

**Knowledge Discovery in Databases II**  
SoSe 2008

**Übungsblatt 8: Multirepräsentiertes Clustering und Minimal Matching Distanz**

Besprechung am 20.6.2008 und 23.6.2008

**Aufgabe 8-1**     *Multirepräsentiertes Clustering*

Gegeben sei ein Datensatz  $X$ , so dass jeder Punkt durch 2 zweidimensionale Vektoren repräsentiert wird.

$$\begin{aligned} A &= (0, 1); (3, 0) \\ B &= (-1, -1); (2, 0) \\ C &= (0, 0); (3, 1) \\ D &= (0, -3); (-2, 2) \\ E &= (2, 1); (-2, -3) \end{aligned}$$

Wir wollen auf diesem Datensatz multirepräsentiertes Clustering mittels DBSCAN durchführen.

- (a) Wie unterscheidet sich multirepräsentiertes Clustering von gewöhnlichem Clustering? Welche besonderen Schwierigkeiten sind damit verbunden?
- (b) Es sei  $MinPoints = 3$ . Für welche Werte von  $\epsilon_1, \epsilon_2$  sind die Objekte C und D Kernobjekte nach
  - der Vereinigungsmethode?
  - der Schnittmethode?

**Aufgabe 8-2**     *Kuhn-Munkres Algorithmus*

Gegeben sei folgende Kostenmatrix  $K$ . Führen Sie ein Hungarian Matching auf dieser Matrix durch.

$$K = \begin{bmatrix} 2 & 1 & 3 \\ 6 & 5 & 4 \\ 7 & 8 & 9 \end{bmatrix};$$