Visualise undrawable Euler diagrams

12th International Conference on Information Visualisation (IV 2008)

Authors: Paolo Simonetto
David Auber

Laboratoire Bordelais
de Recherche en Informatique

University of Bordeaux 1

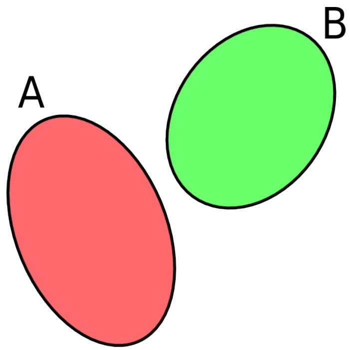
Euler Diagrams

Euler diagrams are plots often used to represent sets and their reciprocal relationships.

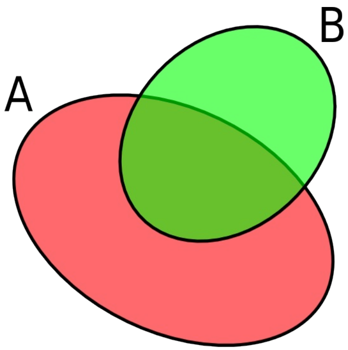
Closed lines are used to identify regions of the plan. Each region is associated to a set.

Overlap, inclusion and exclusion of the regions are used to depict the analogue set concepts.

Euler diagrams must represent only not-null intersections through the intersection of these regions.

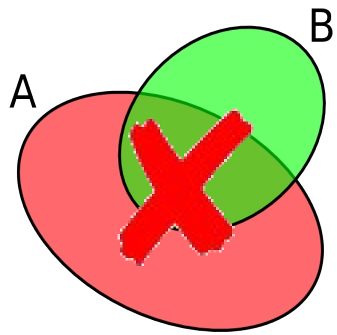


$$A \cap B = \emptyset$$



$$A \cap B \neq \emptyset$$

$$A \cap B = \emptyset$$



Euler Diagrams – General definitions

We will call the sets we aim to represent classes.

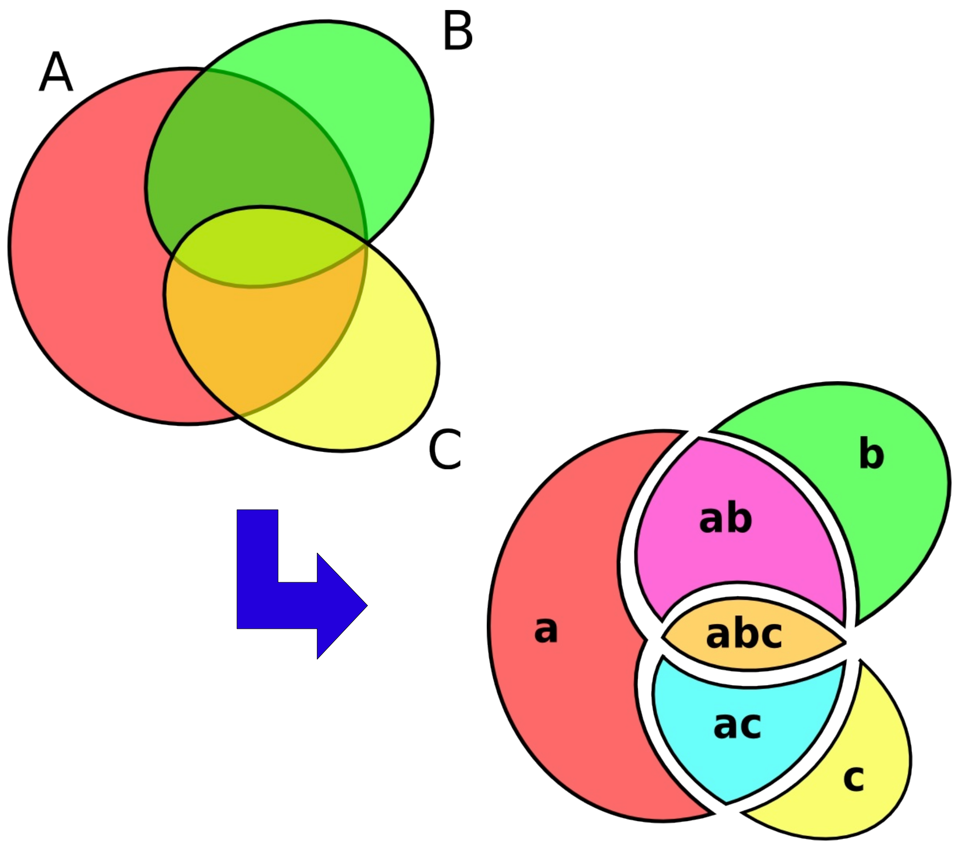
The way classes overlap allows identification of zones, according to the formula:

