Clusters and quartet topologies

Hans-Jürgen Bendelt and Andreas Dress:
Reconstructing the shape of a tree from observed dissimilarity data.
Advances in Applied Mathematics **7** (1986) 309–343.

Speaker: Joseph, Chuang-Chieh Lin
Supervisor: Professor Maw-Shang Chang

Computation Theory Laboratory
Department of Computer Science and Information Engineering
National Chung Cheng University, Taiwan

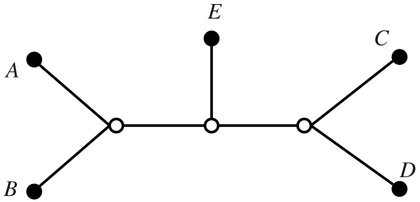
May 19, 2009

Outline

Outline

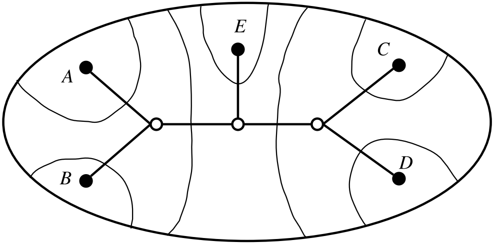
Evolutionary trees

- A set of n taxa S.
- An evolutionary tree $T = (V, E)$ on S:
  - internal vertices with degree 3.
  - bijection from S to the leaves of T.



Clusters

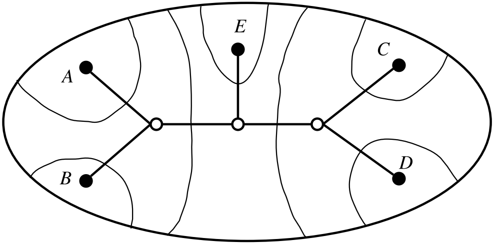
- Clusters: nonempty proper subsets of S according to T.
  - $\mathbb{C} = \{\mathcal{X} \subset S \mid \mathcal{X} \neq \emptyset, \exists e \in E$ such that any two taxa in $\mathcal{X}$ are connected by a path in T avoiding e, and $\mathcal{X}$ is **maximal** with respect to this property$\}$.
- Splits: two complementary clusters.



- There are $(2n-3)$ splits.
- $\mathbb{C} = \{\{A\},\{B\},\{C\},\{D\},\{E\},\{A,B\},\{C,D\},\{A,B,E\},\{C,D,E\}$ $\{A,B,C,D\},\{A,B,C,E\},\{A,B,D,E\},\{A,C,D,E\},\{B,C,D,E\}\}$.

Clusters

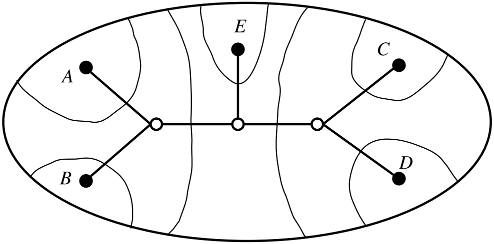
- Clusters: nonempty proper subsets of S according to T.
  - $\mathbb{C} = \{\mathcal{X} \subset S \mid \mathcal{X} \neq \emptyset, \exists e \in E$ such that any two taxa in $\mathcal{X}$ are connected by a path in T avoiding e, and $\mathcal{X}$ is **maximal** with respect to this property$\}$.
- Splits: two complementary clusters.



- There are $(2n - 3)$ splits.
- $\mathbb{C} = \{\{A\}, \{B\}, \{C\}, \{D\}, \{E\}, \{A, B\}, \{C, D\}, \{A, B, E\}, \{C, D, E\}$ $\{A, B, C, D\}, \{A, B, C, E\}, \{A, B, D, E\}, \{A, C, D, E\}, \{B, C, D, E\}\}$.

Clusters

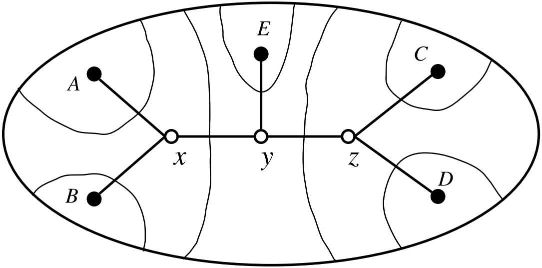
- Clusters: nonempty proper subsets of S according to T.
  - $\mathbb{C} = \{\mathcal{X} \subset S \mid \mathcal{X} \neq \emptyset, \exists e \in E$ such that any two taxa in $\mathcal{X}$ are connected by a path in T avoiding e, and $\mathcal{X}$ is **maximal** with respect to this property$\}$.
- Splits: two complementary clusters.



- There are $(2n-3)$ splits.
- $\mathbb{C} = \{\{A\}, \{B\}, \{C\}, \{D\}, \{E\}, \{A, B\}, \{C, D\}, \{A, B, E\}, \{C, D, E\}$ $\{A, B, C, D\}, \{A, B, C, E\}, \{A, B, D, E\}, \{A, C, D, E\}, \{B, C, D, E\}\}$.

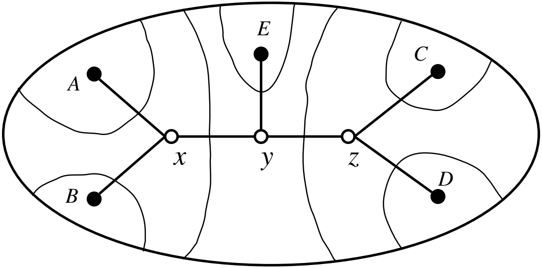
A distinguished subsystem of clusters

- $\mathcal{X} \in \rho(v)$ iff the deletion of some edges e of T incident with v results in two subtrees of T one of which contains the vertex v and is defined on $\mathcal{X}$.



A distinguished subsystem of clusters (contd.)

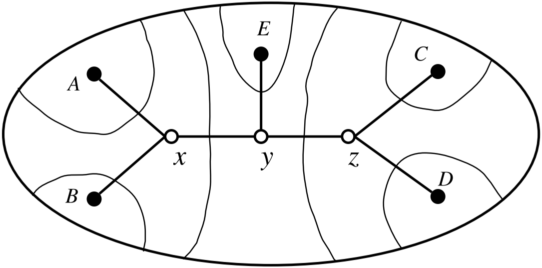
- Assume that
  $\mathbb{C} = \{\{A\}, \{B\}, \{C\}, \{D\}, \{E\}, \{A, B\}, \{C, D\}, \{A, B, E\}, \{C, D, E\}$
  $\{A, B, C, D\}, \{A, B, C, E\}, \{A, B, D, E\}, \{A, C, D, E\}, \{B, C, D, E\}\}$.



- $\rho(A) = \{\{A\}\}, \ \rho(B) = \{\{B\}\}, \ \rho(C) = \{\{C\}\}, \ldots$
- $\rho(x) = \{\{A, B\}, \{A, C, D, E\}, \{B, C, D, E\}\}$;
  $\rho(y) = \{\{A, B, E\}, \{A, B, C, D\}, \{C, D, E\}\}$;
  $\rho(z) = \{\{C, D\}, \{A, B, C, E\}, \{A, B, D, E\}\}$.

A distinguished subsystem of clusters (contd.)

- Assume that
  $\mathbb{C} = \{\{A\}, \{B\}, \{C\}, \{D\}, \{E\}, \{A, B\}, \{C, D\}, \{A, B, E\}, \{C, D, E\}$
  $\{A, B, C, D\}, \{A, B, C, E\}, \{A, B, D, E\}, \{A, C, D, E\}, \{B, C, D, E\}\}$.



- $\rho(A) = \{\{A\}\},\ \ \rho(B) = \{\{B\}\},\ \ \rho(C) = \{\{C\}\}, \ldots$
- $\rho(x) = \{\{A, B\}, \{A, C, D, E\}, \{B, C, D, E\}\}$;
  $\rho(y) = \{\{A, B, E\}, \{A, B, C, D\}, \{C, D, E\}\}$;
  $\rho(z) = \{\{C, D\}, \{A, B, C, E\}, \{A, B, D, E\}\}$.

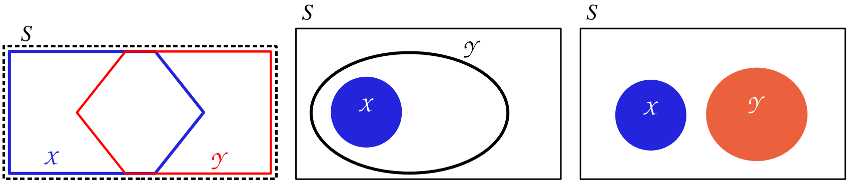
Outline

Proposition 1

$\mathbb{C}$ is the cluster system of an evolutionary tree on S

if and only if

- $\{A\} \in \mathbb{C}$ *for each* $A \in S$,
- $\mathcal{Y} \in \mathbb{C} \Leftrightarrow \bar{\mathcal{Y}} \in \mathbb{C}$ *(totally* $(4n-6)$ *clusters) and*
- *for* $\mathcal{X}, \mathcal{Y} \in \mathbb{C}$, *either* $\mathcal{X} \subseteq \mathcal{Y}$, $\mathcal{X} \subseteq \bar{\mathcal{Y}}$, $\bar{\mathcal{X}} \subseteq \mathcal{Y}$, *or* $\bar{\mathcal{X}} \subseteq \bar{\mathcal{Y}}$ *(compatible).*



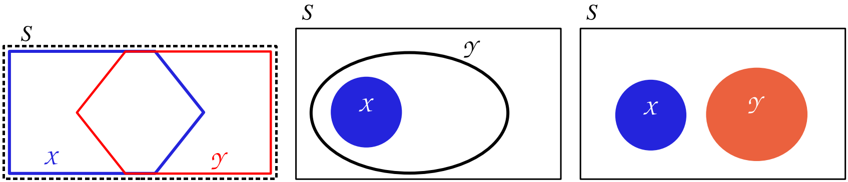
- Equivalent definition: $\mathcal{X} \equiv \mathcal{Y}$ iff **at least one** of the following intersections are empty:
  - $\mathcal{X} \cap \mathcal{Y}$, $\mathcal{X} \cap \bar{\mathcal{Y}}$, $\bar{\mathcal{X}} \cap \mathcal{Y}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}}$.

Proposition 1

$\mathbb{C}$ is the cluster system of an evolutionary tree on S

if and only if

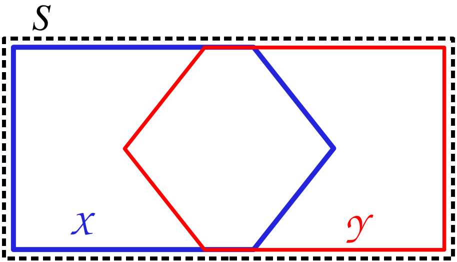
- $\{A\} \in \mathbb{C}$ *for each* $A \in S$,
- $\mathcal{Y} \in \mathbb{C} \Leftrightarrow \bar{\mathcal{Y}} \in \mathbb{C}$ *(totally* $(4n-6)$ *clusters) and*
- *for* $\mathcal{X}, \mathcal{Y} \in \mathbb{C}$, *either* $\mathcal{X} \subseteq \mathcal{Y}$, $\mathcal{X} \subseteq \bar{\mathcal{Y}}$, $\bar{\mathcal{X}} \subseteq \mathcal{Y}$, *or* $\bar{\mathcal{X}} \subseteq \bar{\mathcal{Y}}$ *(compatible).*



- Equivalent definition: $\mathcal{X} \equiv \mathcal{Y}$ iff **at least one** of the following intersections are empty:
  - $\mathcal{X} \cap \mathcal{Y}$, $\mathcal{X} \cap \bar{\mathcal{Y}}$, $\bar{\mathcal{X}} \cap \mathcal{Y}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}}$.

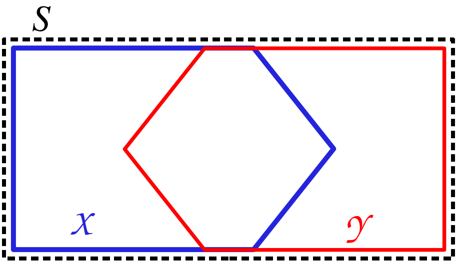
Compatible clusters

- $S = \{A, B, C, D, E\}$
- $\mathcal{X} = \{A, B, C, D\}, \mathcal{Y} = \{C, D, E\}$, $\bar{\mathcal{X}} = \{E\}, \bar{\mathcal{Y}} = \{A, B\}$.
- $\mathcal{X} \cap \mathcal{Y} = \{C, D\}$, $\mathcal{X} \cap \bar{\mathcal{Y}} = \{A, B\}$, $\bar{\mathcal{X}} \cap \mathcal{Y} = \{E\}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}} = \emptyset$.



