



VNIVERSITAT ID VALÈNCIA

Laboratorio de ESTRUCTURAS DE DATOS Y ALGORITMOS

Grado en Ciencia de Datos (1º)
Curso 2024-25

Práctica Nº 1: Algoritmos de Búsqueda

Periodo de realización: Semana del 10 a 14/02/2025

Metodología de trabajo

Fases	Tipo	Dedicación
1. Resolver las tareas planteadas en este guion <u>antes</u> de iniciar la sesión presencial en el laboratorio.	No presencial	Máx. 4,5 h.
2. Resolver en el laboratorio un <u>nuevo ejercicio</u> basado en la resolución de las tareas previas.	Presencial	Máx. 3 h.

Introducción

La búsqueda es un proceso de análisis de la información ampliamente utilizado, en especial cuando se dispone de un volumen elevado de datos sobre un determinado tema (censo de habitantes de una ciudad, datos sanitarios de la población, actividades de una red social, etc.). Es importante que el algoritmo empleado sea rápido, ya que es una tarea que se puede ejecutar muy frecuentemente (miles/millones de veces) y sobre colecciones de datos muy grandes (miles/millones de registros).

En esta práctica vamos a abordar el cálculo empírico/experimental de costes de diversos algoritmos de búsqueda. En concreto, se pretende obtener resultados experimentales del coste de la implementación de los algoritmos típicos de búsqueda sobre vectores/listas: (1) Secuencial (sin parada), (2) Secuencial con parada, (3) Secuencial con centinela y (4) Búsqueda binaria.

A modo de ejemplo, trabajaremos con una lista de personas a las que se les ha administrado dosis de la vacuna contra Covid-19¹ y se medirá (solo) el número total de pasos ejecutados dentro de los bucles de búsqueda de cada uno de los algoritmos, tanto si la búsqueda finaliza con éxito (se localiza el elemento buscado) o con fracaso (no se localiza el elemento).

Para realizar los ejercicios se usará como material inicial el programa "**pr1_v0.py**" y los módulos auxiliares "**pr1_persona.py**" y "**pr1_algoritmos.py**", que se pueden descargar desde el Aula Virtual de la asignatura. Estos archivos incluyen la definición del tipo de datos a utilizar (*PersonaVacunada*), el mecanismo de lectura de datos desde un archivo externo y la versión preliminar de los algoritmos de búsqueda que se desean comparar. A partir de aquí se debe:

1. Leer el código proporcionado e interpretar su significado y funcionamiento.
2. Realizar las tareas indicadas en los siguientes ejercicios, como paso previo al trabajo a realizar en la sesión de laboratorio.

¹ Estos datos son evidentemente falsos, pues los datos personales de salud están catalogados con nivel de seguridad alto y no son públicos.

Ejercicios (Fase 1, previa al laboratorio)

Tarea 1

Se debe modificar el programa "**pr1_v0.py**" (cambiale el nombre por "**pr1_v1.py**") para que realice las siguientes tareas (indicadas en el mismo archivo dentro de la función `main`):

1. Leer y cargar en una lista la información de personas vacunadas disponible en el archivo "**vacunaciones_small.csv**".
2. Mostrar por pantalla los datos cargados en la lista.

Verifica que los datos mostrados corresponden con los disponibles en el archivo.

Tarea 2

Ahora debes añadir al programa el proceso de búsqueda en la lista de las personas con los siguientes identificadores: 0002JTX, 1111VAC, 0006HYR, 1543GYM y 0014FGV. Para cada identificador se deben aplicar los 4 métodos de búsqueda y mostrar sus resultados en pantalla: (1) el identificador buscado, (2) si ha sido encontrado o no y (3) cuántos pasos se han realizado con cada método de búsqueda utilizado².

Para esta tarea será necesario importar desde el programa el módulo "**pr1_algoritmos**", pero también modificar previamente las funciones de búsqueda contenidas en este módulo, ya que en su estado actual comparan elementos completos de la lista (todo el registro). Sin embargo, lo que nos interesa ahora es buscar un identificador concreto y, por lo tanto, se deben buscar personas considerando como clave de búsqueda (exclusivamente) el campo **id** y no todo el registro. Por eso, comparaciones del estilo `v[i]==x` no tendrán sentido puesto que `v[i]` será un registro (*PersonaVacunada*) y `x` será un identificador. Son este tipo de comparaciones las que deben ser adaptadas al problema concreto que se debe resolver. Además, debes prestar especial atención a la definición del centinela empleado en la función *BusquedaSecuencialCentinela*, que deberá ser una *PersonaVacunada* con el **id** buscado (el resto de los campos no importan). Por ese motivo, desde este archivo también deberás importar el módulo "**pr1_persona**".

Verifica que los resultados obtenidos para los 5 identificadores de prueba son correctos, de acuerdo con el contenido del archivo.

Tarea 3

Realiza una nueva modificación de las funciones de búsqueda del módulo "**pr1_algoritmos**" para que devuelvan la posición en la que se encuentra el elemento, en lugar del valor booleano. Si el elemento no se encuentra en la lista las funciones deberán devolver el valor **-1**. Modifica también el programa para que cuando se encuentre un identificador en el catálogo de vacunaciones, se muestre la posición en la que ha sido localizado.

Verifica los resultados usando el banco de prueba proporcionado al final de este documento.

² Fíjate que las funciones proporcionadas ya calculan y devuelven el número de pasos ejecutados dentro de los bucles de búsqueda de cada uno de los algoritmos, tal como se ha indicado en la introducción.

Tarea 4

Revisa toda la documentación del programa y de los módulos que estás utilizando para que todos los comentarios incluidos en los archivos sean coherentes y reflejen la realidad de las modificaciones realizadas hasta ahora, incluida la autoría y fecha de modificación de los archivos.

Banco de pruebas. Verificación de resultados

Clave	Posición	Pasos			
		B.Secuencial	B.Sec.Parada	B.Centinela	B.Binaria
0002JTX	2	31	6	4	10
1111VAC	-1	30	30	20	20
0006HYR	5	31	15	10	10
1543GYM	-1	30	30	20	20
0014FGV	9	31	27	18	15

Las diferencias entre algoritmos no son, en absoluto, significativas. Fundamentalmente porque se trata de una prueba con un número muy pequeño de elementos en la lista (solo 10).

Comentario final

No se requiere entregar para su evaluación el trabajo realizado en esta fase 1, aunque el profesorado puede solicitar revisarlo durante el desarrollo de la sesión de laboratorio. También se debe tener en cuenta que este trabajo será necesario para completar la fase 2 de la práctica durante la sesión de laboratorio, que sí será evaluable.

Fase 2: Tarea final

Resolución de un ejercicio sobre el mismo problema planteado durante la sesión en el laboratorio.