$$Z_S = \left(\bigcap_{C_x \in S} C_x \right) \cap \left(\bigcap_{C_x \notin S} \overline{C_x} \right)$$

where S is a subset of classes.

Each zone enclose all the, and only the, elements that are contained exactly in the same subset of classes.

We can see graphically what zones are, supposing we already have the final diagram:



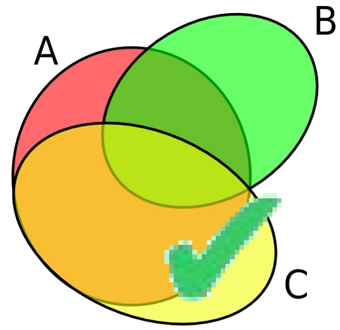
Euler Diagrams – Diagram definitions

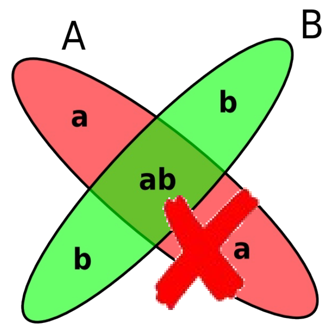
There is no complete agreement on the definition of Euler diagrams.

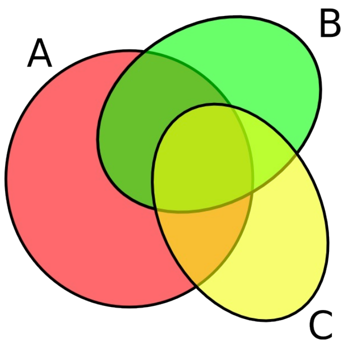
The most accepted definition states that in a well formed Euler diagram each zone is a single connected region.

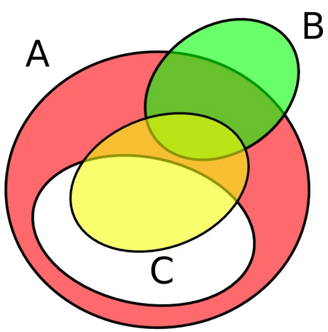
Some definitions are more restricting, other are more permissive.

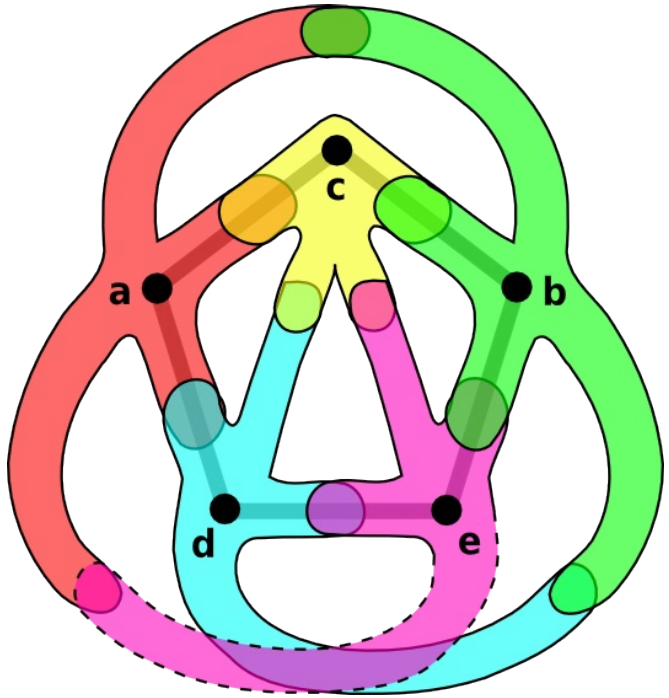
No matter how allowing the definition is, there are instances that cannot be represented in this way.