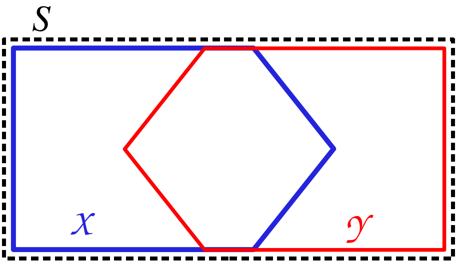
Compatible clusters

- $S = \{A, B, C, D, E\}$
- $\mathcal{X} = \{A, B, C, D\}, \mathcal{Y} = \{C, D, E\}$, $\bar{\mathcal{X}} = \{E\}, \bar{\mathcal{Y}} = \{A, B\}$.
- $\mathcal{X} \cap \mathcal{Y} = \{C, D\}$, $\mathcal{X} \cap \bar{\mathcal{Y}} = \{A, B\}$, $\bar{\mathcal{X}} \cap \mathcal{Y} = \{E\}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}} = \emptyset$.



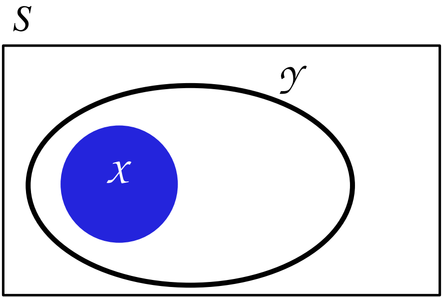
Compatible clusters

- $S = \{A, B, C, D, E\}$
- $\mathcal{X} = \{A, B, C, D\}, \mathcal{Y} = \{C, D, E\}$, $\bar{\mathcal{X}} = \{E\}, \bar{\mathcal{Y}} = \{A, B\}$.
- $\mathcal{X} \cap \mathcal{Y} = \{C, D\}$, $\mathcal{X} \cap \bar{\mathcal{Y}} = \{A, B\}$, $\bar{\mathcal{X}} \cap \mathcal{Y} = \{E\}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}} = \emptyset$.



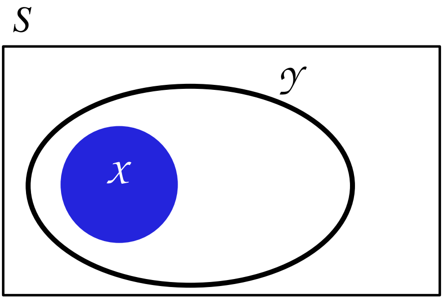
Compatible clusters (contd.)

- $S = \{A, B, C, D, E\}$
- $\mathcal{X} = \{A, B\}, \mathcal{Y} = \{A, B, C, D\}, \bar{\mathcal{X}} = \{C, D, E\}, \bar{\mathcal{Y}} = \{E\}$.
- $\mathcal{X} \cap \mathcal{Y} = \{A, B\}$, $\mathcal{X} \cap \bar{\mathcal{Y}} = \emptyset$, $\bar{\mathcal{X}} \cap \mathcal{Y} = \{C, D\}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}} = \{E\}$.



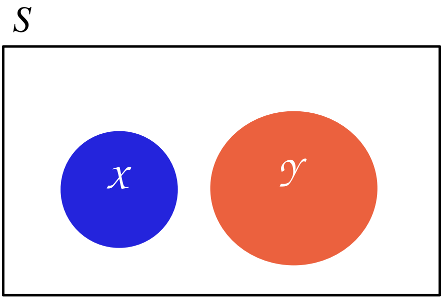
Compatible clusters (contd.)

- $S = \{A, B, C, D, E\}$
- $\mathcal{X} = \{A, B\}, \mathcal{Y} = \{A, B, C, D\}$, $\bar{\mathcal{X}} = \{C, D, E\}, \bar{\mathcal{Y}} = \{E\}$.
- $\mathcal{X} \cap \mathcal{Y} = \{A, B\}$, $\mathcal{X} \cap \bar{\mathcal{Y}} = \emptyset$, $\bar{\mathcal{X}} \cap \mathcal{Y} = \{C, D\}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}} = \{E\}$.



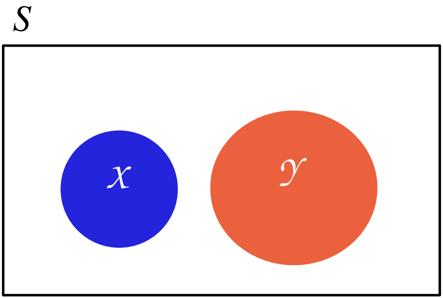
Compatible clusters (contd.)

- $S = \{A, B, C, D, E\}$
- $\mathcal{X} = \{A, B\}, \mathcal{Y} = \{C, D\}$, $\bar{\mathcal{X}} = \{C, D, E\}, \bar{\mathcal{Y}} = \{A, B, E\}$.
- $\mathcal{X} \cap \mathcal{Y} = \emptyset$, $\mathcal{X} \cap \bar{\mathcal{Y}} = \{A, B\}$, $\bar{\mathcal{X}} \cap \mathcal{Y} = \{C, D\}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}} = \{E\}$.



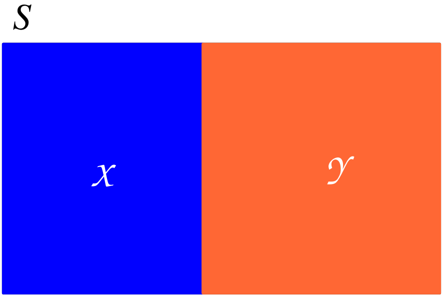
Compatible clusters (contd.)

- $S = \{A, B, C, D, E\}$
- $\mathcal{X} = \{A, B\}, \mathcal{Y} = \{C, D\}$, $\bar{\mathcal{X}} = \{C, D, E\}, \bar{\mathcal{Y}} = \{A, B, E\}$.
- $\mathcal{X} \cap \mathcal{Y} = \emptyset$, $\mathcal{X} \cap \bar{\mathcal{Y}} = \{A, B\}$, $\bar{\mathcal{X}} \cap \mathcal{Y} = \{C, D\}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}} = \{E\}$.



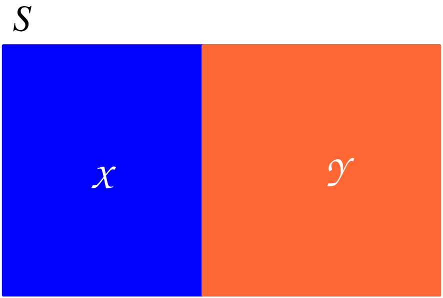
Compatible clusters (contd.)

- $S = \{A, B, C, D, E\}$
- $\mathcal{X} = \{A, B\}, \mathcal{Y} = \{C, D, E\}$, $\bar{\mathcal{X}} = \{C, D, E\}, \bar{\mathcal{Y}} = \{A, B\}$.
- $\mathcal{X} \cap \mathcal{Y} = \emptyset$, $\mathcal{X} \cap \bar{\mathcal{Y}} = \{A, B\}$, $\bar{\mathcal{X}} \cap \mathcal{Y} = \{C, D\}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}} = \emptyset$.



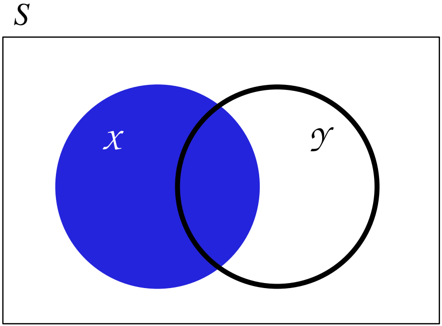
Compatible clusters (contd.)

- $S = \{A, B, C, D, E\}$
- $\mathcal{X} = \{A, B\}, \mathcal{Y} = \{C, D, E\}$, $\bar{\mathcal{X}} = \{C, D, E\}, \bar{\mathcal{Y}} = \{A, B\}$.
- $\mathcal{X} \cap \mathcal{Y} = \emptyset$, $\mathcal{X} \cap \bar{\mathcal{Y}} = \{A, B\}$, $\bar{\mathcal{X}} \cap \mathcal{Y} = \{C, D\}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}} = \emptyset$.



Non-compatible clusters

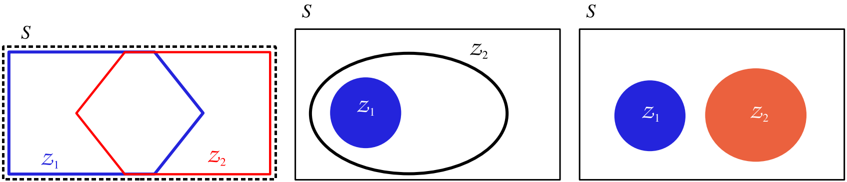
- $S = \{A, B, C, D, E\}$
- $\mathcal{X} = \{A, B, C\}, \mathcal{Y} = \{B, C, D\}$, $\bar{\mathcal{X}} = \{D, E\}, \bar{\mathcal{Y}} = \{A, E\}$.
- $\mathcal{X} \cap \mathcal{Y} = \{B, C\}$, $\mathcal{X} \cap \bar{\mathcal{Y}} = \{A\}$, $\bar{\mathcal{X}} \cap \mathcal{Y} = \{D\}$, $\bar{\mathcal{X}} \cap \bar{\mathcal{Y}} = \{E\}$.



Some observations (for a compatible $\mathbb{C}$)

- $\mathcal{X} \subseteq \mathcal{Y} \Leftrightarrow \bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$ for any two clusters $\mathcal{X}, \mathcal{Y}$.

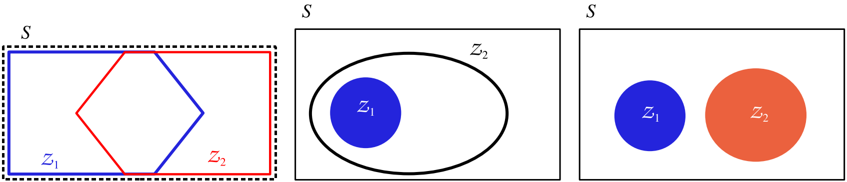
- $\forall \mathcal{X}, \mathcal{Y} \in \mathbb{C}$, where $\mathcal{X} \subseteq \mathcal{Y}$, if $\mathcal{X} \subseteq \mathcal{Z}_1, \ldots, \mathcal{Z}_k \subseteq \mathcal{Y}$, and $|\mathcal{Z}_i| \leq |\mathcal{Z}_j|$ for $i \leq j$, then $\mathcal{Z}_1 \subseteq \mathcal{Z}_2 \ldots \subseteq \mathcal{Z}_k$ ($\mathcal{Z}_1, \ldots, \mathcal{Z}_k$ forms a chain).
  - $\because \mathcal{X} \subseteq \mathcal{Z}_i, \mathcal{Z}_j \subseteq \mathcal{Y} \Rightarrow \emptyset \neq \mathcal{X} \subseteq \mathcal{Z}_i \cap \mathcal{Z}_j$ and $\emptyset \neq \bar{\mathcal{Y}} \subseteq \bar{\mathcal{Z}}_i \cap \bar{\mathcal{Z}}_j$
  - then we have $\mathcal{Z}_i \subseteq \mathcal{Z}_j$ or $\mathcal{Z}_j \subseteq \mathcal{Z}_i$ for all i, j.



Some observations (for a compatible $\mathbb{C}$)

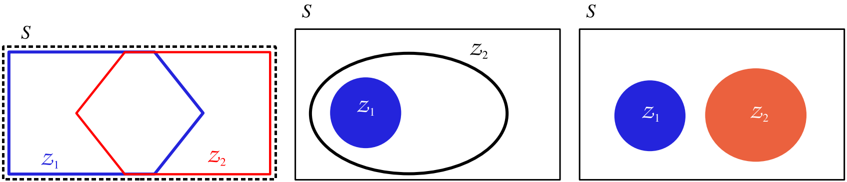
- $\mathcal{X} \subseteq \mathcal{Y} \Leftrightarrow \bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$ for any two clusters $\mathcal{X}, \mathcal{Y}$.

