gramas (sub-cadenas de tamaño Q). La siguiente figura muestra un ejemplo de *tokenización* en gramas de tamaño 2:

Cadena original c	Posición de la ventana	Q-Grama extraído	Lista de Q-gramas
Glucosa	[#G]lucosa\$	[#G]	[#G]
	#[G]lucosa\$	[G]	[#G][G]
	#G[lu]cosa\$	[lu]	[#G][G][lu]
	#G[uc]osa\$	[uc]	[#G][G][lu][uc]
	#Glu[co]sa\$	[co]	[#G][G][lu][uc][co]
	#Gluc[os]a\$	[os]	[#G][G][lu][uc][co]
	#Gluco[sa]\$	[sa]	[#G][G][lu][uc][co][sa]
	#Glucos[a\$]	[a\$]	[#G][G][lu][uc][co][sa][a\$]

Figura 1 - Ejemplo del proceso de tokenización: Dada la cadena "Glucosa" se introducen caracteres de inicio y de final de cadena (# y \$ símbolos no existentes en el alfabeto utilizado) y se obtiene una lista de Q-gramas mediante el uso de una ventana de tamaño 2 que se desliza a través de los caracteres de la cadena.

Usted debe:

- a) Implementar un diccionario de gramas que almacenará el resultado de la factorización de palabras en Q-gramas, con Q=2, donde las claves del diccionario serán las gramas y el valor almacenado en el mismo serán las listas de palabras que contiene el grama de la clave.
- b) Implementar un motor de comparación basado en la siguiente técnica de **string matching**:
 - o **Distancia de Levenshtein menor a tres:** Las cadenas comparadas deben poseer una distancia de Levenshtein menor a tres. La distancia de Levenshtein, distancia de edición o distancia entre palabras es el número mínimo de operaciones requeridas para transformar una cadena de

caracteres en otra. Por ejemplo, la distancia de Levenshtein entre "casa" y "calle" es de 3 porque se necesitan al menos tres ediciones elementales para cambiar uno en el otro.

(http://es.wikipedia.org/wiki/Distancia_de_Levenshtein).

El motor de comparación definirá un método que recibirá como parámetro un diccionario de gramas de tamaño 2 y una cadena. Dicho método dividirá la cadena en gramas de tamaño 2, explorará recolectando las palabras almacenadas en el diccionario de gramas que contenga alguno de los gramas de la cadena enviada como parámetro y retornará como resultado la colección de palabras que cumplen con el criterio de comparación.