

GRAU EN ENGINYERIA DE SISTEMES AUDIOVISUALS

ESEIAAT

ESTRUCTURA DE DADES I ORIENTACIÓ A OBJECTES

20/04/2021, 18h

EXÀMEN PARCIAL

LLEGIU ATENTAMENT AQUESTA INFORMACIÓ

1. SOBRE LA CÒPIA: En cas d'indicis de còpia:

- L'alumne/els alumnes que han copiat i els que han facilitat la còpia tindran un zero en el control
- S'informarà als Estudis de l'ESEIAAT sobre el cas per tal que prenguin mesures disciplinàries i que consti en l'expedient

2. Sobre solucions que no es corresponen amb els continguts de l'assignatura:

- En cas d'indicis de continguts en la solució que no es corresponen amb els continguts de l'assignatura, per l'ajuda de tercers, es demanarà una entrevista amb el(s) implicat(s) i si els indicis es confirmen tindran un zero en el control. Igualment s'informarà als Estudis de l'ESEIAAT pel cas.

3. No s'acceptaran lliuraments fora del termini establert en les tasques corresponents habilitades a ATENEA. No s'acceptaran en cap cas lliuraments per email o altres vies.

ENUNCIAT –PROBLEMA 1

PHASE 1: Lliureu la solució del problema 1 (només els fitxers font del programa) a la tasca corresponent dins el termini.

[5 punts, 1h:15min]

Es vol implementar un programa que permet detectar la similitud entre dos texts. Aquest tipus de programa són particularment útils per detectar plagi en els texts.

Implementeu una classe **TextSimil** que manté les següents dades:

- Dos texts (com vectors de paraules)
- Altres dades que pugueu necessitar per implementar l'API de la classe

L'API de la classe ha d'incloure:

- Constructors: per defecte, amb paràmetres (cal llegir els texts des de dos fitxers de text) i el de còpia.
- **Mètodes consultors:**
 - a) Calcular el vector de les paraules **comunes** en els dos texts.
 - b) Consultar el nombre de paraules comunes en els dos texts.
 - c) Calcular l'índex de similitud entre els dos texts:

$$\frac{\text{Nombre de les paraules comunes en els dos texts}}{\text{Nombre total de les paraules diferents en els dos texts}}$$

- **Mètodes modificadors:** Eliminar paraules consecutives repetides d'un text.
- **Operadors:**
 - a) <= (cert, quan l'índex de similitud de un objecte TS1 de la classe TextSimil és <= que el d'un objecte TS2 de la classe TextSimil).
 - b) << per escriure un text a un fitxer de sortida (haver eliminat les paraules consecutives repetides).

Nota1: podeu fer la solució amb .h i .cpp o amb .cpp directament.

Nota2: no cal que lliureu el programa main però se us recomana que el feu per provar la vostra solució.

Exemple d'execució:

Per dos fitxers de text

Hola bon bon dia com va el divendres dia dia número cinc de de la setmana

i

Hola avui és és divendres el darrer dia dia laboral de la setmana amb vista al cap de setmana per descansar i trobar-se amb amics

el resultat seria:

txt1: té 16 paraules

txt2: té 25 paraules

txt1 netejat: té 13 paraules

txt2 netejat: té 23 paraules

Report de similitud:

Nombre de paraules en comú: 10

Index similitud: 62.5%