- $\forall \mathcal{X}, \mathcal{Y} \in \mathbb{C}$, where $\mathcal{X} \subseteq \mathcal{Y}$, if $\mathcal{X} \subseteq \mathcal{Z}_1, \ldots, \mathcal{Z}_k \subseteq \mathcal{Y}$, and $|\mathcal{Z}_i| \leq |\mathcal{Z}_j|$ for $i \leq j$, then $\mathcal{Z}_1 \subseteq \mathcal{Z}_2 \ldots \subseteq \mathcal{Z}_k$ ($\mathcal{Z}_1, \ldots, \mathcal{Z}_k$ forms a chain).
  - $\because \mathcal{X} \subseteq \mathcal{Z}_i, \mathcal{Z}_j \subseteq \mathcal{Y} \Rightarrow \emptyset \neq \mathcal{X} \subseteq \mathcal{Z}_i \cap \mathcal{Z}_j$ and $\emptyset \neq \bar{\mathcal{Y}} \subseteq \bar{\mathcal{Z}}_i \cap \bar{\mathcal{Z}}_j$
  - then we have $\mathcal{Z}_i \subseteq \mathcal{Z}_j$ or $\mathcal{Z}_j \subseteq \mathcal{Z}_i$ for all i, j.



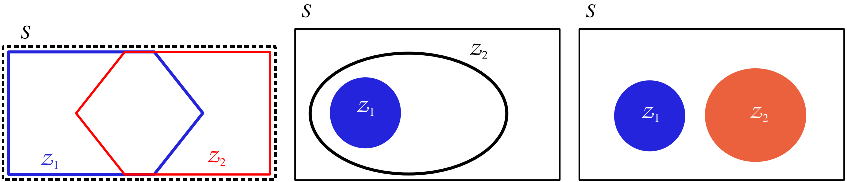
Some observations (for a compatible $\mathbb{C}$)

- $\mathcal{X} \subseteq \mathcal{Y} \Leftrightarrow \bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$ for any two clusters $\mathcal{X}, \mathcal{Y}$.
- $\forall \mathcal{X}, \mathcal{Y} \in \mathbb{C}$, where $\mathcal{X} \subseteq \mathcal{Y}$, if $\mathcal{X} \subseteq \mathcal{Z}_1, \dots, \mathcal{Z}_k \subseteq \mathcal{Y}$, and $|\mathcal{Z}_i| \leq |\mathcal{Z}_j|$ for $i \leq j$, then $\mathcal{Z}_1 \subseteq \mathcal{Z}_2 \dots \subseteq \mathcal{Z}_k$ ($\mathcal{Z}_1, \dots, \mathcal{Z}_k$ forms a chain).
  - $\because \mathcal{X} \subseteq \mathcal{Z}_i, \mathcal{Z}_j \subseteq \mathcal{Y} \Rightarrow \emptyset \neq \mathcal{X} \subseteq \mathcal{Z}_i \cap \mathcal{Z}_j$ and $\emptyset \neq \bar{\mathcal{Y}} \subseteq \bar{\mathcal{Z}}_i \cap \bar{\mathcal{Z}}_j$
  - then we have $\mathcal{Z}_i \subseteq \mathcal{Z}_j$ or $\mathcal{Z}_j \subseteq \mathcal{Z}_i$ for all i, j.



Some observations (for a compatible $\mathbb{C}$)

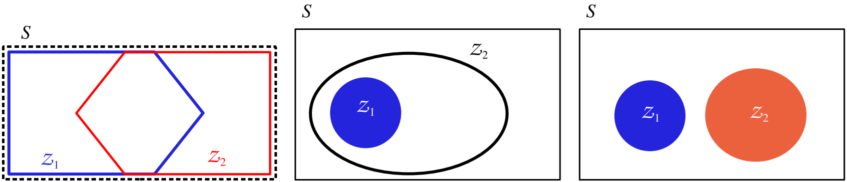
- $\mathcal{X} \subseteq \mathcal{Y} \Leftrightarrow \bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$ for any two clusters $\mathcal{X}, \mathcal{Y}$.
- $\forall \mathcal{X}, \mathcal{Y} \in \mathbb{C}$, where $\mathcal{X} \subseteq \mathcal{Y}$, if $\mathcal{X} \subseteq \mathcal{Z}_1, \ldots, \mathcal{Z}_k \subseteq \mathcal{Y}$, and $|\mathcal{Z}_i| \leq |\mathcal{Z}_j|$ for $i \leq j$, then $\mathcal{Z}_1 \subseteq \mathcal{Z}_2 \ldots \subseteq \mathcal{Z}_k$ ($\mathcal{Z}_1, \ldots, \mathcal{Z}_k$ forms a chain).
  - $\because \mathcal{X} \subseteq \mathcal{Z}_i, \mathcal{Z}_j \subseteq \mathcal{Y} \Rightarrow \emptyset \neq \mathcal{X} \subseteq \mathcal{Z}_i \cap \mathcal{Z}_j$ and $\emptyset \neq \bar{\mathcal{Y}} \subseteq \bar{\mathcal{Z}}_i \cap \bar{\mathcal{Z}}_j$
  - then we have $\mathcal{Z}_i \subseteq \mathcal{Z}_j$ or $\mathcal{Z}_j \subseteq \mathcal{Z}_i$ for all i, j.



Some observations (for a compatible $\mathbb{C}$)

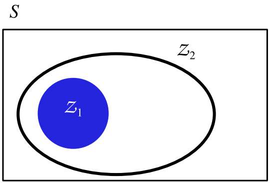
- $\mathcal{X} \subseteq \mathcal{Y} \Leftrightarrow \bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$ for any two clusters $\mathcal{X}, \mathcal{Y}$.

- $\forall \mathcal{X}, \mathcal{Y} \in \mathbb{C}$, where $\mathcal{X} \subseteq \mathcal{Y}$, if $\mathcal{X} \subseteq \mathcal{Z}_1, \ldots, \mathcal{Z}_k \subseteq \mathcal{Y}$, and $|\mathcal{Z}_i| \leq |\mathcal{Z}_j|$ for $i \leq j$, then $\mathcal{Z}_1 \subseteq \mathcal{Z}_2 \ldots \subseteq \mathcal{Z}_k$ ($\mathcal{Z}_1, \ldots, \mathcal{Z}_k$ forms a chain).
  - $\because \mathcal{X} \subseteq \mathcal{Z}_i, \mathcal{Z}_j \subseteq \mathcal{Y} \Rightarrow \emptyset \neq \mathcal{X} \subseteq \mathcal{Z}_i \cap \mathcal{Z}_j$ and $\emptyset \neq \bar{\mathcal{Y}} \subseteq \bar{\mathcal{Z}}_i \cap \bar{\mathcal{Z}}_j$
  - ▷ then we have $\mathcal{Z}_i \subseteq \mathcal{Z}_J$ or $\mathcal{Z}_j \subseteq \mathcal{Z}_i$ for all i, j.

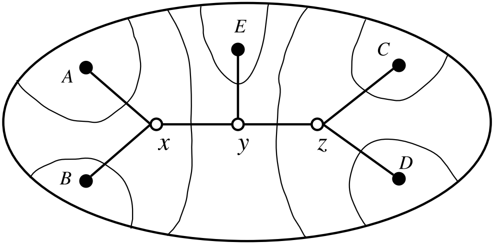


Some observations (for a compatible $\mathbb{C}$)

- $\mathcal{X} \subseteq \mathcal{Y} \Leftrightarrow \bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$ for any two clusters $\mathcal{X}, \mathcal{Y}$.
- $\forall \mathcal{X}, \mathcal{Y} \in \mathbb{C}$, where $\mathcal{X} \subseteq \mathcal{Y}$, if $\mathcal{X} \subseteq \mathcal{Z}_1, \ldots, \mathcal{Z}_k \subseteq \mathcal{Y}$, and $|\mathcal{Z}_i| \leq |\mathcal{Z}_j|$ for $i \leq j$, then $\mathcal{Z}_1 \subseteq \mathcal{Z}_2 \ldots \subseteq \mathcal{Z}_k$ ($\mathcal{Z}_1, \ldots, \mathcal{Z}_k$ forms a chain).
  - $\because \mathcal{X} \subseteq \mathcal{Z}_i, \mathcal{Z}_j \subseteq \mathcal{Y} \Rightarrow \emptyset \neq \mathcal{X} \subseteq \mathcal{Z}_i \cap \mathcal{Z}_j$ and $\emptyset \neq \bar{\mathcal{Y}} \subseteq \bar{\mathcal{Z}}_i \cap \bar{\mathcal{Z}}_j$
  - ▷ then we have $\mathcal{Z}_i \subseteq \mathcal{Z}_J$ or $\mathcal{Z}_j \subseteq \mathcal{Z}_i$ for all i, j.



Recall the distinguished subsystem of $\mathbb{C}$ ...



- $\rho(A) = \{\{A\}\},\ \ \rho(B) = \{\{B\}\},\ \ \rho(C) = \{\{C\}\}, \ldots,$
  $\rho(x) = \{\{A, B\}, \{A, C, D, E\}, \{B, C, D, E\}\};$
  $\rho(y) = \{\{A, B, E\}, \{A, B, C, D\}, \{C, D, E\}\};$
  $\rho(z) = \{\{C, D\}, \{A, B, C, E\}, \{A, B, D, E\}\}.$
- We can **DEFINE** a corresponding equivalence relation on $\mathbb{C}$:
  - For $\mathcal{X}, \mathcal{Y} \in \mathbb{C}$, we say $\mathcal{X} \equiv \mathcal{Y}$ if and only if either $\mathcal{X} = \mathcal{Y}$ or $\bar{\mathcal{X}}$ is a maximal proper subcluster of $\mathcal{Y}$.

How to reconstruct T from a corresponding $\mathbb{C}$?

- Each equivalence class $\rho(x)$ of $\equiv$ represents a vertex of T.
- $\rho(x) = \{A\} \Leftrightarrow$ a leaf node A.
- $\rho(x), \rho(y)$ represent **adjacent** vertices x, y iff $\exists \mathcal{Y} \in \mathbb{C}$ such that $\mathcal{Y} \in \rho(x)$ and $\bar{\mathcal{Y}} \in \rho(y)$.

Sketch of the proof of Proposition 1

Proposition 1

$\mathbb{C}$ is the cluster system of an evolutionary tree T on S

if and only if

- $\{A\} \in \mathbb{C}$ for each $A \in S$,
- $\mathcal{Y} \in \mathbb{C} \Leftrightarrow \bar{\mathcal{Y}} \in \mathbb{C}$ and
- every two $\mathcal{X}, \mathcal{Y} \in \mathbb{C}$ are compatible.

- The if-part is easy (by observing clusters corresponding to T).
- For the only-if-part:
  1. Prove that $\equiv$ is an equivalence relation (reflexive, symmetric, transitive).
     - ⋆ $\mathcal{X} \equiv \mathcal{Y}$ iff either $\mathcal{X} = \mathcal{Y}$ or $\bar{\mathcal{X}}$ is a maximal proper subcluster of $\mathcal{Y}$.
  2. Construct a corresponding graph T by $\equiv$.
  3. Show that the graph T is an evolutionary tree on S.

Sketch of the proof of Proposition 1

Proposition 1

$\mathbb{C}$ is the cluster system of an evolutionary tree T on S

if and only if

- $\{A\} \in \mathbb{C}$ for each $A \in S$,
- $\mathcal{Y} \in \mathbb{C} \Leftrightarrow \bar{\mathcal{Y}} \in \mathbb{C}$ and
- every two $\mathcal{X}, \mathcal{Y} \in \mathbb{C}$ are compatible.

- The if-part is easy (by observing clusters corresponding to T).
- For the only-if-part:
  1. Prove that $\equiv$ is an equivalence relation (reflexive, symmetric, transitive).
     - ★ $\mathcal{X} \equiv \mathcal{Y}$ iff either $\mathcal{X} = \mathcal{Y}$ or $\bar{\mathcal{X}}$ is a maximal proper subcluster of $\mathcal{Y}$.
  2. Construct a corresponding graph T by $\equiv$.
  3. Show that the graph T is an evolutionary tree on S.

Sketch of the proof of Proposition 1

Proposition 1

$\mathbb{C}$ is the cluster system of an evolutionary tree T on S

if and only if

- $\{A\} \in \mathbb{C}$ for each $A \in S$,
- $\mathcal{Y} \in \mathbb{C} \Leftrightarrow \bar{\mathcal{Y}} \in \mathbb{C}$ and
- every two $\mathcal{X}, \mathcal{Y} \in \mathbb{C}$ are compatible.