Intersection graph

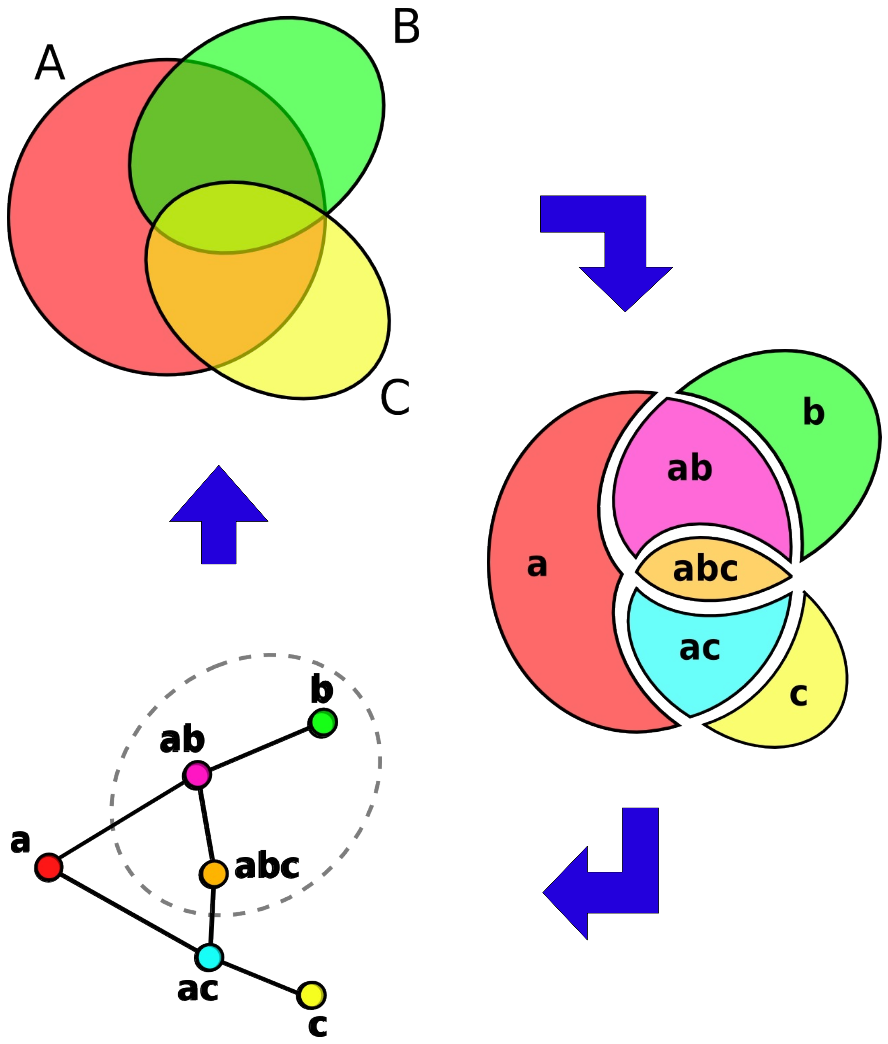
From an Euler diagram, it is possible to deduce a graph that expresses its structure: the intersection graph.

An intersection graph uses nodes to represent zones, and edges to indicate that the zones are contiguous.

For the reverse operation, we need to draw a class boundary:

- enclosing each class node,
- crossing each cut edge.

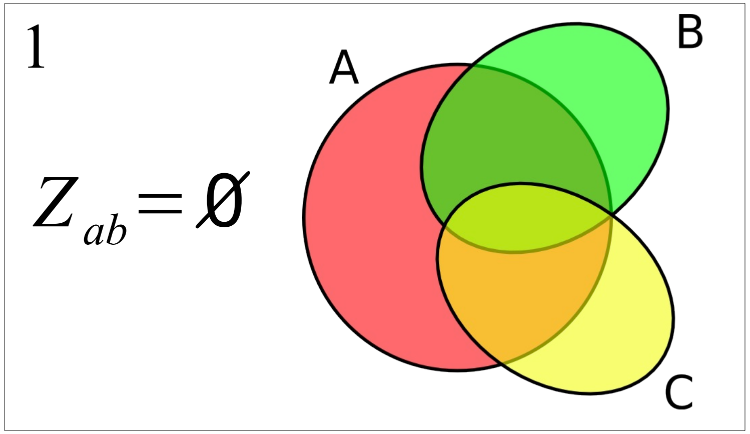
Intersection graphs allow us to work only on graphs.



