Equivalence

Nitesh Nandy

02CS1002 {niteshnandy@gmail.com}

25.1.2007

1 Structural Equivalence

Two nodes are structurally equivalence if they have same relationships to all other nodes.

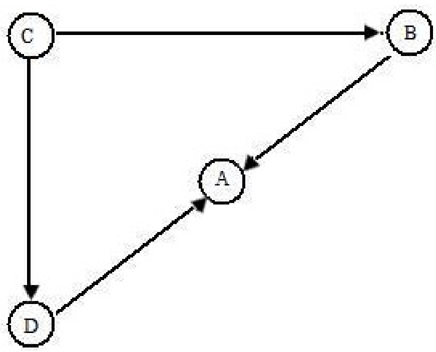

Figure 1: An example with four nodes

The adjacency matrix representation of the graph:

	A	B	C	D
A	0	0	0	0
B	1	0	0	0
C	0	1	0	1
D	1	0	0	0

1.1 Pearson's Correlation Coefficient

This topic has already been covered in previous lecture.

1.2 Eucledian Distance

Eucledian distance is the measure of dissimilarity between two nodes. It uses Geodesic Distance to calculate the shortest distance between two nodes.
**Geodesic Distance: Actual shortest distance on earth.*

G =

	A	B	C	D
A	0	∞	∞	∞
B	1	0	∞	∞
C	2	1	0	1
D	1	∞	∞	0

G^T =

	A	B	C	D
A	0	1	2	1
B	∞	0	1	∞
C	∞	∞	0	∞
D	∞	∞	1	0

now putting G and G^T together

$$D = \begin{pmatrix} G \\ G^T \end{pmatrix} = \begin{pmatrix} 0 & \infty & \infty & \infty \\ 1 & 0 & \infty & \infty \\ 2 & 1 & 0 & 1 \\ 1 & \infty & \infty & 0 \\ 0 & 1 & 2 & 1 \\ \infty & 0 & 1 & \infty \\ \infty & \infty & 0 & \infty \\ \infty & \infty & 1 & 0 \end{pmatrix}$$

Eucledian Distance = ED (A,B) = $\sqrt{\sum(A_i - B_i)^2}$

If Eucledian Distance between two nodes is more then it implies that those two nodes are different. Overall ED is less sensitive to differences.

To Normalize:

$$\frac{G.E.D - S.E.D}{G.E.D}$$

$$\frac{G.E.D - S.E.D}{G.E.D} < \frac{G.P.C - S.P.C}{G.P.C}$$

[*where G.E.D $\rightarrow$ Greatest Eucledian Distance, S.E.D $\rightarrow$ Smallest Eucledian Distance & G.P.C $\rightarrow$ Greatest Pearson's Correlation Coefficient*]

If the graph is sparse then we use Eucledian Distance Calculation.

1.3 % of Exact Match

Transposing the matrix and checking the number of columns that match gives the percent of exact match.

1.4 Jaccard's Coefficient

Let, S is No. of positive matches
P_1 is Non Null elements in A
P_2 is Non Null elements in B

Then JC = $\frac{S}{S+P_1+P_2}$

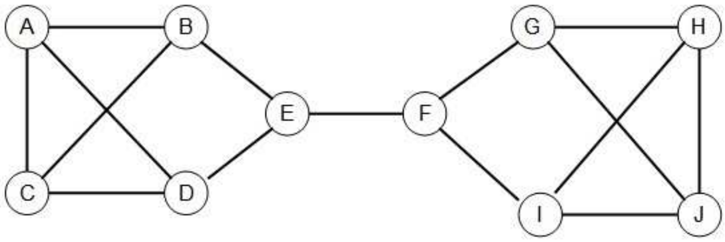

Figure 2: Calculate % Match and Jaccard's Coefficient

$$Match_{AB} = \begin{array}{|c|c|} \hline A & B \\ \hline 0 & 1 \\ 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ 1 & 1 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 1 \\ \hline \end{array}$$

% of Exact Match $= \frac{7}{10} = 0.7$

Jaccard's Coefficient $= \frac{3}{3+0+3} = 0.5$

2 Regular Equivalence

Also known as Automorphic Equivalence. Actually Automorphic equivalence is a stricter form of Regular Equivalence.

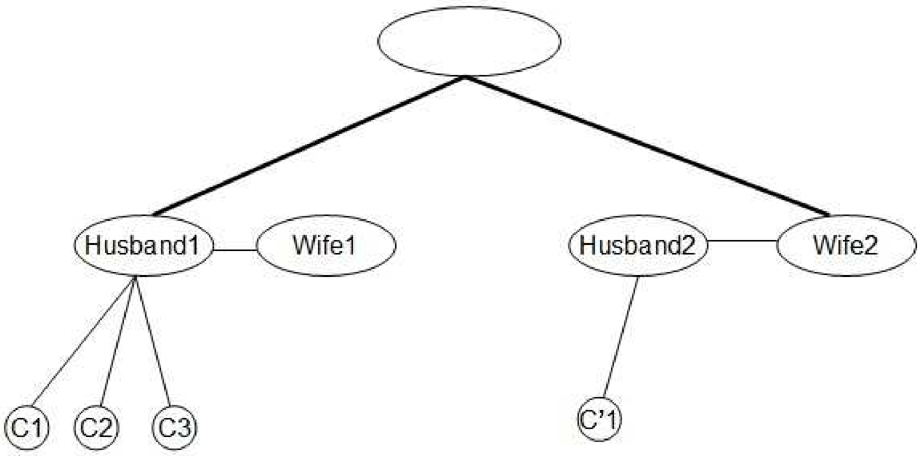

Figure 3: Difference between Regular Eqv. and Automorphic Eqv.

$R_1 : xRz \Rightarrow \exists w \in u(yRw \bigwedge w \approx x)$
$R_2 : zRx \Rightarrow \exists w \in u(wRy \bigwedge w \approx z)$

If swapping of node *Husband1* and node *Husband2* is allowed then it is a Regular Relation. There is no Automorphic Relation between *Husband1* and *Husband2* nodes.

2.1 Measuring similarity (Regular Equivalence)

We make profile of each node to measure their regular equivalence.

Referring *figure 2*:

Calculating for Node A and Node H

PROFILE(A) = {0,1,1,1,2,3,4,4,5,5}
PROFILE(H) = {5,5,4,4,3,2,1,1,0,1}

Now sort the profile columns and check if they are the same.

(After Sorting) PROFILE(A) = {0,1,1,1,2,3,4,4,5,5}

(After Sorting) PROFILE(H) = {0,1,1,1,2,3,4,4,5,5}

Therefore the nodes A and H have regular equivalence.