Hierarchical clustering is a method used to group similar data points into clusters. One of the commonly used approaches is agglomerative hierarchical clustering. While expressing the entire algorithm in a mathematical equation can be complex due to its iterative and step-wise nature, I can provide you with a simplified representation of the linkage criterion used in agglomerative hierarchical clustering.

Let's denote:

* *D* as the distance matrix between data points.
* *n* as the number of data points.

The basic idea is to iteratively merge the two closest clusters until only one cluster remains. The linkage criterion defines the distance between clusters at each step. One of the common linkage criteria is the complete linkage, which considers the maximum distance between points in different clusters. The formula for complete linkage between clusters *i* and *j* is:

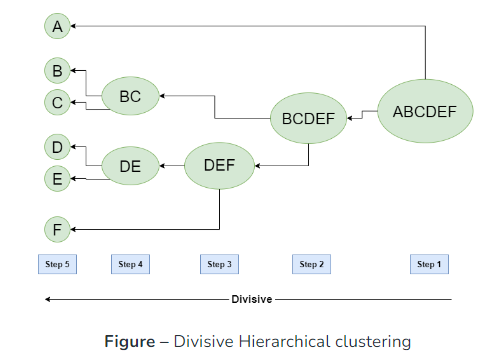


Here:

* *Ci*​ and *Cj*​ are the data points in clusters *i* and *j*, respectively.
* *D*(*u*,*v*) is the distance between data points *u* and *v*.

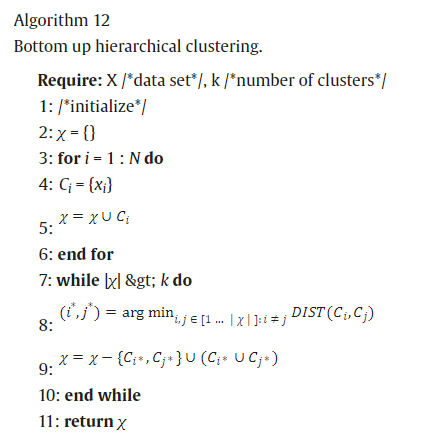
The algorithm proceeds by iteratively merging the closest clusters based on this linkage criterion until only one cluster remains.

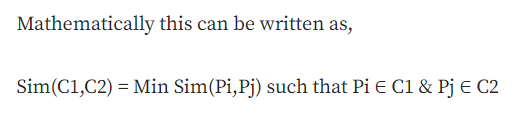
It's important to note that there are other linkage criteria, such as single linkage, average linkage, and Ward's method, each with its own formula for calculating the distance between clusters. The specific choice of linkage criterion depends on the characteristics of the data and the goals of the analysis.



<https://www.mdpi.com/2079-9292/12/10/2308?type=check_update&version=1>

<https://www.sciencedirect.com/topics/mathematics/hierarchical-clustering>





<https://towardsdatascience.com/understanding-the-concept-of-hierarchical-clustering-technique-c6e8243758ec>

<https://www.solver.com/xlminer/help/hierarchical-clustering-intro>