Why Euler diagrams are not drawable?

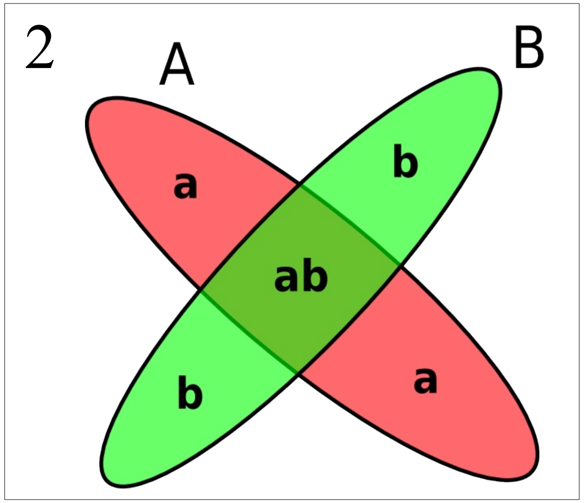
Euler diagrams might not be possible to draw for different reasons:

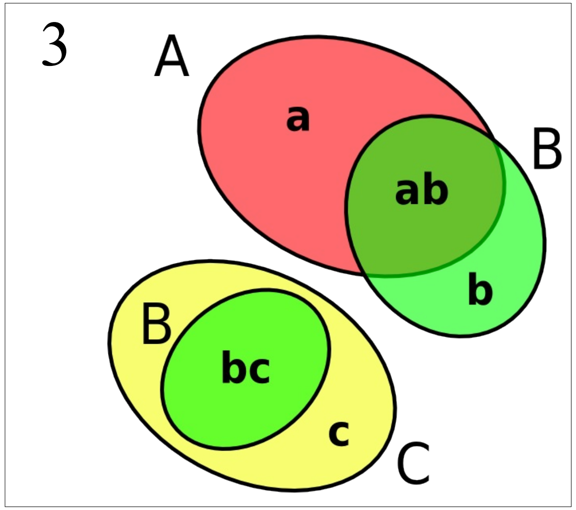
1. undesired overlaps,
2. zones not connected,
3. classes not connected.



We now need to study how these Euler diagrams properties are mapped into intersection graphs.

