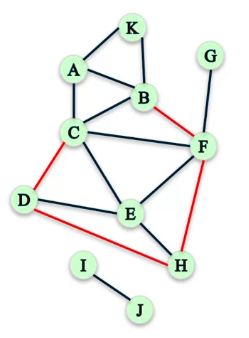
**Clustering Coefficient:**

**Triadic Closure:**

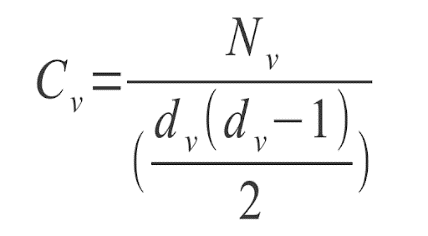
**Triadic closure –** This is the tendency for people who share connections in a social network to become connected. E.g. people with lots of mutual friends are more likely to become friends. Triadic closure is likely to occur for people in networks where their connection to someone new forms a new triangle. (red edges are the most likely connections to form due to triadic closure).

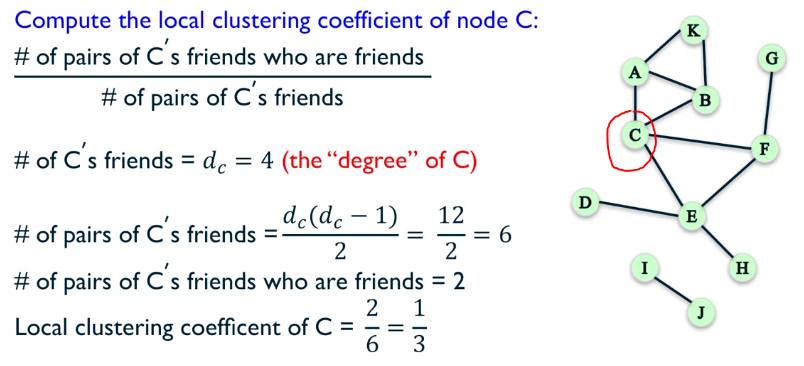
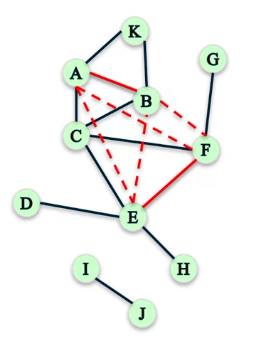


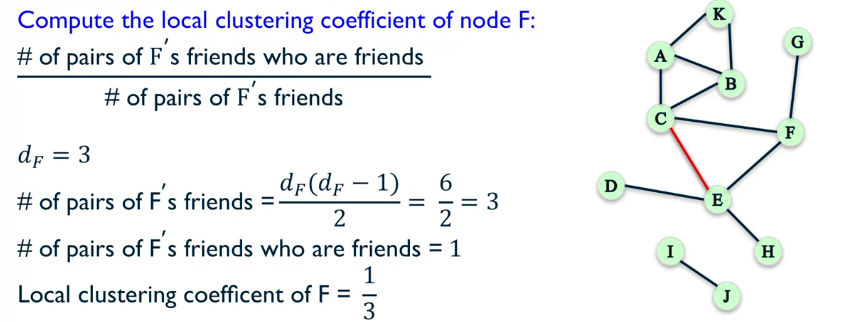
Sometimes we don’t have information one when the edges are formed, and we would like to know how present triadic closure is in this network. That is, how can we measure the prevalence of triadic closure in a network?

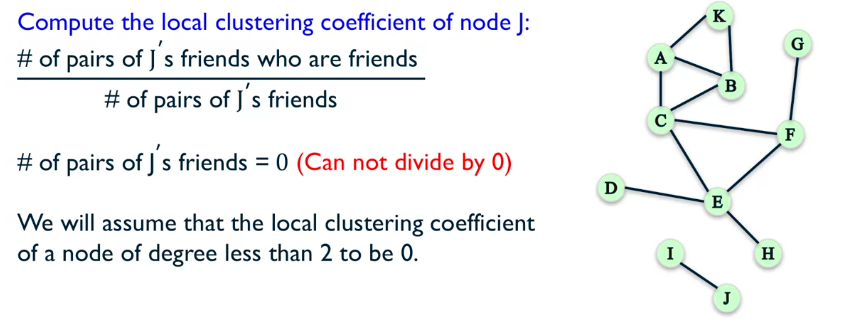
A useful way of measuring triadic closure is by using clustering.

