

# Estructures de dades

## Curs 2020-21

### Práctica 2 - Códigos de Huffman - Sprint 1

10 d'abril de 2021

Esta semana empezaremos con la implementación de la práctica 2 construyendo la **tabla de frecuencias** a partir del texto contenido en el archivo *entrada.txt*. La tabla de frecuencias resultante debe imprimirse dentro de otro archivo de texto llamado *entrada\_freq.txt*.

Por ejemplo, si el contenido del archivo *entrada.txt* es:

```
1 això es un exemple darbre de huffman
```

Una vez finalizada la ejecución, el archivo *entrada\_freq.txt* debería contener:

```
1  :   6
2  a:   3
3  b:   1
4  d:   2
5  e:   6
6  f:   2
7  h:   1
8  i:   1
9  l:   1
10 m:   2
11 n:   2
12 o:   1
13 p:   1
14 r:   2
15 s:   1
16 u:   2
17 x:   2
```

# 1 Tareas a realizar

Las tareas a realizar durante esta semana son:

1. Implementar un **mapping** utilizando la implementación **array indexado por claves**.
  - La clave será el carácter.
  - El valor será la frecuencia observada del carácter en el texto contenido en el archivo *entrada.txt*.
2. Para cada uno de los caracteres contenidos en el fichero *entrada.txt*:
  - En el caso de que el carácter no exista dentro del conjunto, hay que poner la llave dentro del conjunto con un valor = 1.
  - En el caso de que el carácter ya exista dentro del conjunto, hay que actualizar el valor correspondiente a la clave incrementándolo en una unidad.
  - Para determinar si un carácter (clave) existe dentro del conjunto, debe realizar una **búsqueda** mediante un **iterador**.
3. Una vez todos los caracteres contenidos en el fichero *entrada.txt* han sido procesados, se debe **recorrer** el conjunto mediante un **iterador** y escribir todas las parejas clave-valor en el fichero *entrada-freq.txt* (una pareja por línea).

## 2 Instancia genérica del TAD mapping utilizando la implementación array indexado por claves

El TAD mapping utilizando la implementación array indexado por claves tiene 2 elementos genéricos:

```
1 generic
2   -- tipo discreto
3   type key is (<>);
4   type item is private;
5 package d_conjunto is
```

Recordad que para poder utilizar el TAD, debemos realizar la instancia genérica del paquete *d\_conjunto* concretando cada uno de sus elementos genéricos (key, item).

Para esta implementación (array indexado por claves), el tipo *key* debe corresponder a un tipo discreto.

Para esta práctica, consideraremos que el archivo de entrada puede contener cualquier carácter alfanumérico contenido entre el **espacio** en blanco y la **letra minúscula z**. Por lo tanto, puede utilizarse la siguiente instancia genérica:

```
1 subtype alfabet is Character range ' '.. 'z';  
2  
3 package d_taula_frequencies is new d_conjunto(key => alfabet,  
    item => Integer);
```

Recordad que la implementación **array indexado por claves** es útil cuando el *mapping* o *conjunto* contiene muchos de los elementos que puede contener.