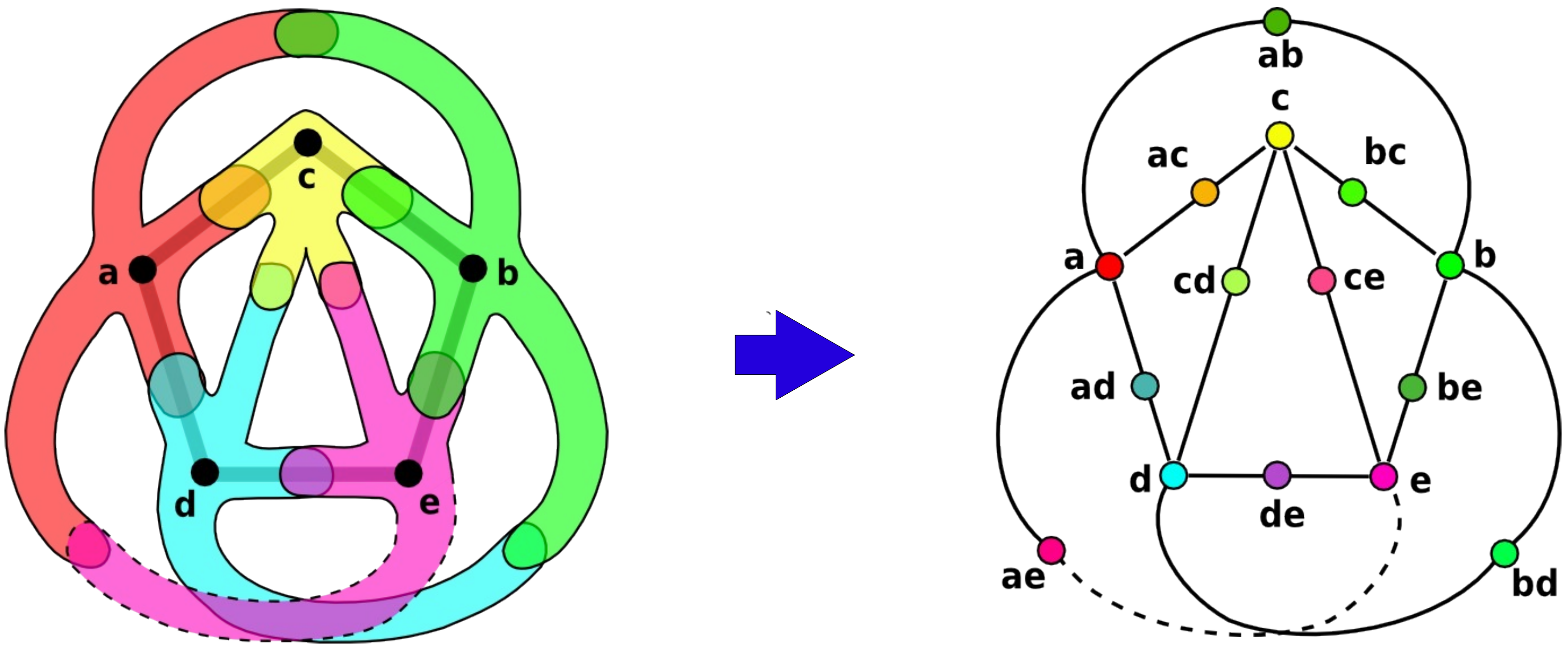
Then, we will be able to state whether a graph describes a correct Euler diagram or not.





Undesired overlaps

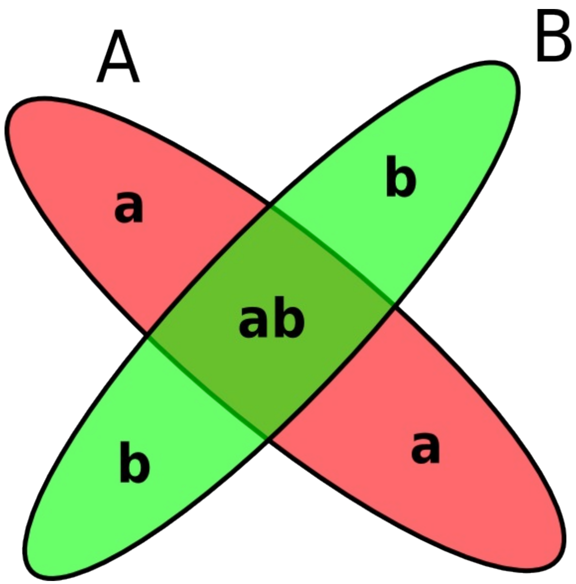
Undesired overlaps in Euler diagrams are related to edge crossings in the intersection graphs.

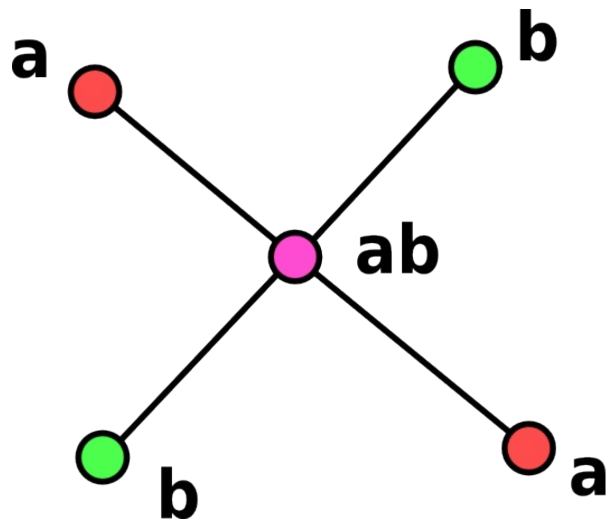


Planar intersection graph

Disconnected zones

We have disconnected zones when the intersection graph presents more instances of the same zone.