**Local clustering coefficient of a node (Cv):** This is the fraction of the pair of the node’s friends divided by the fraction of the nodes common friend pairs. N is the number of common friend pairs; d is the degree of the node, so the number of its edges.

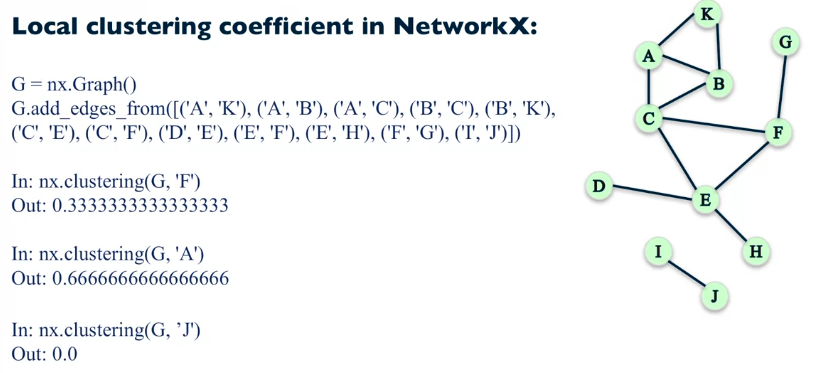


**Examples:**





**How to compute the local network coefficient using NetworkX?**

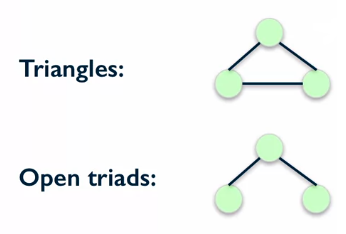


This is good for measuring local triadic closure, but how do we find how prevalent it is in the whole network? There are two different methods of doing this:

**Approach 1:** Average all the local clustering coefficients over all nodes in the graph.

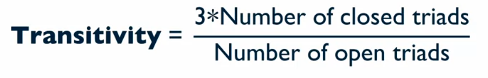


**Approach 2:** To measure the percentage of open triads (close to being a triangle) in the network.



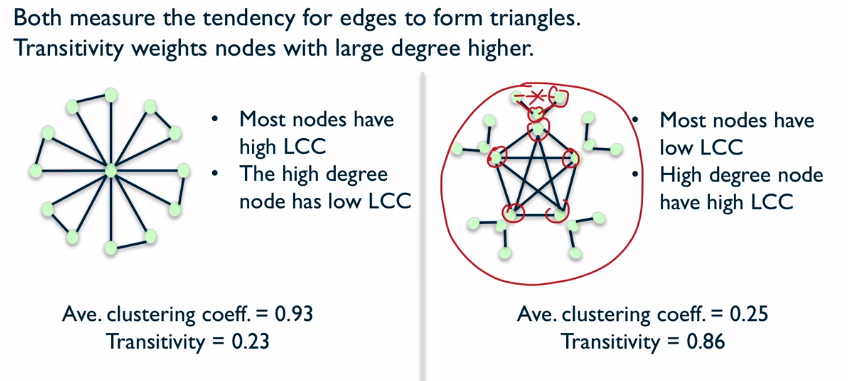
Inside each triangle there are 3 open triads, just think about taking away one edge for each.

For this approach we will calculate the **Transitivity** of the network. This is the 3 times the number of closed triads (triangles) divided by the number of open triads.





**Transitivity vs Average Clustering:**



For the network on the left transitivity is low because there is only one node with high degree and this one node has a low coefficient. Conversely, ACC is high because all of the outer nodes have a coefficient of 1, making the average large. The right network has 5 very well-connected nodes with high degrees, this makes the transitivity score high. The ACC score is low because there are 15 nodes that have very low coefficients and therefore the average is low.