- The if-part is easy (by observing clusters corresponding to T).
- For the only-if-part:
  1. Prove that $\equiv$ is an equivalence relation (reflexive, symmetric, transitive).
     - ⋆ $\mathcal{X} \equiv \mathcal{Y}$ iff either $\mathcal{X} = \mathcal{Y}$ or $\bar{\mathcal{X}}$ is a maximal proper subcluster of $\mathcal{Y}$.
  2. Construct a corresponding graph T by $\equiv$.
  3. Show that the graph T is an evolutionary tree on S.

Sketch of the proof of Proposition 1

Proposition 1

$\mathbb{C}$ is the cluster system of an evolutionary tree T on S

if and only if

- $\{A\} \in \mathbb{C}$ for each $A \in S$,
- $\mathcal{Y} \in \mathbb{C} \Leftrightarrow \bar{\mathcal{Y}} \in \mathbb{C}$ and
- every two $\mathcal{X}, \mathcal{Y} \in \mathbb{C}$ are compatible.

- The if-part is easy (by observing clusters corresponding to T).
- For the only-if-part:
  1. Prove that $\equiv$ is an equivalence relation (reflexive, symmetric, transitive).
     - ⋆ $\mathcal{X} \equiv \mathcal{Y}$ iff either $\mathcal{X} = \mathcal{Y}$ or $\bar{\mathcal{X}}$ is a maximal proper subcluster of $\mathcal{Y}$.
  2. Construct a corresponding graph T by $\equiv$.
  3. Show that the graph T is an evolutionary tree on S.

Sketch of the proof of Proposition 1

Proposition 1

$\mathbb{C}$ is the cluster system of an evolutionary tree T on S

if and only if

- $\{A\} \in \mathbb{C}$ for each $A \in S$,
- $\mathcal{Y} \in \mathbb{C} \Leftrightarrow \bar{\mathcal{Y}} \in \mathbb{C}$ and
- every two $\mathcal{X}, \mathcal{Y} \in \mathbb{C}$ are compatible.

- The if-part is easy (by observing clusters corresponding to T).
- For the only-if-part:
  1. Prove that $\equiv$ is an equivalence relation (reflexive, symmetric, transitive).
     - ⋆ $\mathcal{X} \equiv \mathcal{Y}$ iff either $\mathcal{X} = \mathcal{Y}$ or $\bar{\mathcal{X}}$ is a maximal proper subcluster of $\mathcal{Y}$.
  2. Construct a corresponding graph T by $\equiv$.
  3. Show that the graph T is an evolutionary tree on S.

Sketch of the proof of Proposition 1

Proposition 1

$\mathbb{C}$ is the cluster system of an evolutionary tree T on S

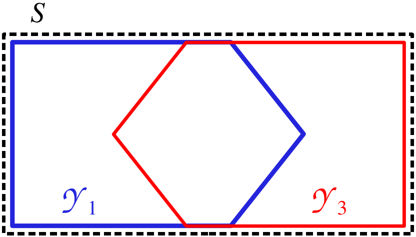
if and only if

- $\{A\} \in \mathbb{C}$ for each $A \in S$,
- $\mathcal{Y} \in \mathbb{C} \Leftrightarrow \bar{\mathcal{Y}} \in \mathbb{C}$ and
- every two $\mathcal{X}, \mathcal{Y} \in \mathbb{C}$ are compatible.

- The if-part is easy (by observing clusters corresponding to T).
- For the only-if-part:
  1. Prove that $\equiv$ is an equivalence relation (reflexive, symmetric, transitive).
     - ⋆ $\mathcal{X} \equiv \mathcal{Y}$ iff either $\mathcal{X} = \mathcal{Y}$ or $\bar{\mathcal{X}}$ is a maximal proper subcluster of $\mathcal{Y}$.
  2. Construct a corresponding graph T by $\equiv$.
  3. Show that the graph T is an evolutionary tree on S.

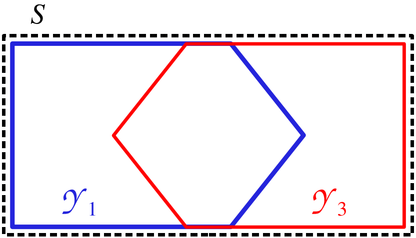
$\equiv$ is an equivalence relation

- It is easy to see that $\equiv$ is reflexive and symmetric.
- Assume that $\mathcal{Y}_1 \equiv \mathcal{Y}_2$ and $\mathcal{Y}_2 \equiv \mathcal{Y}_3$ and $\mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3$ are distinct.
- $\bar{\mathcal{Y}}_1$ and $\bar{\mathcal{Y}}_3$ are maximal proper subsets of $\mathcal{Y}_2$ in $\mathbb{C}$ by assumption.
- We have $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$.
  - $\mathcal{Y}_1 \not\subseteq \mathcal{Y}_3$ and $\mathcal{Y}_3 \not\subseteq \mathcal{Y}_1$ (by maximality)
  - $\mathcal{Y}_1 \cap \mathcal{Y}_3 \neq \emptyset$ ($\because \bar{\mathcal{Y}}_2 \subseteq \mathcal{Y}_1 \cap \mathcal{Y}_3$).
  - Hence $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$ (by the assumption of compatibility).



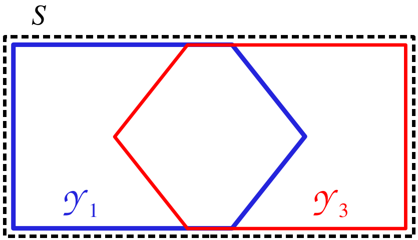
$\equiv$ is an equivalence relation

- It is easy to see that $\equiv$ is reflexive and symmetric.
- Assume that $\mathcal{Y}_1 \equiv \mathcal{Y}_2$ and $\mathcal{Y}_2 \equiv \mathcal{Y}_3$ and $\mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3$ are distinct.
- $\bar{\mathcal{Y}}_1$ and $\bar{\mathcal{Y}}_3$ are maximal proper subsets of $\mathcal{Y}_2$ in $\mathbb{C}$ by assumption.
- We have $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$.
  - $\mathcal{Y}_1 \not\subseteq \mathcal{Y}_3$ and $\mathcal{Y}_3 \not\subseteq \mathcal{Y}_1$ (by maximality)
  - $\mathcal{Y}_1 \cap \mathcal{Y}_3 \neq \emptyset$ ($\because \bar{\mathcal{Y}}_2 \subseteq \mathcal{Y}_1 \cap \mathcal{Y}_3$).
  - Hence $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$ (by the assumption of compatibility).



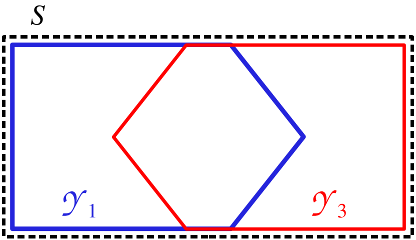
$\equiv$ is an equivalence relation

- It is easy to see that $\equiv$ is reflexive and symmetric.
- Assume that $\mathcal{Y}_1 \equiv \mathcal{Y}_2$ and $\mathcal{Y}_2 \equiv \mathcal{Y}_3$ and $\mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3$ are distinct.
- $\bar{\mathcal{Y}}_1$ and $\bar{\mathcal{Y}}_3$ are maximal proper subsets of $\mathcal{Y}_2$ in $\mathbb{C}$ by assumption.
- We have $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$.
  - $\mathcal{Y}_1 \not\subseteq \mathcal{Y}_3$ and $\mathcal{Y}_3 \not\subseteq \mathcal{Y}_1$ (by maximality)
  - $\mathcal{Y}_1 \cap \mathcal{Y}_3 \neq \emptyset$ ($\because \bar{\mathcal{Y}}_2 \subseteq \mathcal{Y}_1 \cap \mathcal{Y}_3$).
  - Hence $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$ (by the assumption of compatibility).



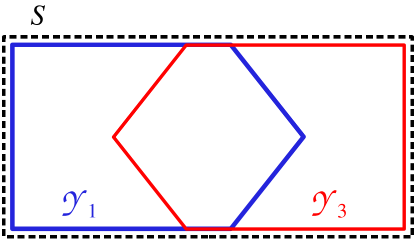
$\equiv$ is an equivalence relation

- It is easy to see that $\equiv$ is reflexive and symmetric.
- Assume that $\mathcal{Y}_1 \equiv \mathcal{Y}_2$ and $\mathcal{Y}_2 \equiv \mathcal{Y}_3$ and $\mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3$ are distinct.
- $\bar{\mathcal{Y}}_1$ and $\bar{\mathcal{Y}}_3$ are maximal proper subsets of $\mathcal{Y}_2$ in $\mathbb{C}$ by assumption.
- We have $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$.
    - $\mathcal{Y}_1 \not\subseteq \mathcal{Y}_3$ and $\mathcal{Y}_3 \not\subseteq \mathcal{Y}_1$ (by maximality)
    - $\mathcal{Y}_1 \cap \mathcal{Y}_3 \neq \emptyset$ ($\because \bar{\mathcal{Y}}_2 \subseteq \mathcal{Y}_1 \cap \mathcal{Y}_3$).
    - Hence $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$ (by the assumption of compatibility).



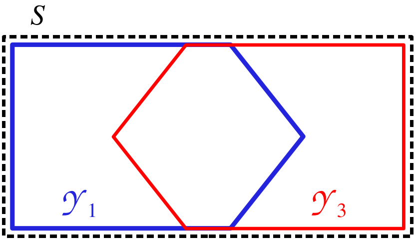
$\equiv$ is an equivalence relation

- It is easy to see that $\equiv$ is reflexive and symmetric.
- Assume that $\mathcal{Y}_1 \equiv \mathcal{Y}_2$ and $\mathcal{Y}_2 \equiv \mathcal{Y}_3$ and $\mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3$ are distinct.
- $\bar{\mathcal{Y}}_1$ and $\bar{\mathcal{Y}}_3$ are maximal proper subsets of $\mathcal{Y}_2$ in $\mathbb{C}$ by assumption.
- We have $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$.
  - $\mathcal{Y}_1 \not\subseteq \mathcal{Y}_3$ and $\mathcal{Y}_3 \not\subseteq \mathcal{Y}_1$ (by maximality)
  - $\mathcal{Y}_1 \cap \mathcal{Y}_3 \neq \emptyset$ ($\because \bar{\mathcal{Y}}_2 \subseteq \mathcal{Y}_1 \cap \mathcal{Y}_3$).
  - Hence $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$ (by the assumption of compatibility).



$\equiv$ is an equivalence relation

- It is easy to see that $\equiv$ is reflexive and symmetric.
- Assume that $\mathcal{Y}_1 \equiv \mathcal{Y}_2$ and $\mathcal{Y}_2 \equiv \mathcal{Y}_3$ and $\mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3$ are distinct.
- $\bar{\mathcal{Y}}_1$ and $\bar{\mathcal{Y}}_3$ are maximal proper subsets of $\mathcal{Y}_2$ in $\mathbb{C}$ by assumption.
- We have $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$.
  - $\mathcal{Y}_1 \not\subseteq \mathcal{Y}_3$ and $\mathcal{Y}_3 \not\subseteq \mathcal{Y}_1$ (by maximality)
  - $\mathcal{Y}_1 \cap \mathcal{Y}_3 \neq \emptyset$ ($\because \bar{\mathcal{Y}}_2 \subseteq \mathcal{Y}_1 \cap \mathcal{Y}_3$).
  - Hence $\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_3$ (by the assumption of compatibility).



$\equiv$ is an equivalence relation (contd.)

$$\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_2 \text{ and } \bar{\mathcal{Y}}_3 \subseteq \mathcal{Y}_2$$

- Let $\mathcal{X} \in \mathbb{C}$ such that $\bar{\mathcal{Y}}_1 \subseteq \mathcal{X} \subseteq \mathcal{Y}_3$. Then $\mathcal{X} = \bar{\mathcal{Y}}_1$ or $\mathcal{X} = \mathcal{Y}_3$ (Thus $\mathcal{Y}_1 \equiv \mathcal{Y}_3$).
  - Either $\mathcal{X} \subseteq \mathcal{Y}_2$ or $\bar{\mathcal{X}} \subseteq \mathcal{Y}_2$ ($\because \emptyset \neq \bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_2 \cap \mathcal{X}$ and $\emptyset \neq \bar{\mathcal{Y}}_3 \subseteq \mathcal{Y}_2 \cap \bar{\mathcal{X}}$).
    - $\mathcal{X} \subseteq \bar{\mathcal{Y}}_2$? $\bar{\mathcal{X}} \subseteq \bar{\mathcal{Y}}_2$?
  - Hence $\bar{\mathcal{Y}}_1 \subseteq \mathcal{X} \subseteq \mathcal{Y}_2$ or $\bar{\mathcal{Y}}_3 \subseteq \bar{\mathcal{X}} \subseteq \mathcal{Y}_2$.
  - By maximality we have either $\mathcal{X} = \bar{\mathcal{Y}}_1$ or $\mathcal{X} = \mathcal{Y}_3$.

