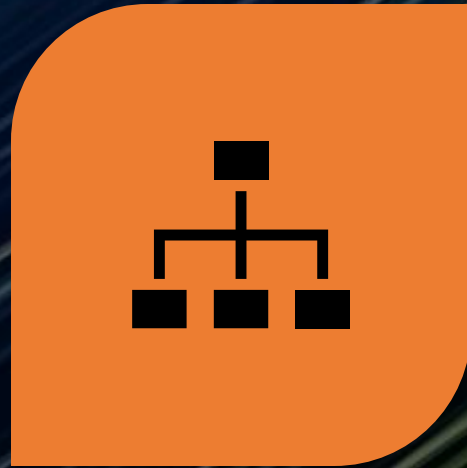


# Hierarchical Clustering Algorithms

Alok Yadav

# Hierarchical Clustering Algorithms



**AGGLOMERATIVE  
HIERARCHICAL CLUSTERING**



**DIVISIVE HIERARCHICAL  
CLUSTERING**



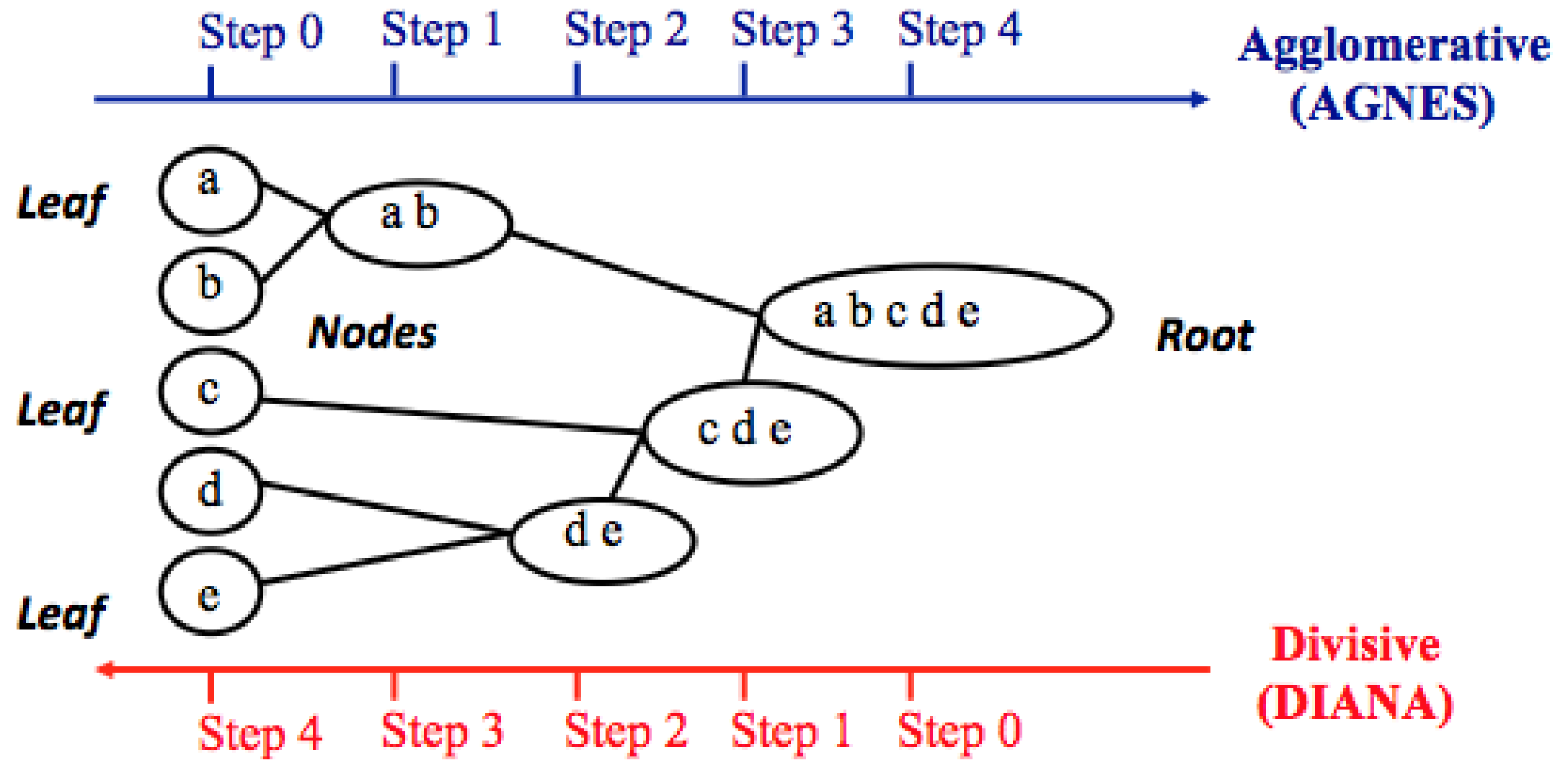
# Agglomerative clustering:

- It's also known as **AGNES** (Agglomerative Nesting).
- It works in a bottom-up manner.
- That is, each object is initially considered as a single-element cluster (leaf).
- At each step of the algorithm, the two clusters that are the most similar are combined into a new bigger cluster (nodes).
- This procedure is iterated until all points are member of just one single big cluster (root) (see figure below).
- The result is a tree which can be plotted as a dendrogram.



## Divisive hierarchical clustering:

- It's also known as **DIANA** (Divise Analysis) and it works in a top-down manner.
- The algorithm is an inverse order of AGNES.
- It begins with the root, in which all objects are included in a single cluster.
- At each step of iteration, the most heterogeneous cluster is divided into two.
- The process is iterated until all objects are in their own cluster (see figure below).



# Hierarchical Clustering with R

---

There are different functions available in R for computing hierarchical clustering. The commonly used functions are:

---

**hclust** [in stats package] and **agnes** [in cluster package] for agglomerative hierarchical clustering (HC)

---

**diana** [in cluster package] for divisive HC

# Agglomerative Hierarchical Clustering

- # Dissimilarity matrix
- `d <- dist(df, method = "euclidean")`
- 
- # Hierarchical clustering using Complete Linkage
- `hc1 <- hclust(d, method = "complete" )`
- 
- # Plot the obtained dendrogram
- `plot(hc1, cex = 0.6, hang = -1)`

## Cluster Dendrogram

