

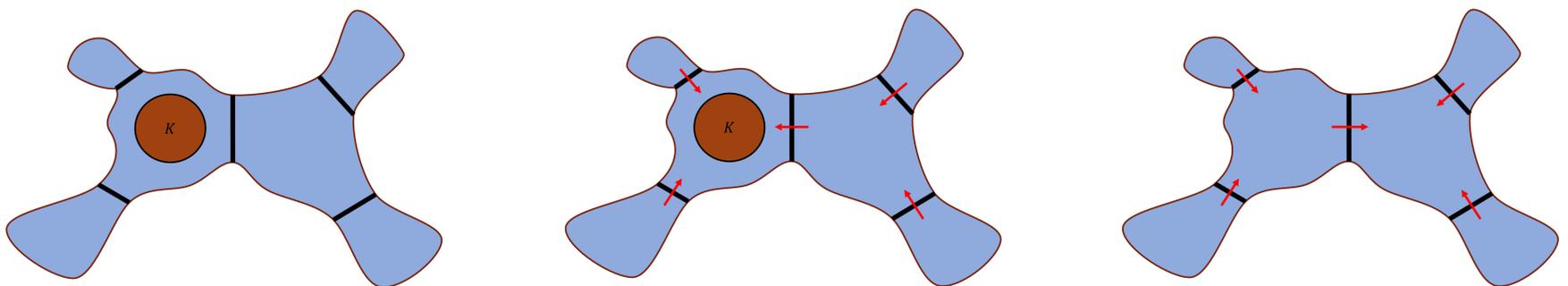
Bachelor Projects in Algorithms

Section for Algorithms, Logic and Graphs

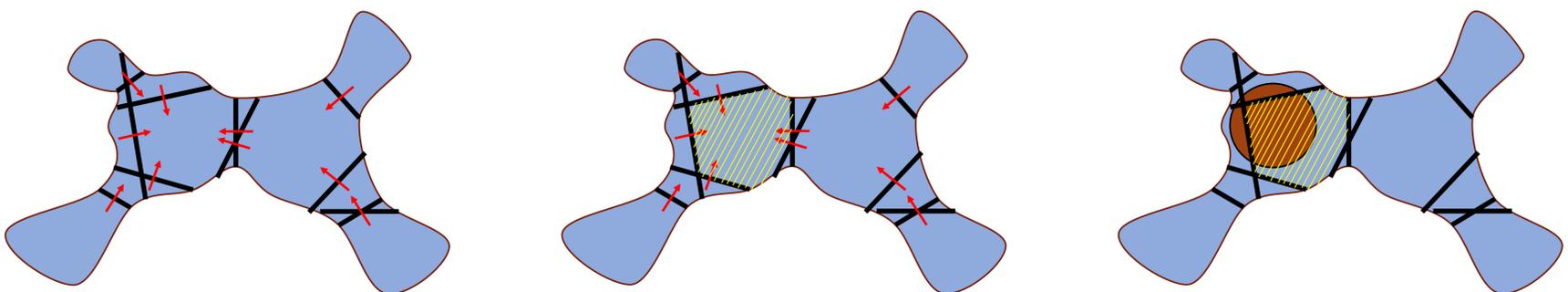
DTU Compute, November 2022

A new graph-theoretic approach for hierarchical clustering: Tangles

- **Tangles** provide:
 - a way to **identify / find clusters**.
 - a way to **hierarchically** cluster data sets.
 - **(soft) dendrograms**.
- Originally *tangles* are objects from structural Graph Theory.
 - Tangles are witnesses for highly connected substructures in graphs.
 - They *consistently* orient each cheap *cut / separation* of a graph.



- Tangles can be defined for **arbitrary data sets**, based on (consistent) ways to **cut the data set**.
- They provide more *fuzzy* information about clusters, e.g. where **most of the cluster** is located.



Potential project components:

- Implement and test different versions of a tangle clustering algorithm from the literature, e.g. w.r.t overlapping clusters / communities.
- Find and test heuristics for generating initial cuts to start the clustering.
- Analyse and compare performance of the tangle algorithm with other clustering algorithms.
- Enhance the algorithm w.r.t. e.g. type, shape or amount of data.
- Develop a software framework for testing different versions of the tangle algorithm.
- Produce a software library for the tangle algorithm

Background:

- Basic knowledge in algorithms (required)
- Clustering and Data Analysis (desired)
- Graph Theory (helpful)