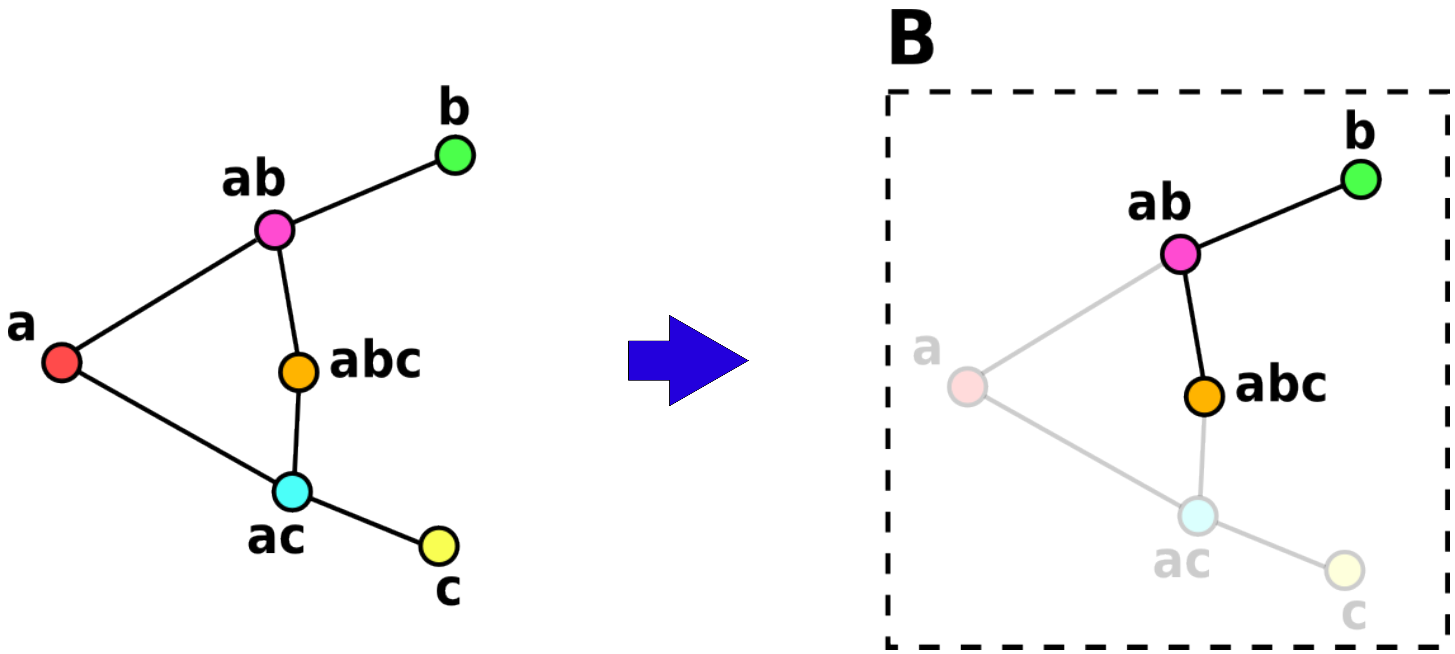




Avoid duplicated nodes

Disconnected classes

We have a disconnected class when the subgraph induced by the nodes of the class is not connected. We will call these subgraphs class schemas.



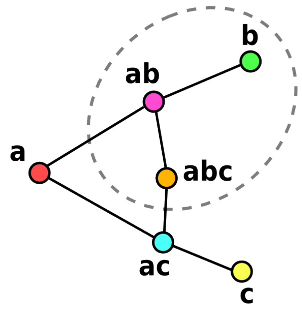
Connected class schemas

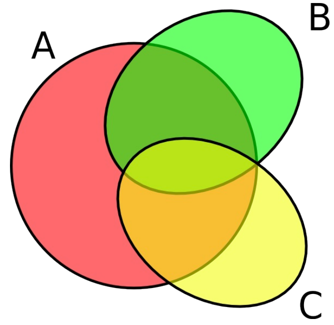
Building an Euler diagram

To construct the Euler diagram we could:

1. Detect the not-null zones
2. Construct a graph with nodes associated to the zones, paying attention to make it:
   - planar,
   - with connected class schemas,
   - without duplicated nodes.
3. Transform the graph obtained in a diagram.