$\equiv$ is an equivalence relation (contd.)

$$\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_2 \text{ and } \bar{\mathcal{Y}}_3 \subseteq \mathcal{Y}_2$$

- Let $\mathcal{X} \in \mathbb{C}$ such that $\bar{\mathcal{Y}}_1 \subseteq \mathcal{X} \subseteq \mathcal{Y}_3$. Then $\mathcal{X} = \bar{\mathcal{Y}}_1$ or $\mathcal{X} = \mathcal{Y}_3$ (Thus $\mathcal{Y}_1 \equiv \mathcal{Y}_3$).
  - Either $\mathcal{X} \subseteq \mathcal{Y}_2$ or $\bar{\mathcal{X}} \subseteq \mathcal{Y}_2$ ($\because \emptyset \neq \bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_2 \cap \mathcal{X}$ and $\emptyset \neq \bar{\mathcal{Y}}_3 \subseteq \mathcal{Y}_2 \cap \bar{\mathcal{X}}$).
    - $\mathcal{X} \subseteq \bar{\mathcal{Y}}_2$? $\bar{\mathcal{X}} \subseteq \bar{\mathcal{Y}}_2$?
  - Hence $\bar{\mathcal{Y}}_1 \subseteq \mathcal{X} \subseteq \mathcal{Y}_2$ or $\bar{\mathcal{Y}}_3 \subseteq \bar{\mathcal{X}} \subseteq \mathcal{Y}_2$.
  - By maximality we have either $\mathcal{X} = \bar{\mathcal{Y}}_1$ or $\mathcal{X} = \mathcal{Y}_3$.

$\equiv$ is an equivalence relation (contd.)

$$\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_2 \text{ and } \bar{\mathcal{Y}}_3 \subseteq \mathcal{Y}_2$$

- Let $\mathcal{X} \in \mathbb{C}$ such that $\bar{\mathcal{Y}}_1 \subseteq \mathcal{X} \subseteq \mathcal{Y}_3$. Then $\mathcal{X} = \bar{\mathcal{Y}}_1$ or $\mathcal{X} = \mathcal{Y}_3$ (Thus $\mathcal{Y}_1 \equiv \mathcal{Y}_3$).
  - Either $\mathcal{X} \subseteq \mathcal{Y}_2$ or $\bar{\mathcal{X}} \subseteq \mathcal{Y}_2$ ($\because \emptyset \neq \bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_2 \cap \mathcal{X}$ and $\emptyset \neq \bar{\mathcal{Y}}_3 \subseteq \mathcal{Y}_2 \cap \bar{\mathcal{X}}$).
    - $\mathcal{X} \subseteq \bar{\mathcal{Y}}_2$? $\bar{\mathcal{X}} \subseteq \bar{\mathcal{Y}}_2$?
  - Hence $\bar{\mathcal{Y}}_1 \subseteq \mathcal{X} \subseteq \mathcal{Y}_2$ or $\bar{\mathcal{Y}}_3 \subseteq \bar{\mathcal{X}} \subseteq \mathcal{Y}_2$.
  - By maximality we have either $\mathcal{X} = \bar{\mathcal{Y}}_1$ or $\mathcal{X} = \mathcal{Y}_3$.

$\equiv$ is an equivalence relation (contd.)

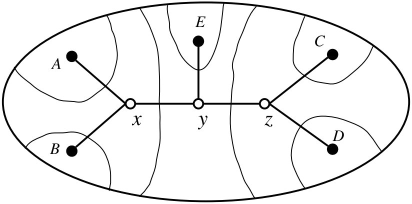
$$\bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_2 \text{ and } \bar{\mathcal{Y}}_3 \subseteq \mathcal{Y}_2$$

- Let $\mathcal{X} \in \mathbb{C}$ such that $\bar{\mathcal{Y}}_1 \subseteq \mathcal{X} \subseteq \mathcal{Y}_3$. Then $\mathcal{X} = \bar{\mathcal{Y}}_1$ or $\mathcal{X} = \mathcal{Y}_3$ (Thus $\mathcal{Y}_1 \equiv \mathcal{Y}_3$).
  - Either $\mathcal{X} \subseteq \mathcal{Y}_2$ or $\bar{\mathcal{X}} \subseteq \mathcal{Y}_2$ ($\because \emptyset \neq \bar{\mathcal{Y}}_1 \subseteq \mathcal{Y}_2 \cap \mathcal{X}$ and $\emptyset \neq \bar{\mathcal{Y}}_3 \subseteq \mathcal{Y}_2 \cap \bar{\mathcal{X}}$).
    - $\mathcal{X} \subseteq \bar{\mathcal{Y}}_2$? $\bar{\mathcal{X}} \subseteq \bar{\mathcal{Y}}_2$?
  - Hence $\bar{\mathcal{Y}}_1 \subseteq \mathcal{X} \subseteq \mathcal{Y}_2$ or $\bar{\mathcal{Y}}_3 \subseteq \bar{\mathcal{X}} \subseteq \mathcal{Y}_2$.
  - By maximality we have either $\mathcal{X} = \bar{\mathcal{Y}}_1$ or $\mathcal{X} = \mathcal{Y}_3$.

Constructing a graph T according to '$\equiv$'

- The vertices correspond to the equivalence classes of $\equiv$.
- Two classes are adjacent iff they contain some complementary pair $\mathcal{Y}, \bar{\mathcal{Y}}$.
- An equivalence class represents a taxa A if it has only one member $\{A\}$.

An observation



Note:
A **subcluster** means a subset of a cluster which is also a cluster.

- From the point of view of clusters, two clusters $\mathcal{Y}_1, \mathcal{Y}_2$ represent adjacent vertices iff
  - $\bar{\mathcal{Y}}_1 = \mathcal{Y}_2$ ($\mathcal{Y}_1 \cap \mathcal{Y}_2 = \emptyset$ in this case);
  - $\mathcal{Y}_1$ is a maximal proper subcluster of $\mathcal{Y}_2$ ($\mathcal{Y}_1 \leftrightarrow \bar{\mathcal{Y}}_1 \equiv \mathcal{Y}_2$);
  - $\mathcal{Y}_2$ is a maximal proper subcluster of $\mathcal{Y}_1$ ($\mathcal{Y}_2 \leftrightarrow \bar{\mathcal{Y}}_2 \equiv \mathcal{Y}_1$);
  - $\exists! \mathcal{X} \in \mathbb{C}$ such that $\bar{\mathcal{Y}}_1 \subset \mathcal{X} \subset \mathcal{Y}_2$ ($\mathcal{Y}_1 \equiv \mathcal{X} \leftrightarrow \bar{\mathcal{X}} \equiv \mathcal{Y}_2$).

How about the connectivity of the constructed graph?

Proof of the connectivity and the cycle-freeness of T

- Let $\mathcal{X}, \bar{\mathcal{X}}$ be an arbitrary pair of complementary clusters.
- For any $\mathcal{Y} \in \mathbb{C}$, $\exists ! \mathcal{Y}_0$ with $\mathcal{Y}_0 \equiv \mathcal{Y}$ such that $\mathcal{Y}_0 \subseteq \mathcal{X}$ or $\mathcal{Y}_0 \subseteq \bar{\mathcal{X}}$.
  - Existence: either $\mathcal{Y} \subseteq \mathcal{X}$, $\mathcal{Y} \subseteq \bar{\mathcal{X}}$, $\bar{\mathcal{Y}} \subseteq \mathcal{X}$, or $\bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$.
    - Either $\mathcal{Y}$ or the minimal subcluster of $\mathcal{X}$ (or $\bar{\mathcal{X}}$) containing $\bar{\mathcal{Y}}$ can be chosen as $\mathcal{Y}_0$.
  - Uniqueness: $\mathcal{Y}_0 \equiv \mathcal{Y}_1 \neq \mathcal{Y}_0 \Rightarrow \mathcal{Y}_0 \cup \mathcal{Y}_1 = S$ so we cannot have $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \mathcal{X}$ or $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \bar{\mathcal{X}}$.
- WLOG, let $\mathcal{Y}_0 \subseteq \mathcal{X}$. Then $\mathcal{Y}_0 \subseteq \mathcal{Y}_1 \subseteq \ldots \subseteq \mathcal{Y}_n = \mathcal{X}$ gives a path on T joining the vertices represented by $\mathcal{Y}$ and $\mathcal{X}$.
  - Thus T is **connected**.
- Moreover, **no cycle** in T.
  - No edge between $\mathcal{Y}_i$ and $\mathcal{Y}_j$ for $|i-j| > 1$ (by maximality).
  - $\mathcal{Z}_i \subseteq \mathcal{X}$ cannot be adjacent to $\mathcal{Z}_j \subseteq \bar{\mathcal{X}}$ unless $\mathcal{Z}_i = \mathcal{X}$ and $\mathcal{Z}_j = \bar{\mathcal{X}}$.

Proof of the connectivity and the cycle-freeness of T

- Let $\mathcal{X}, \bar{\mathcal{X}}$ be an arbitrary pair of complementary clusters.
- For any $\mathcal{Y} \in \mathbb{C}$, $\exists! \mathcal{Y}_0$ with $\mathcal{Y}_0 \equiv \mathcal{Y}$ such that $\mathcal{Y}_0 \subseteq \mathcal{X}$ or $\mathcal{Y}_0 \subseteq \bar{\mathcal{X}}$.
  - Existence: either $\mathcal{Y} \subseteq \mathcal{X}$, $\mathcal{Y} \subseteq \bar{\mathcal{X}}$, $\bar{\mathcal{Y}} \subseteq \mathcal{X}$, or $\bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$.
    - Either $\mathcal{Y}$ or the minimal subcluster of $\mathcal{X}$ (or $\bar{\mathcal{X}}$) containing $\bar{\mathcal{Y}}$ can be chosen as $\mathcal{Y}_0$.
  - Uniqueness: $\mathcal{Y}_0 \equiv \mathcal{Y}_1 \neq \mathcal{Y}_0 \Rightarrow \mathcal{Y}_0 \cup \mathcal{Y}_1 = S$ so we cannot have $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \mathcal{X}$ or $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \bar{\mathcal{X}}$.
- WLOG, let $\mathcal{Y}_0 \subseteq \mathcal{X}$. Then $\mathcal{Y}_0 \subseteq \mathcal{Y}_1 \subseteq \ldots \subseteq \mathcal{Y}_n = \mathcal{X}$ gives a path on T joining the vertices represented by $\mathcal{Y}$ and $\mathcal{X}$.
  - Thus T is **connected**.
- Moreover, **no cycle** in T.
  - No edge between $\mathcal{Y}_i$ and $\mathcal{Y}_j$ for $|i-j| > 1$ (by maximality).
  - $\mathcal{Z}_i \subseteq \mathcal{X}$ cannot be adjacent to $\mathcal{Z}_j \subseteq \bar{\mathcal{X}}$ unless $\mathcal{Z}_i = \mathcal{X}$ and $\mathcal{Z}_j = \bar{\mathcal{X}}$.

Proof of the connectivity and the cycle-freeness of T

- Let $\mathcal{X}, \bar{\mathcal{X}}$ be an arbitrary pair of complementary clusters.
- For any $\mathcal{Y} \in \mathbb{C}$, $\exists ! \mathcal{Y}_0$ with $\mathcal{Y}_0 \equiv \mathcal{Y}$ such that $\mathcal{Y}_0 \subseteq \mathcal{X}$ or $\mathcal{Y}_0 \subseteq \bar{\mathcal{X}}$.
  - Existence: either $\mathcal{Y} \subseteq \mathcal{X}$, $\mathcal{Y} \subseteq \bar{\mathcal{X}}$, $\bar{\mathcal{Y}} \subseteq \mathcal{X}$, or $\bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$.
    - ★ Either $\mathcal{Y}$ or the minimal subcluster of $\mathcal{X}$ (or $\bar{\mathcal{X}}$) containing $\bar{\mathcal{Y}}$ can be chosen as $\mathcal{Y}_0$.
  - Uniqueness: $\mathcal{Y}_0 \equiv \mathcal{Y}_1 \neq \mathcal{Y}_0 \Rightarrow \mathcal{Y}_0 \cup \mathcal{Y}_1 = S$ so we cannot have $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \mathcal{X}$ or $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \bar{\mathcal{X}}$.
- WLOG, let $\mathcal{Y}_0 \subseteq \mathcal{X}$. Then $\mathcal{Y}_0 \subseteq \mathcal{Y}_1 \subseteq \ldots \subseteq \mathcal{Y}_n = \mathcal{X}$ gives a path on T joining the vertices represented by $\mathcal{Y}$ and $\mathcal{X}$.
  - Thus T is **connected**.
- Moreover, **no cycle** in T.
  - No edge between $\mathcal{Y}_i$ and $\mathcal{Y}_j$ for $|i-j| > 1$ (by maximality).
  - $\mathcal{Z}_i \subseteq \mathcal{X}$ cannot be adjacent to $\mathcal{Z}_j \subseteq \bar{\mathcal{X}}$ unless $\mathcal{Z}_i = \mathcal{X}$ and $\mathcal{Z}_j = \bar{\mathcal{X}}$.

