

## Prova Finale – Livello Avanzato Compito A1

---

COGNOME E NOME: \_\_\_\_\_ MATRICOLA: \_\_\_\_\_

**Tempo a disposizione: 4 ore**

### Clustering in Java

E' necessario scrivere una libreria che fornisce i principali metodi a supporto di un semplice algoritmo di "clustering". Dato un insieme di punti nello spazio (per semplicità consideriamo lo spazio a due dimensioni), si dice "cluster" un gruppo di punti particolarmente "vicini" tra di loro. Si chiama "clustering" il processo secondo il quale i punti vengono raggruppati per costruire i cluster. In questo esempio utilizziamo come misura della distanza tra i punti la tradizionale distanza euclidea (ovvero, dati i due punti  $(x_1, y_1)$  e  $(x_2, y_2)$ , la radice quadrata di  $(x_1 - x_2)^2 + (y_1 - y_2)^2$ ). La distanza tra due cluster C1 e C2 (ovvero tra due gruppi di punti) corrisponde alla distanza tra i loro baricentri. Si definisce "baricentro" di un cluster il punto dello spazio le cui coordinate sono ottenute dalla media delle corrispondenti coordinate dei punti del cluster (ovvero, dato un cluster che contiene i punti  $(x_1, y_1)$ ,  $(x_2, y_2)$  e  $(x_3, y_3)$ , il baricentro avrà coordinate  $(x_1 + x_2 + x_3)/3$  e  $(y_1 + y_2 + y_3)/3$ ).

La libreria deve fornire i seguenti servizi:

#### “Calcolo della distanza tra punti”

- dati due punti, la libreria deve calcolare e restituire la loro distanza (ovvero la distanza euclidea)

#### “Calcolo del baricentro di un cluster”

- dato un cluster, la libreria deve calcolare e restituire il baricentro del cluster

#### “Calcolo della distanza tra cluster”

- dati due cluster, la libreria deve calcolare e restituire la loro distanza (ovvero la distanza tra i rispettivi baricentri)

#### “Fusione di cluster”

- dati due cluster, la libreria deve generare un nuovo cluster ottenuto fondendo i punti del primo e del secondo cluster

#### “Fusione dei cluster più vicini”

- data una collezione di cluster, la libreria deve (a) trovare i due cluster più vicini, ovvero quelli con la distanza minima; (b) eliminare i due cluster dalla collezione (c) aggiungere alla collezione un nuovo cluster ottenuto dalla fusione dei due precedenti

Scrivere l'applicazione che effettua le operazioni elencate sopra secondo le seguenti specifiche:

- Sviluppare il package che implementa la libreria. **NON** è necessario sviluppare interfaccia e controllo
- Sviluppare un **piano completo** di test di unità per i componenti della libreria
- Utilizzare uno stile di programmazione difensiva nello sviluppo delle classi
- Utilizzare il sistema di logging fornito dalla piattaforma per il debugging del codice