Equivalence & Hierarchical Clustering

Lecture delivered by: Prof. Niloy Ganguly
Scribed by: Himika Biswas (05CS6027)

01-03-2006

1 Equivalence

In a network, if two nodes are similar then we can say that these two nodes are **Equivalent**. This similarity is of two types:

1. **Regular Equivalence :** Two nodes are said to be *regularly equivalent* if they have the same profile of ties with other nodes who are also regularly equivalent. For example, in Figure1, *Child1* and *Child2* are regularly equivalent as *Mother1* and *Mother2* are also regularly equivalent. So we can say, *A_1 and A_2 are regularly equivalent if $A_1 -> B1$ and $A_2 -> B_2$ where B_1 and B_2 are regularly equivalent for all B_is.*

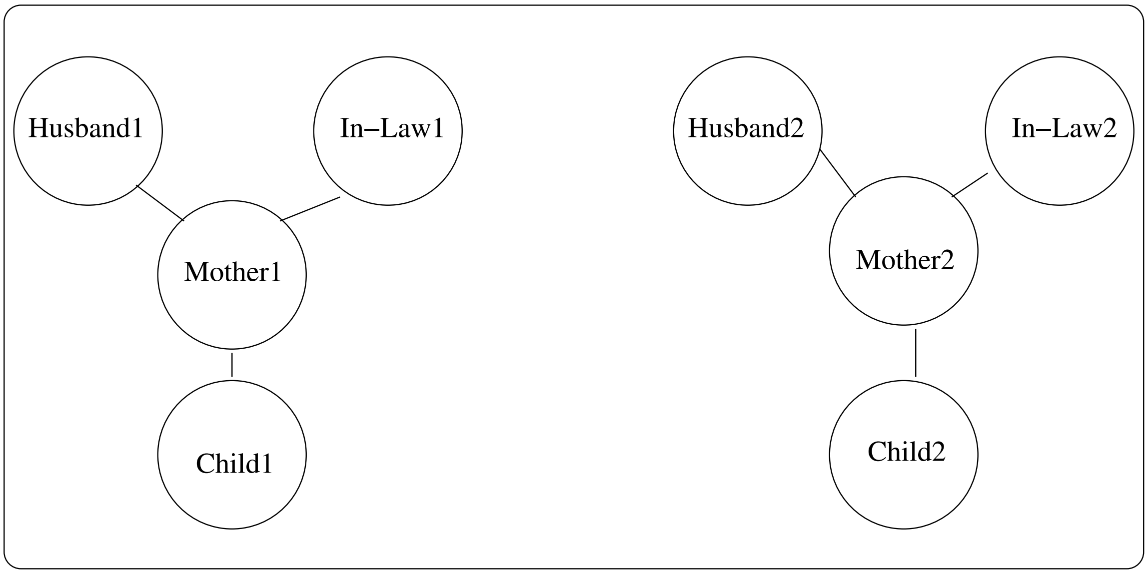

Figure 1: Regular Equivalence.

2. **Structural Equivalence :** *Two nodes are said to be structurally equivalent if they have the same relationships to all other nodes. For example, in Figure2,*

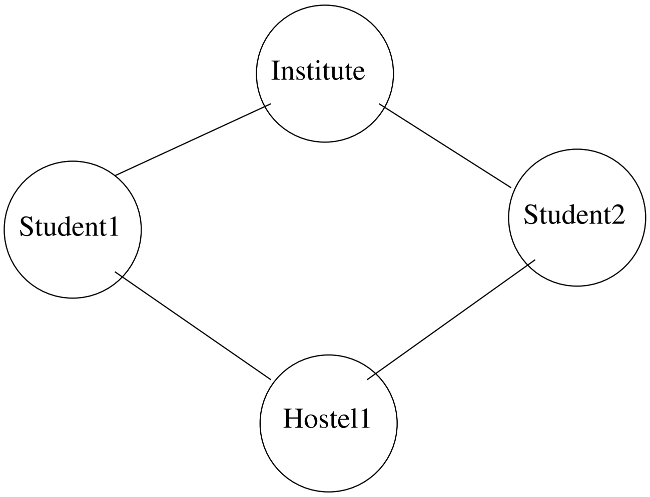

Figure 2: Structural Equivalence.

student1 and student2 are structurally equivalent as they have similar relationships with all other connected nodes. So we can say, A_1 *and* A_2 *are structurally equivalent if* $A_1 -> B_1$ *and* $A_2 -> B_2$ *where* $B_1 = B_2$ *for all* B_1 *and* B_2*. In real life (social life) exact structural equivalence is not possible. So we will essentially have a percentage of structural equivalence. Structural equivalence gives the idea of class, i.e., cluster analysis.*

	$City1$	$City2$	$City3$	$City4$	$City5$
$City1$	*0*	*1075*	*2013*	*2054*	*996*
$City2$	*1075*	*0*	*3273*	*2687*	*2037*
$City3$	*2013*	*3273*	*0*	*808*	*1307*
$City4$	*2054*	*2687*	*808*	*0*	*1059*
$City5$	*996*	*2037*	*1307*	*1059*	*0*

Table 1: Inter city distances matrix.

	$City1$	$City2$	$City3\|4$	$City5$
$City1$	*0*	*1075*	*2013*	*996*
$City2$	*1075*	*0*	*2687*	*2037*
$City3\|4$	*2013*	*2687*	*0*	*1059*
$City5$	*996*	*2037*	*1059*	*0*

Table 2: Inter city distances matrix after merging cities 3 & 4.

	$City1\|5$	$City2$	$City3\|4$
$City1\|5$	*0*	*1075*	*1059*
$City2$	*1075*	*0*	*2687*
$City3\|4$	*1059*	*2687*	*0*

Table 3: Inter city distances matrix after merging cities 1 & 5.

	$City1\|5\|3\|4$	$City2$
$City1\|5\|3\|4$	*0*	*1075*
$City2$	*1075*	*0*

Table 4: Inter city distances matrix after merging cities 1, 5 & 3, 4.

1.1 Hierarchical Clustering

Let there are 5 cities. Inter-city distances are given in Table1. If we perform Hierarchical Clustering on these cities we will search for pair of cities with minimum distance. We can see from the Table1 that, distance between cities 3 & 4 is minimum. So at the first level we shall merge these two cities. Now after merging these two cities we can determine the distance between other cities and this merged cities based on **maximum distance** *or based on* **minimum distance** *or based on* **average distance**. *Here we are considering the minimum distance and obtained result is shown in Table2. In Table3 & Table 4 we can see the results of subsequent merges. In Figure3 we can see the structure of hierarchical clustering*

2	○	3	○	4	○	5	○	1
○	○	○	○	○	○	○	○	○
○	○	x	x	x	○	○	○	○
○	○	x	x	x	○	x	x	x
○	○	x	x	x	x	x	x	x
x	x	x	x	x	x	x	x	x

Figure 3: Structure of Hierarchical Clustering.

for this example.