Proof of the connectivity and the cycle-freeness of T

- Let $\mathcal{X}, \bar{\mathcal{X}}$ be an arbitrary pair of complementary clusters.
- For any $\mathcal{Y} \in \mathbb{C}$, $\exists! \mathcal{Y}_0$ with $\mathcal{Y}_0 \equiv \mathcal{Y}$ such that $\mathcal{Y}_0 \subseteq \mathcal{X}$ or $\mathcal{Y}_0 \subseteq \bar{\mathcal{X}}$.
  - Existence: either $\mathcal{Y} \subseteq \mathcal{X}$, $\mathcal{Y} \subseteq \bar{\mathcal{X}}$, $\bar{\mathcal{Y}} \subseteq \mathcal{X}$, or $\bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$.
    - ★ Either $\mathcal{Y}$ or the minimal subcluster of $\mathcal{X}$ (or $\bar{\mathcal{X}}$) containing $\bar{\mathcal{Y}}$ can be chosen as $\mathcal{Y}_0$.
  - Uniqueness: $\mathcal{Y}_0 \equiv \mathcal{Y}_1 \neq \mathcal{Y}_0 \Rightarrow \mathcal{Y}_0 \cup \mathcal{Y}_1 = S$ so we cannot have $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \mathcal{X}$ or $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \bar{\mathcal{X}}$.
- WLOG, let $\mathcal{Y}_0 \subseteq \mathcal{X}$. Then $\mathcal{Y}_0 \subseteq \mathcal{Y}_1 \subseteq \ldots \subseteq \mathcal{Y}_n = \mathcal{X}$ gives a path on T joining the vertices represented by $\mathcal{Y}$ and $\mathcal{X}$.
  - Thus T is **connected**.
- Moreover, **no cycle** in T.
  - No edge between $\mathcal{Y}_i$ and $\mathcal{Y}_j$ for $|i - j| > 1$ (by maximality).
  - $\mathcal{Z}_i \subseteq \mathcal{X}$ cannot be adjacent to $\mathcal{Z}_j \subseteq \bar{\mathcal{X}}$ unless $\mathcal{Z}_i = \mathcal{X}$ and $\mathcal{Z}_j = \bar{\mathcal{X}}$.

Proof of the connectivity and the cycle-freeness of T

- Let $\mathcal{X}, \bar{\mathcal{X}}$ be an arbitrary pair of complementary clusters.
- For any $\mathcal{Y} \in \mathbb{C}$, $\exists! \mathcal{Y}_0$ with $\mathcal{Y}_0 \equiv \mathcal{Y}$ such that $\mathcal{Y}_0 \subseteq \mathcal{X}$ or $\mathcal{Y}_0 \subseteq \bar{\mathcal{X}}$.
  - Existence: either $\mathcal{Y} \subseteq \mathcal{X}$, $\mathcal{Y} \subseteq \bar{\mathcal{X}}$, $\bar{\mathcal{Y}} \subseteq \mathcal{X}$, or $\bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$.
    - Either $\mathcal{Y}$ or the minimal subcluster of $\mathcal{X}$ (or $\bar{\mathcal{X}}$) containing $\bar{\mathcal{Y}}$ can be chosen as $\mathcal{Y}_0$.
  - Uniqueness: $\mathcal{Y}_0 \equiv \mathcal{Y}_1 \neq \mathcal{Y}_0 \Rightarrow \mathcal{Y}_0 \cup \mathcal{Y}_1 = S$ so we cannot have $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \mathcal{X}$ or $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \bar{\mathcal{X}}$.
- WLOG, let $\mathcal{Y}_0 \subseteq \mathcal{X}$. Then $\mathcal{Y}_0 \subseteq \mathcal{Y}_1 \subseteq \ldots \subseteq \mathcal{Y}_n = \mathcal{X}$ gives a path on T joining the vertices represented by $\mathcal{Y}$ and $\mathcal{X}$.
  - Thus T is **connected**.
- Moreover, **no cycle** in T.
  - No edge between $\mathcal{Y}_i$ and $\mathcal{Y}_j$ for $|i-j| > 1$ (by maximality).
  - $\mathcal{Z}_i \subseteq \mathcal{X}$ cannot be adjacent to $\mathcal{Z}_j \subseteq \bar{\mathcal{X}}$ unless $\mathcal{Z}_i = \mathcal{X}$ and $\mathcal{Z}_j = \bar{\mathcal{X}}$.

Proof of the connectivity and the cycle-freeness of T

- Let $\mathcal{X}, \bar{\mathcal{X}}$ be an arbitrary pair of complementary clusters.
- For any $\mathcal{Y} \in \mathbb{C}$, $\exists! \mathcal{Y}_0$ with $\mathcal{Y}_0 \equiv \mathcal{Y}$ such that $\mathcal{Y}_0 \subseteq \mathcal{X}$ or $\mathcal{Y}_0 \subseteq \bar{\mathcal{X}}$.
  - Existence: either $\mathcal{Y} \subseteq \mathcal{X}$, $\mathcal{Y} \subseteq \bar{\mathcal{X}}$, $\bar{\mathcal{Y}} \subseteq \mathcal{X}$, or $\bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$.
    - ★ Either $\mathcal{Y}$ or the minimal subcluster of $\mathcal{X}$ (or $\bar{\mathcal{X}}$) containing $\bar{\mathcal{Y}}$ can be chosen as $\mathcal{Y}_0$.
  - Uniqueness: $\mathcal{Y}_0 \equiv \mathcal{Y}_1 \neq \mathcal{Y}_0 \Rightarrow \mathcal{Y}_0 \cup \mathcal{Y}_1 = S$ so we cannot have $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \mathcal{X}$ or $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \bar{\mathcal{X}}$.
- WLOG, let $\mathcal{Y}_0 \subseteq \mathcal{X}$. Then $\mathcal{Y}_0 \subseteq \mathcal{Y}_1 \subseteq \ldots \subseteq \mathcal{Y}_n = \mathcal{X}$ gives a path on T joining the vertices represented by $\mathcal{Y}$ and $\mathcal{X}$.
  - Thus T is **connected**.
- Moreover, **no cycle** in T.
  - No edge between $\mathcal{Y}_i$ and $\mathcal{Y}_j$ for $|i-j| > 1$ (by maximality).
  - $\mathcal{Z}_i \subseteq \mathcal{X}$ cannot be adjacent to $\mathcal{Z}_j \subseteq \bar{\mathcal{X}}$ unless $\mathcal{Z}_i = \mathcal{X}$ and $\mathcal{Z}_j = \bar{\mathcal{X}}$.

Proof of the connectivity and the cycle-freeness of T

- Let $\mathcal{X}, \bar{\mathcal{X}}$ be an arbitrary pair of complementary clusters.
- For any $\mathcal{Y} \in \mathbb{C}$, $\exists! \mathcal{Y}_0$ with $\mathcal{Y}_0 \equiv \mathcal{Y}$ such that $\mathcal{Y}_0 \subseteq \mathcal{X}$ or $\mathcal{Y}_0 \subseteq \bar{\mathcal{X}}$.
  - Existence: either $\mathcal{Y} \subseteq \mathcal{X}$, $\mathcal{Y} \subseteq \bar{\mathcal{X}}$, $\bar{\mathcal{Y}} \subseteq \mathcal{X}$, or $\bar{\mathcal{Y}} \subseteq \bar{\mathcal{X}}$.
    - ★ Either $\mathcal{Y}$ or the minimal subcluster of $\mathcal{X}$ (or $\bar{\mathcal{X}}$) containing $\bar{\mathcal{Y}}$ can be chosen as $\mathcal{Y}_0$.
  - Uniqueness: $\mathcal{Y}_0 \equiv \mathcal{Y}_1 \neq \mathcal{Y}_0 \Rightarrow \mathcal{Y}_0 \cup \mathcal{Y}_1 = S$ so we cannot have $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \mathcal{X}$ or $\mathcal{Y}_0, \mathcal{Y}_1 \subseteq \bar{\mathcal{X}}$.
- WLOG, let $\mathcal{Y}_0 \subseteq \mathcal{X}$. Then $\mathcal{Y}_0 \subseteq \mathcal{Y}_1 \subseteq \ldots \subseteq \mathcal{Y}_n = \mathcal{X}$ gives a path on T joining the vertices represented by $\mathcal{Y}$ and $\mathcal{X}$.
  - Thus T is **connected**.
- Moreover, **no cycle** in T.
  - No edge between $\mathcal{Y}_i$ and $\mathcal{Y}_j$ for $|i-j| > 1$ (by maximality).
  - $\mathcal{Z}_i \subseteq \mathcal{X}$ cannot be adjacent to $\mathcal{Z}_j \subseteq \bar{\mathcal{X}}$ unless $\mathcal{Z}_i = \mathcal{X}$ and $\mathcal{Z}_j = \bar{\mathcal{X}}$.

The one-to-one correspondence with complementary cluster pairs

- $\mathcal{Y} \equiv \mathcal{X}$ and $\bar{\mathcal{Y}} \equiv \bar{\mathcal{X}} \Rightarrow \mathcal{Y} = \mathcal{X}$.
- So the edges of T are in one-to-one correspondence with the complementary cluster pairs $\mathcal{X}, \bar{\mathcal{X}}$.
- Hence the clusters in a given equivalence class correspond in a one-to-one manner to the edges incident with this equivalence class (regarded as a vertex).

The one-to-one correspondence with complementary cluster pairs

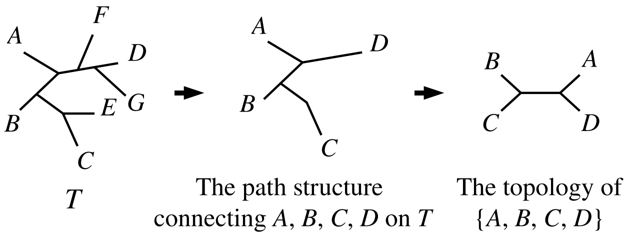
- $\mathcal{Y} \equiv \mathcal{X}$ and $\bar{\mathcal{Y}} \equiv \bar{\mathcal{X}} \Rightarrow \mathcal{Y} = \mathcal{X}$.
- So the edges of T are in one-to-one correspondence with the complementary cluster pairs $\mathcal{X}, \bar{\mathcal{X}}$.
- Hence the clusters in a given equivalence class correspond in a one-to-one manner to the edges incident with this equivalence class (regarded as a vertex).

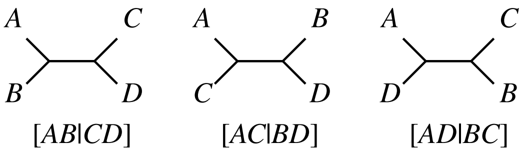
Is every taxon A represented by a unique equivalence class $\rho(A)$?

- Yes.
- $\{A\} \in \mathbb{C}$ for all $A \in S$ and $\rho(A) = \{A\}$.

Outline

Quartet topologies





Quartet topologies (contd.)

- Let Q be a set of quartet topologies over S.
- Assume that Q is complete: every four taxa in S has exactly one quartet topology in Q.

Translation between clusters and quartet topologies

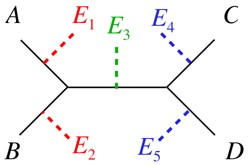
- $[AB|CD] \in Q$ if and only if $A, B \in Y$ and $C, D \in \bar{Y}$ for some cluster $Y \in \mathbb{C}$.
- Y is a cluster of size at least two if and only if $Y \neq S$ and $[AB|CD] \in Q$ for all $A, B \in Y$ and for all $C, D \in \bar{Y}$.

