

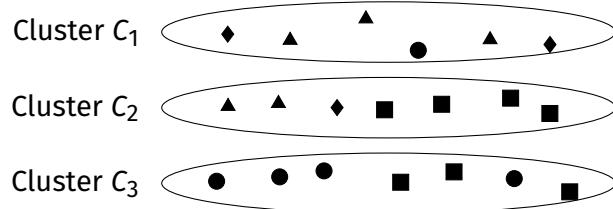
**Assignment 9: Cluster Evaluation and SCHISM**  
Due: Thursday, 30.6.2022

---

**Problem 9-1 Evaluation of Clusterings**

**8**

Given the following clustering into clusters  $C_1$  to  $C_3$ , symbols indicate classes.



- a) Compute the cross-tabulation matrix, and its marginals. *Label* your rows and columns!
- b) Compute the number of pairs in each intersection.
- c) Compute the true positives, false positives, true negatives and false negatives.
- d) Compute the Jaccard and Rand indexes.
- e) Compute the expected value of the Rand index, and the Adjusted Rand Index (ARI).

**Problem 9-2 SCHISM**

**6**

Compare CLIQUE and SCHISM:

- a) What is the difference in their density threshold definition?
- b) Explain the advantage of a variable threshold for subspace clustering.
- c) Are there also disadvantages? Explain how they possibly affect the result.