$$Z_a = \{\ldots\}$$
$$Z_{ab} = \{\ldots\}$$
$$\ldots$$





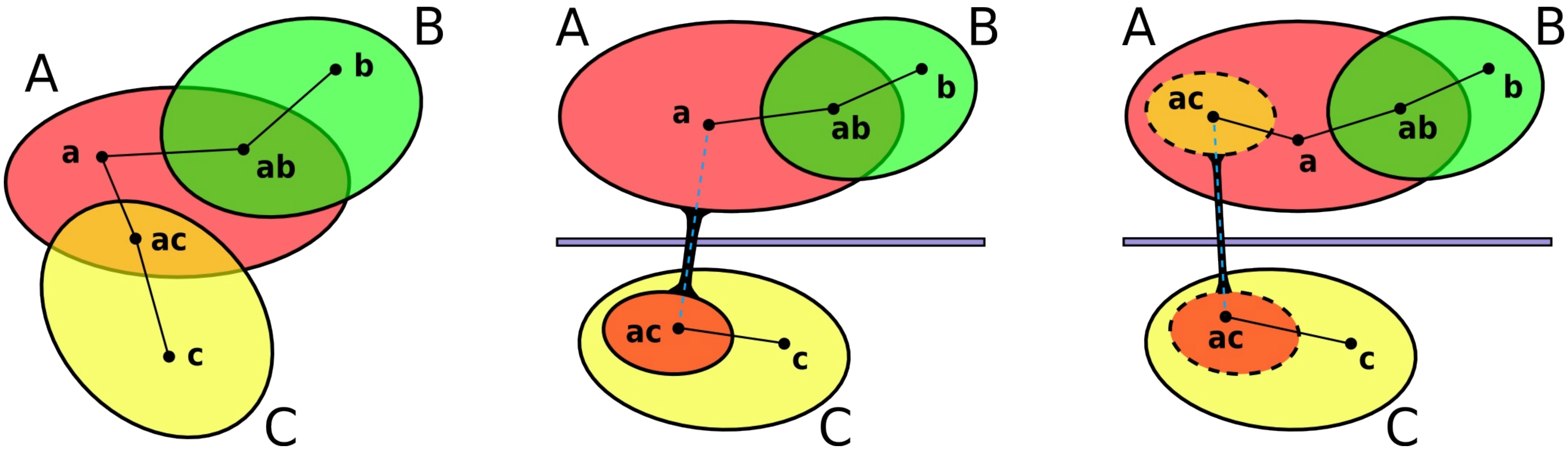
Clearly, for the undrawable diagrams such graph does not exist.

To construct those instances we will have to violate one or more rules. We will have to limit as much as possible these infractions, considering their importance according to the above list.

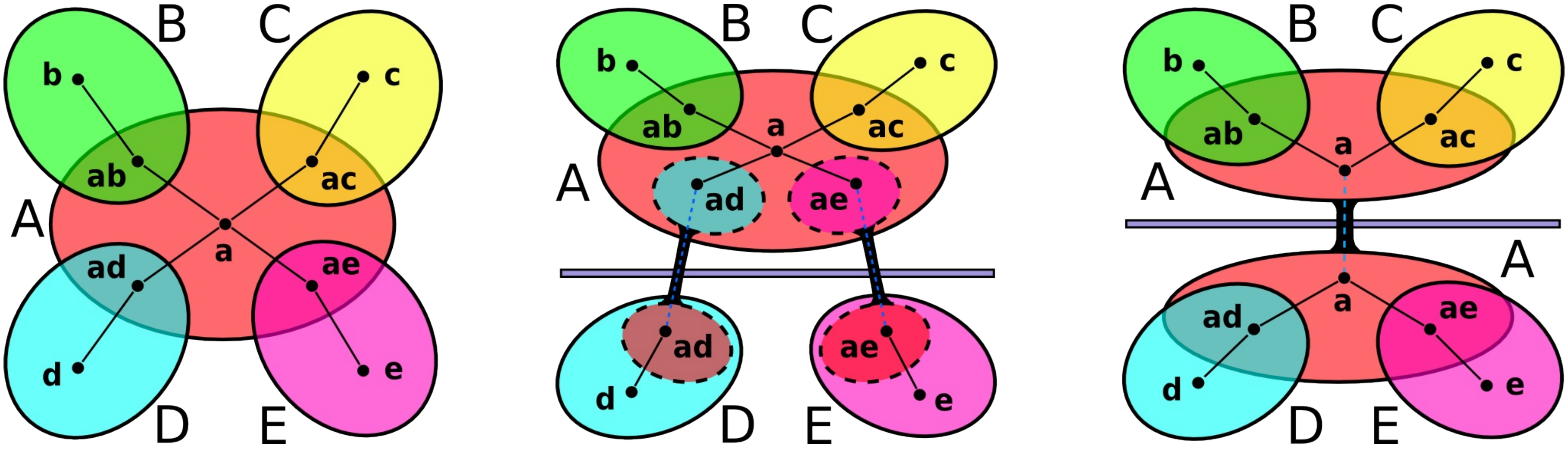
Undrawable diagrams patterns (1)