The substitution property

$[AB|CD] \in Q \;\Rightarrow$

- $\star$ $[AB|CE], [AB|DE] \in Q$ or $[AE|CD], [BE|CD] \in Q$

for any $E \in S \setminus \{A, B, C, D\}$.



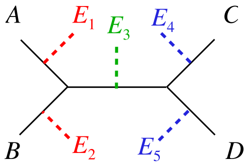
- We say a quintet $q = \{s_1, s_2, s_3, s_4, s_5\}$ is consistent if for every bijection $\sigma : q \longrightarrow \{A, B, C, D, E\}$, we have $[AB|CD] \in Q \Rightarrow [AB|CE], [AB|DE] \in Q$ or $[AE|CD], [BE|CD] \in Q$.

Transitive property

Lemma 1

If every quintet over S satisfies the substitution property, then for every quintet $\{A, B, C, D, E\}$, we have

$$[AB|CD], [AB|DE] \in Q \Rightarrow [AB|CE] \in Q.$$



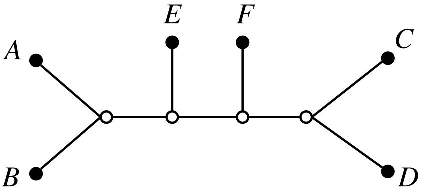
Q is tree-like:

$\exists$ an evolutionary tree T whose set of induced quartet topologies is exactly Q.

Quartet topologies and tree-likeness

Proposition 2

Q is tree-like $\Leftrightarrow$ every quintet over S is consistent.

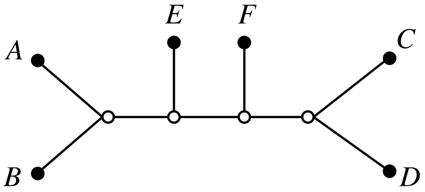


- Assume that $Q = \{[AB|CD], [AB|CE], [AB|CF], [AB|DE], [AB|DF], [AB|EF], [AE|CD], [AF|CD], [AE|CF], [AD|EF], [BC|DE], [BF|CD], [BE|CF], [BE|DF], [CD|EF]\}$.

Quartet topologies and tree-likeness

Proposition 2

Q is tree-like $\Leftrightarrow$ every quintet over S is consistent.

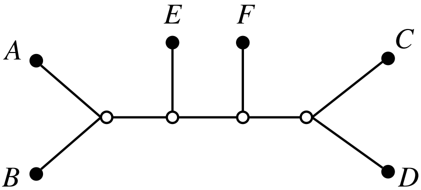


- Assume that $Q = \{[AB|CD], [AB|CE], [AB|CF], [AB|DE], [AB|DF], [AB|EF], [AE|CD], [AF|CD], [AE|CF], [AD|EF], [BC|DE], [BF|CD], [BE|CF], [BE|DF], [CD|EF]\}$.

Quartet topologies and tree-likeness

Proposition 2

Q is tree-like $\Leftrightarrow$ every quintet over S is consistent.

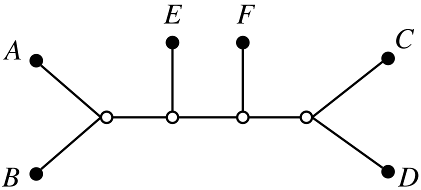


- Assume that $Q = \{[AB|CD], [AB|CE], [AB|CF], [AB|DE], [AB|DF], [AB|EF], [AE|CD], [AF|CD], [AE|CF], [AD|EF], [BC|DE], [BF|CD], [BE|CF], [BE|DF], [CD|EF]\}$.

Quartet topologies and tree-likeness

Proposition 2

Q is tree-like $\Leftrightarrow$ every quintet over S is consistent.

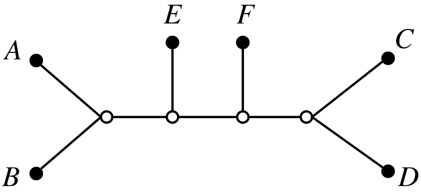


- Assume that $Q = \{[AB|CD], [AB|CE], [AB|CF], [AB|DE], [AB|DF], [AB|EF], [AE|CD], [AF|CD], [AE|CF], [AD|EF], [BC|DE], [BF|CD], [BE|CF], [BE|DF], [CD|EF]\}$.

Quartet topologies and tree-likeness

Proposition 2

Q is tree-like $\Leftrightarrow$ every quintet over S is consistent.



- Assume that $Q = \{[AB|CD], [AB|CE], [AB|CF], [AB|DE], [AB|DF], [AB|EF], [AE|CD], [AF|CD], [AE|CF], [AD|EF], [BC|DE], [BF|CD], [BE|CF], [BE|DF], [CD|EF]\}$.

Proof of Proposition 2

- The if-part is clearly true.
- We construct abstract clusters w.r.t. Q as follows.
  - Construct clusters $\{A\}$ and their complementary clusters $S \setminus \{A\}$ for each $A \in S$. (Trivial clusters)
  - Construct a cluster $\mathcal{Y}$ w.r.t. Q when $1 < |\mathcal{Y}| < n - 1$ and $[AB|CD] \in Q$ for all $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
  - $\mathcal{Y}$ is a cluster $\Leftrightarrow \bar{\mathcal{Y}}$ is a cluster.
- Any two clusters $\mathcal{X}, \mathcal{Y}$ w.r.t. Q are compatible.
  - Assume $A \in \mathcal{X} \cap \mathcal{Y}$, $B \in \mathcal{X} \cap \bar{\mathcal{Y}}$, $C \in \bar{\mathcal{X}} \cap \mathcal{Y}$, $D \in \bar{\mathcal{X}} \cap \bar{\mathcal{Y}}$. We have $[AB|CD], [AC|BD] \in Q$ ($\Rightarrow\Leftarrow$)

Proof of Proposition 2

- The if-part is clearly true.
- We construct abstract clusters w.r.t. Q as follows.
  - Construct clusters $\{A\}$ and their complementary clusters $S \setminus \{A\}$ for each $A \in S$. (Trivial clusters)
  - Construct a cluster $\mathcal{Y}$ w.r.t. Q when $1 < |\mathcal{Y}| < n - 1$ and $[AB|CD] \in Q$ for all $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
  - $\mathcal{Y}$ is a cluster $\Leftrightarrow \bar{\mathcal{Y}}$ is a cluster.
- Any two clusters $\mathcal{X}, \mathcal{Y}$ w.r.t. Q are compatible.
  - Assume $A \in \mathcal{X} \cap \mathcal{Y}$, $B \in \mathcal{X} \cap \bar{\mathcal{Y}}$, $C \in \bar{\mathcal{X}} \cap \mathcal{Y}$, $D \in \bar{\mathcal{X}} \cap \bar{\mathcal{Y}}$. We have $[AB|CD], [AC|BD] \in Q$ ($\Rightarrow\Leftarrow$)

Proof of Proposition 2

- The if-part is clearly true.
- We construct abstract clusters w.r.t. Q as follows.
  - Construct clusters $\{A\}$ and their complementary clusters $S \setminus \{A\}$ for each $A \in S$. (Trivial clusters)
  - Construct a cluster $\mathcal{Y}$ w.r.t. Q when $1 < |\mathcal{Y}| < n - 1$ and $[AB|CD] \in Q$ for all $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
  - $\mathcal{Y}$ is a cluster $\Leftrightarrow$ $\bar{\mathcal{Y}}$ is a cluster.
- Any two clusters $\mathcal{X}, \mathcal{Y}$ w.r.t. Q are compatible.
  - Assume $A \in \mathcal{X} \cap \mathcal{Y}$, $B \in \mathcal{X} \cap \bar{\mathcal{Y}}$, $C \in \bar{\mathcal{X}} \cap \mathcal{Y}$, $D \in \bar{\mathcal{X}} \cap \bar{\mathcal{Y}}$. We have $[AB|CD], [AC|BD] \in Q$ $(\Rightarrow\Leftarrow)$

Proof of Proposition 2

- The if-part is clearly true.
- We construct abstract clusters w.r.t. Q as follows.
  - Construct clusters $\{A\}$ and their complementary clusters $S \setminus \{A\}$ for each $A \in S$. (Trivial clusters)
  - Construct a cluster $\mathcal{Y}$ w.r.t. Q when $1 < |\mathcal{Y}| < n - 1$ and $[AB|CD] \in Q$ for all $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
  - $\mathcal{Y}$ is a cluster $\Leftrightarrow \bar{\mathcal{Y}}$ is a cluster.
- Any two clusters $\mathcal{X}, \mathcal{Y}$ w.r.t. Q are compatible.
  - Assume $A \in \mathcal{X} \cap \mathcal{Y}$, $B \in \mathcal{X} \cap \bar{\mathcal{Y}}$, $C \in \bar{\mathcal{X}} \cap \mathcal{Y}$, $D \in \bar{\mathcal{X}} \cap \bar{\mathcal{Y}}$. We have $[AB|CD], [AC|BD] \in Q$ ($\Rightarrow\!\Leftarrow$)

Proof of Proposition 2

- The if-part is clearly true.
- We construct abstract clusters w.r.t. Q as follows.
  - Construct clusters $\{A\}$ and their complementary clusters $S \setminus \{A\}$ for each $A \in S$. (Trivial clusters)
  - Construct a cluster $\mathcal{Y}$ w.r.t. Q when $1 < |\mathcal{Y}| < n - 1$ and $[AB|CD] \in Q$ for all $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
  - $\mathcal{Y}$ is a cluster $\Leftrightarrow \bar{\mathcal{Y}}$ is a cluster.
- Any two clusters $\mathcal{X}, \mathcal{Y}$ w.r.t. Q are compatible.
  - Assume $A \in \mathcal{X} \cap \mathcal{Y}$, $B \in \mathcal{X} \cap \bar{\mathcal{Y}}$, $C \in \bar{\mathcal{X}} \cap \mathcal{Y}$, $D \in \bar{\mathcal{X}} \cap \bar{\mathcal{Y}}$. We have $[AB|CD], [AC|BD] \in Q$ ($\Rightarrow\!\Leftarrow$)

Proof of Proposition 2

- The if-part is clearly true.
- We construct abstract clusters w.r.t. Q as follows.
  - Construct clusters $\{A\}$ and their complementary clusters $S \setminus \{A\}$ for each $A \in S$. (Trivial clusters)
  - Construct a cluster $\mathcal{Y}$ w.r.t. Q when $1 < |\mathcal{Y}| < n - 1$ and $[AB|CD] \in Q$ for all $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
  - $\mathcal{Y}$ is a cluster $\Leftrightarrow$ $\bar{\mathcal{Y}}$ is a cluster.
- Any two clusters $\mathcal{X}, \mathcal{Y}$ w.r.t. Q are compatible.
  - Assume $A \in \mathcal{X} \cap \mathcal{Y}$, $B \in \mathcal{X} \cap \bar{\mathcal{Y}}$, $C \in \bar{\mathcal{X}} \cap \mathcal{Y}$, $D \in \bar{\mathcal{X}} \cap \bar{\mathcal{Y}}$. We have $[AB|CD], [AC|BD] \in Q$ ($\Rightarrow\!\Leftarrow$)

Proof of Proposition 2 (contd.)

Construct the corresponding clusters and show that they coincide with Q.

- Assume that $[AB|CD] \in Q$ and let $\mathcal{Y} = \{E \mid\ [AE|CD] \in Q \text{ or } [BE|CD] \in Q\}$.
  - $[AE|CD] \in Q \Leftrightarrow [BE|CD] \in Q$ (transitivity).
- $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
- If $E \in \mathcal{Y}, F \in \bar{\mathcal{Y}}$, then $[AE|CF] \in Q$ (by the substitution property & $[AF|CD] \notin Q$ since $F \notin \mathcal{Y}$).
- Hence for taxa $M_1, M_2 \in \mathcal{Y}$ and $N_1, N_2 \in \bar{\mathcal{Y}}$ we have $[AM_i|CN_j] \in Q$ for $i, j = 1, 2$.
- By transitivity, $[M_1 M_2|CN_j] \in Q$ for $j = 1, 2$, and further $[M_1 M_2|N_1 N_2] \in Q$.

Proof of Proposition 2 (contd.)

Construct the corresponding clusters and show that they coincide with Q.

