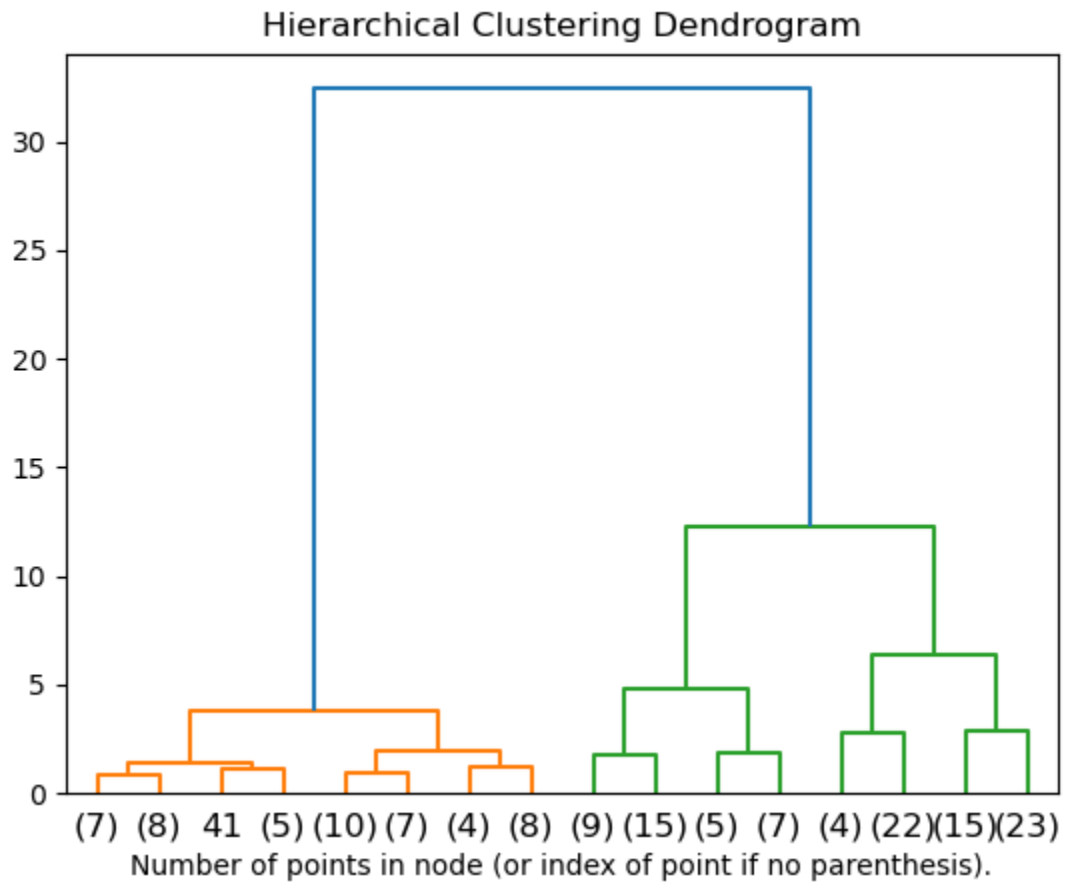
Agglomerative Clustering

Dr. Kalidas Y., IIT Tirupati

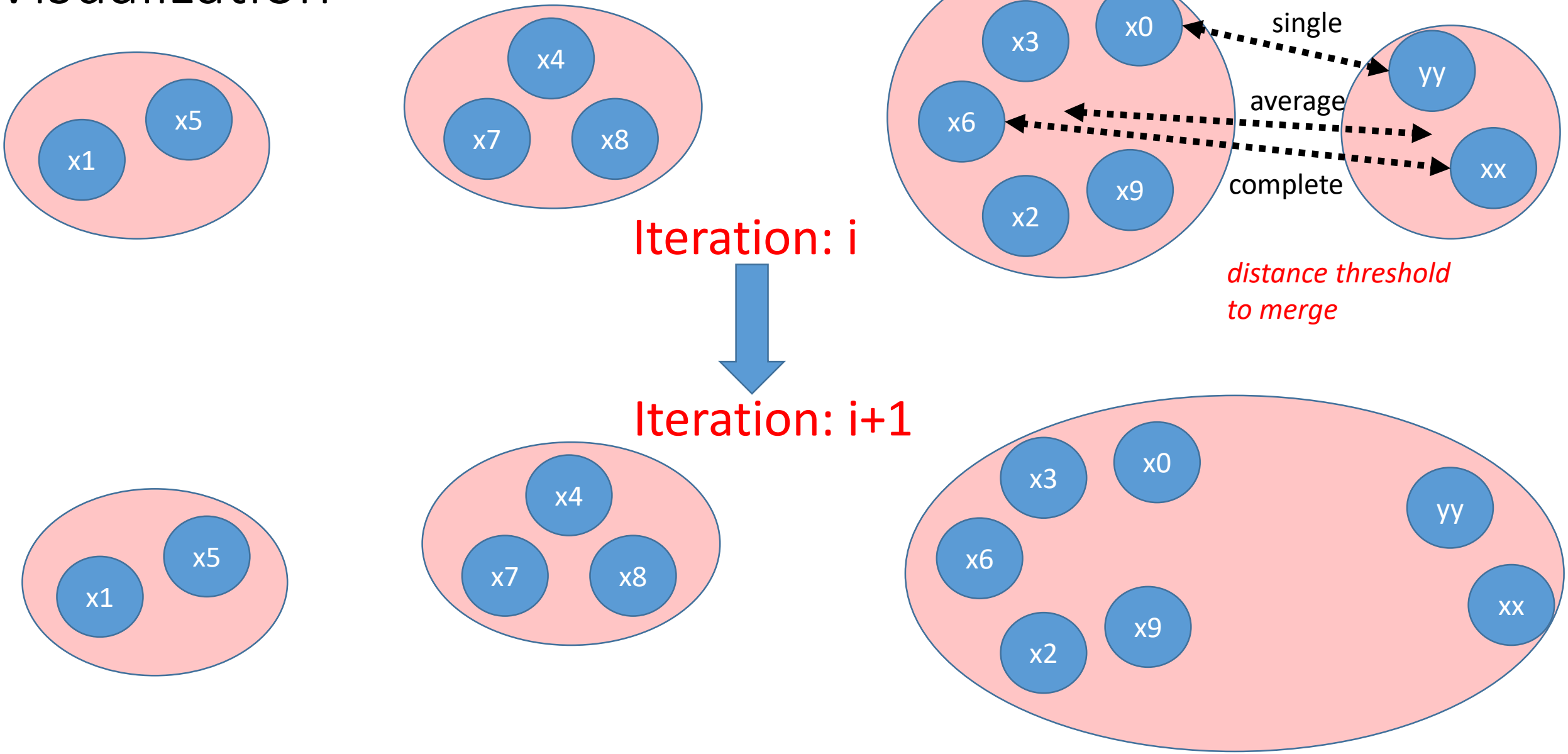
113) key phrase… “Agglomerative/Hierarchical clustering”



Brief Sketch of the Algorithm

- Input: A set of data points {xi}
- Process: Agglomerative Clustering
  - Recursively partition the elements
- Output: A set of clusters C1, C2, …, Ck
  - Each Ci is a set of points
  - Ci and Cj are disjoint

Visualization



Agglomerative clustering – Scope and Limitation

- As with K-Means, this clustering also mainly works when clusters are neatly apart
- A problematic scenario:
  - For example, if two concentric circles are connected by a small trail of points, Agglomerative clustering will merge them!
  - Another example, imagine two blobs of points
    - Connect them by a very thin line of points
    - That's the failure example of agglomerative clustering, it will merge the blobs incorrectly!
- Agglomerative clustering is used widely in life-science discipline
- Phylogenetic analysis
- Gene sequence alignment and relationship between biological species are studied by looking at these clusters