As we want to avoid undesired overlaps in every case, we will have to accept the presence of disconnected classes, when necessary. Allowing disconnected classes we can draw every classification (trivial).

Duplicate nodes of the intersection graph might contribute to obtain clearer diagrams. Node duplication might be intended as zone cloning or zone extension.



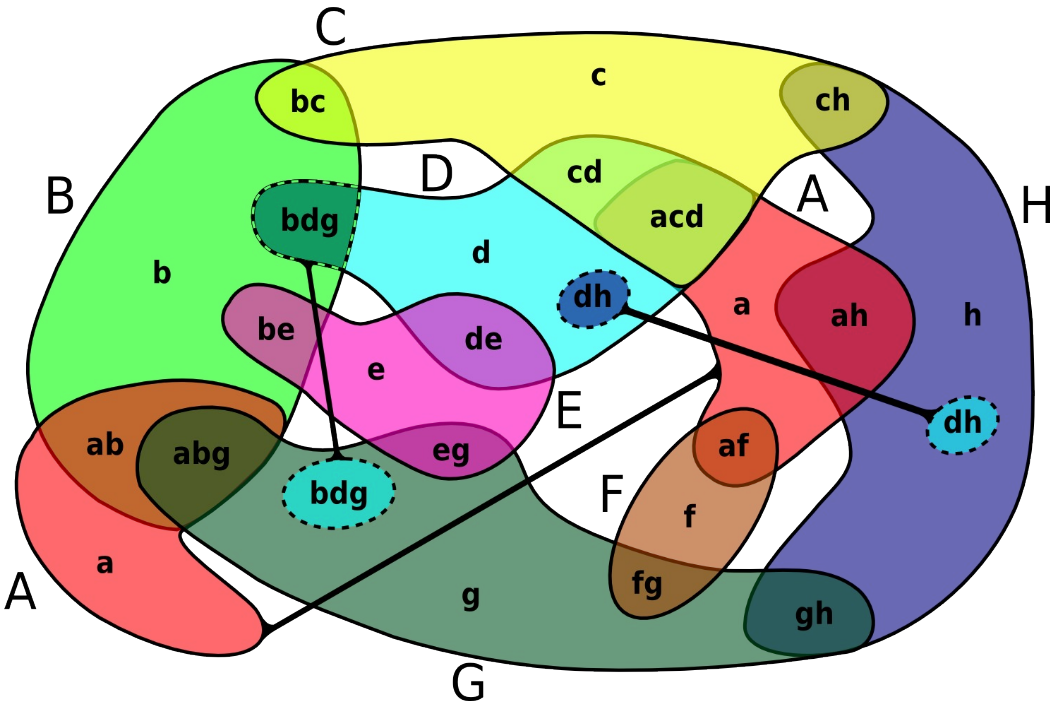
Undrawable diagrams patterns (2)



The choice of the pattern to use must be done carefully. Sometimes multiple patterns are applicable, leading to quite different results.

The only fundamental pattern is “disconnected class”. The other patterns might be allowed or not according to the preference of the user.

Conclusions and future work



In this work we highlighted the problems related to the generation of Euler representations through the generation of intersection graphs, even when the particular instance does not allow a proper Euler diagram.

We aim to publish soon an implementation of this method for the visualization of overlapping graph clusterings.

Thanks for your attention.