- Assume that $[AB|CD] \in Q$ and let $\mathcal{Y} = \{E \mid \ [AE|CD] \in Q \text{ or } [BE|CD] \in Q\}$.
  - $[AE|CD] \in Q \Leftrightarrow [BE|CD] \in Q$ (transitivity).
- $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
- If $E \in \mathcal{Y}, F \in \bar{\mathcal{Y}}$, then $[AE|CF] \in Q$ (by the substitution property & $[AF|CD] \notin Q$ since $F \notin \mathcal{Y}$).
- Hence for taxa $M_1, M_2 \in \mathcal{Y}$ and $N_1, N_2 \in \bar{\mathcal{Y}}$ we have $[AM_i|CN_j] \in Q$ for $i, j = 1, 2$.
- By transitivity, $[M_1M_2|CN_j] \in Q$ for $j = 1, 2$, and further $[M_1M_2|N_1N_2] \in Q$.

Proof of Proposition 2 (contd.)

Construct the corresponding clusters and show that they coincide with Q.

- Assume that $[AB|CD] \in Q$ and let $\mathcal{Y} = \{E \mid [AE|CD] \in Q$ or $[BE|CD] \in Q\}$.
  - $[AE|CD] \in Q \Leftrightarrow [BE|CD] \in Q$ (transitivity).
- $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
- If $E \in \mathcal{Y}, F \in \bar{\mathcal{Y}}$, then $[AE|CF] \in Q$ (by the substitution property & $[AF|CD] \notin Q$ since $F \notin \mathcal{Y}$).
- Hence for taxa $M_1, M_2 \in \mathcal{Y}$ and $N_1, N_2 \in \bar{\mathcal{Y}}$ we have $[AM_i|CN_j] \in Q$ for $i, j = 1, 2$.
- By transitivity, $[M_1 M_2|CN_j] \in Q$ for $j = 1, 2$, and further $[M_1 M_2|N_1 N_2] \in Q$.

Proof of Proposition 2 (contd.)

Construct the corresponding clusters and show that they coincide with Q.

- Assume that $[AB|CD] \in Q$ and let $\mathcal{Y} = \{E \mid [AE|CD] \in Q$ or $[BE|CD] \in Q\}$.
  - $[AE|CD] \in Q \Leftrightarrow [BE|CD] \in Q$ (transitivity).
- $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
- If $E \in \mathcal{Y}, F \in \bar{\mathcal{Y}}$, then $[AE|CF] \in Q$ (by the substitution property & $[AF|CD] \notin Q$ since $F \notin \mathcal{Y}$).
- Hence for taxa $M_1, M_2 \in \mathcal{Y}$ and $N_1, N_2 \in \bar{\mathcal{Y}}$ we have $[AM_i|CN_j] \in Q$ for $i, j = 1, 2$.
- By transitivity, $[M_1M_2|CN_j] \in Q$ for $j = 1, 2$, and further $[M_1M_2|N_1N_2] \in Q$.

Proof of Proposition 2 (contd.)

Construct the corresponding clusters and show that they coincide with Q.

- Assume that $[AB|CD] \in Q$ and let $\mathcal{Y} = \{E \mid [AE|CD] \in Q \text{ or } [BE|CD] \in Q\}$.
  - $[AE|CD] \in Q \Leftrightarrow [BE|CD] \in Q$ (transitivity).
- $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
- If $E \in \mathcal{Y}$, $F \in \bar{\mathcal{Y}}$, then $[AE|CF] \in Q$ (by the substitution property & $[AF|CD] \notin Q$ since $F \notin \mathcal{Y}$).
- Hence for taxa $M_1, M_2 \in \mathcal{Y}$ and $N_1, N_2 \in \bar{\mathcal{Y}}$ we have $[AM_i|CN_j] \in Q$ for $i, j = 1, 2$.
- By transitivity, $[M_1M_2|CN_j] \in Q$ for $j = 1, 2$, and further $[M_1M_2|N_1N_2] \in Q$.

Proof of Proposition 2 (contd.)

Construct the corresponding clusters and show that they coincide with Q.

- Assume that $[AB|CD] \in Q$ and let $\mathcal{Y} = \{E \mid \ [AE|CD] \in Q$ or $[BE|CD] \in Q\}$.
  - $[AE|CD] \in Q \Leftrightarrow [BE|CD] \in Q$ (transitivity).
- $A, B \in \mathcal{Y}$ and $C, D \in \bar{\mathcal{Y}}$.
- If $E \in \mathcal{Y}, F \in \bar{\mathcal{Y}}$, then $[AE|CF] \in Q$ (by the substitution property & $[AF|CD] \notin Q$ since $F \notin \mathcal{Y}$).
- Hence for taxa $M_1, M_2 \in \mathcal{Y}$ and $N_1, N_2 \in \bar{\mathcal{Y}}$ we have $[AM_i|CN_j] \in Q$ for $i, j = 1, 2$.
- By transitivity, $[M_1 M_2|CN_j] \in Q$ for $j = 1, 2$, and further $[M_1 M_2|N_1 N_2] \in Q$.

An improved result...

Proposition 3

Given any fixed taxon F, then:
Q is tree-like $\Leftrightarrow$ every quintet containing F is consistent.

Proof of Proposition 3

- Assume that $[AB|CD] \in Q$ and let E be any taxon in $S \setminus \{A, B, C, D\}$.
- Wish to show: either $[AE|CD] \in Q$ or $[AB|CE] \in Q$.
  - By the assumption, either $[AB|CF] \in Q$ or $[AF|CD] \in Q$ is true.
  - $\Rightarrow [AB|DF] \in Q$.
  - $\Rightarrow$ either $[AB|EF] \in Q$ or $[AE|DF] \in Q$.
    - ★ If $[AB|EF] \in Q$, so does $[AB|CE] \in Q$ (transitivity & $[AB|CF] \in Q$).
    - ★ Otherwise, (i.e., $[AE|DF] \in Q$).
      $\therefore [AB|CE] \in Q$ or $[AE|CF] \in Q$ ($\because [AB|CF] \in Q$).

      The latter with $[AE|DF] \in Q$ gives $[AE|CD] \in Q$.

Proof of Proposition 3

- Assume that $[AB|CD] \in Q$ and let E be any taxon in $S \setminus \{A, B, C, D\}$.
- Wish to show: either $[AE|CD] \in Q$ or $[AB|CE] \in Q$.
  - By the assumption, either $[AB|CF] \in Q$ or $[AF|CD] \in Q$ is true.
  - $\Rightarrow [AB|DF] \in Q$.
  - $\Rightarrow$ either $[AB|EF] \in Q$ or $[AE|DF] \in Q$.
    - ★ If $[AB|EF] \in Q$, so does $[AB|CE] \in Q$ (transitivity & $[AB|CF] \in Q$).
    - ★ Otherwise, (i.e., $[AE|DF] \in Q$).
      $\therefore [AB|CE] \in Q$ or $[AE|CF] \in Q$ ($\because [AB|CF] \in Q$).

      The latter with $[AE|DF] \in Q$ gives $[AE|CD] \in Q$.

Proof of Proposition 3

- Assume that $[AB|CD] \in Q$ and let E be any taxon in $S \setminus \{A, B, C, D\}$.
- Wish to show: either $[AE|CD] \in Q$ or $[AB|CE] \in Q$.
  - By the assumption, either $[AB|CF] \in Q$ or $[AF|CD] \in Q$ is true.
  - $\Rightarrow [AB|DF] \in Q$.
  - $\Rightarrow$ either $[AB|EF] \in Q$ or $[AE|DF] \in Q$.
    - ★ If $[AB|EF] \in Q$, so does $[AB|CE] \in Q$ (transitivity & $[AB|CF] \in Q$).
    - ★ Otherwise, (i.e., $[AE|DF] \in Q$).
      $\because [AB|CE] \in Q$ or $[AE|CF] \in Q$ ($\because [AB|CF] \in Q$).

      The latter with $[AE|DF] \in Q$ gives $[AE|CD] \in Q$.

Proof of Proposition 3

- Assume that $[AB|CD] \in Q$ and let E be any taxon in $S \setminus \{A, B, C, D\}$.
- Wish to show: either $[AE|CD] \in Q$ or $[AB|CE] \in Q$.
  - By the assumption, either $[AB|CF] \in Q$ or $[AF|CD] \in Q$ is true.
  - $\Rightarrow [AB|DF] \in Q$.
  - $\Rightarrow$ either $[AB|EF] \in Q$ or $[AE|DF] \in Q$.
    - ★ If $[AB|EF] \in Q$, so does $[AB|CE] \in Q$ (transitivity & $[AB|CF] \in Q$).
    - ★ Otherwise, (i.e., $[AE|DF] \in Q$).
      $\therefore [AB|CE] \in Q$ or $[AE|CF] \in Q$ ($\because [AB|CF] \in Q$).

      The latter with $[AE|DF] \in Q$ gives $[AE|CD] \in Q$.

Proof of Proposition 3

- Assume that $[AB|CD] \in Q$ and let E be any taxon in $S \setminus \{A, B, C, D\}$.
- Wish to show: either $[AE|CD] \in Q$ or $[AB|CE] \in Q$.
  - By the assumption, either $[AB|CF] \in Q$ or $[AF|CD] \in Q$ is true.
  - $\Rightarrow [AB|DF] \in Q$.
  - $\Rightarrow$ either $[AB|EF] \in Q$ or $[AE|DF] \in Q$.
    - ★ If $[AB|EF] \in Q$, so does $[AB|CE] \in Q$ (transitivity & $[AB|CF] \in Q$).
    - ★ Otherwise, (i.e., $[AE|DF] \in Q$).
      $\because$ $[AB|CE] \in Q$ or $[AE|CF] \in Q$ ($\because$ $[AB|CF] \in Q$).

      The latter with $[AE|DF] \in Q$ gives $[AE|CD] \in Q$.

Proof of Proposition 3

- Assume that $[AB|CD] \in Q$ and let E be any taxon in $S \setminus \{A, B, C, D\}$.
- Wish to show: either $[AE|CD] \in Q$ or $[AB|CE] \in Q$.
  - By the assumption, either $[AB|CF] \in Q$ or $[AF|CD] \in Q$ is true.
  - $\Rightarrow [AB|DF] \in Q$.
  - $\Rightarrow$ either $[AB|EF] \in Q$ or $[AE|DF] \in Q$.
    - ★ If $[AB|EF] \in Q$, so does $[AB|CE] \in Q$ (transitivity & $[AB|CF] \in Q$).
    - ★ Otherwise, (i.e., $[AE|DF] \in Q$).
      $\because [AB|CE] \in Q$ or $[AE|CF] \in Q$ ($\because [AB|CF] \in Q$).

      The latter with $[AE|DF] \in Q$ gives $[AE|CD] \in Q$.

Proof of Proposition 3

- Assume that $[AB|CD] \in Q$ and let E be any taxon in $S \setminus \{A, B, C, D\}$.
- Wish to show: either $[AE|CD] \in Q$ or $[AB|CE] \in Q$.
  - By the assumption, either $[AB|CF] \in Q$ or $[AF|CD] \in Q$ is true.
  - $\Rightarrow [AB|DF] \in Q$.
  - $\Rightarrow$ either $[AB|EF] \in Q$ or $[AE|DF] \in Q$.
    - ⋆ If $[AB|EF] \in Q$, so does $[AB|CE] \in Q$ (transitivity & $[AB|CF] \in Q$).
    - ⋆ Otherwise, (i.e., $[AE|DF] \in Q$).
      $\because [AB|CE] \in Q$ or $[AE|CF] \in Q$ ($\because [AB|CF] \in Q$).

      The latter with $[AE|DF] \in Q$ gives $[AE|CD] \in Q$.

Proof of Proposition 3

- Assume that $[AB|CD] \in Q$ and let E be any taxon in $S \setminus \{A, B, C, D\}$.
- Wish to show: either $[AE|CD] \in Q$ or $[AB|CE] \in Q$.
  - By the assumption, either $[AB|CF] \in Q$ or $[AF|CD] \in Q$ is true.
  - $\Rightarrow [AB|DF] \in Q$.
  - $\Rightarrow$ either $[AB|EF] \in Q$ or $[AE|DF] \in Q$.
    - ★ If $[AB|EF] \in Q$, so does $[AB|CE] \in Q$ (transitivity & $[AB|CF] \in Q$).
    - ★ Otherwise, (i.e., $[AE|DF] \in Q$).
      $\because [AB|CE] \in Q$ or $[AE|CF] \in Q$ ($\because [AB|CF] \in Q$).

      The latter with $[AE|DF] \in Q$ gives $[AE|CD] \in Q$.

Outline

1. Introduction

2. Cluster and tree-likeness

3. Quartet topologies and tree-likeness

4. Conclusions

Conclusions

- The arguments in the paper are very unclear.
- I felt painful when reading this paper.

Thank you!