

Ucinet et Netdraw
Logiciels pour l'analyse des réseaux sociaux

Petit mode d'emploi

Version 1.01
Septembre 2009

Laurent BEAUGUITTE (UMR Géographie-cités, Université Paris 7)
beauguittelaurent@parisgeo.cnrs.fr

L'objectif de ce livret est simple : permettre à des personnes non spécialistes de l'analyse des réseaux sociaux d'utiliser les logiciels Ucinet et Netdraw. Pour compléter les éléments présentés ici, une courte bibliographie commentée permet de repérer les ouvrages essentiels.

Toutes les captures d'écran ont été réalisées sur la version 6.220 du logiciel et sous environnement Windows Vista.

Ce petit mode d'emploi ne prétend pas épuiser toutes les possibilités offertes par ces logiciels mais permettre leur rapide prise en main.

Vos commentaires et remarques sont les bienvenues et serviront à améliorer cet outil. Une version augmentée (et corrigée si besoin) est programmée pour mi-2010.

1. Présentation des logiciels

Ucinet et Netdraw sont tous deux développés par Borgatti, Everett et Freeman et sont téléchargeables à l'adresse suivante <http://www.analytictech.com/>

Si, jusqu'en décembre 2008, il était possible de télécharger les versions antérieures d'Ucinet gratuitement, ce n'est malheureusement plus le cas. Ucinet est utilisable dans sa version d'essai pendant 60 jours (d'utilisation effective)¹, Netdraw est quand à lui gratuit. Ces deux logiciels ne fonctionnent que sous environnement Windows.

Pourquoi alors se fatiguer à les maîtriser dans la mesure où il existe des logiciels gratuits et plus puissants (Pajek, Stocnet, Tulip par exemple) ? La raison est simple : ces deux logiciels sont d'une grande simplicité d'emploi et ils permettent de très rapidement produire des analyses efficaces et claires.

L'un des points les plus perturbants pour un francophone désireux de se frotter aux logiciels de 'Social network analysis' est l'hétérogénéité du vocabulaire d'un logiciel à l'autre (ainsi 'node' dans Ucinet devient 'vertice' dans Pajek). La situation se complique encore quand le dit francophone se base sur les ouvrages de référence français. Le parti pris choisi ici est de systématiquement garder les termes présents dans les deux logiciels. Les termes les plus courants sont traduits (degree devient degré), les termes plus ésotériques sont écrits en anglais, en gras, et placés entre apostrophes. Un lexique placé en annexe permet de devenir à peu près bilingue en analyse des réseaux.

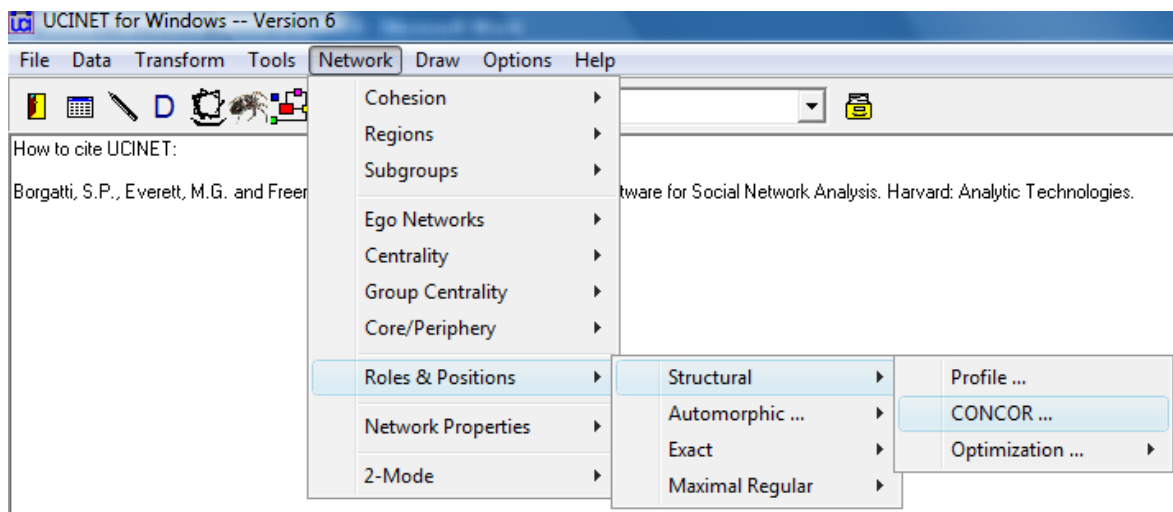
La logique des deux logiciels est fondamentalement Windows : quelques icônes, des menus déroulants et des sous-menus, plus quelques raccourcis. Les choix à effectuer pour lancer un traitement sont indiqués sous cette forme

Intitulé du menu > intitulé du sous-menu > options à choisir

Les captures d'écran présentes dans la partie 2 permettent de repérer les principales options disponibles. Ainsi, si vous désirez lancer l'algorithme CONCOR, vous devrez dérouler les menus et sous-menus suivants.

¹ Les tarifs en septembre 2009 pour une licence sont de 40 dollars pour un étudiant, 150 dollars pour une université (tarifs dégressifs pour plusieurs licences).

Figure 1 : Lancer l'algorithme CONCOR



Cela sera noté

Network > Roles & Positions > Structural > CONCOR

Lorsque la première étape consiste à cliquer sur une icône, puis à choisir parmi les menus déroulants, la notation utilisée sera la même. Ainsi, pour contrôler une matrice, il convient de cliquer sur l'icône 'Spreadsheet' (le nom des icônes apparaît lorsque la souris les survole).



Puis, dans la fenêtre 'Spreadsheet' de dérouler le menu 'File' et de choisir 'Open' avant d'indiquer l'emplacement du fichier à ouvrir.

La notation pour une telle opération est

Spreadsheet > File > Open

Pour installer les logiciels, il suffit de les télécharger et de lancer les fichiers .exe.

2. Les interfaces

Cette partie est indigeste dans la mesure où elle présente la plupart des icônes et des menus déroulants des 2 logiciels. Il est probablement inutile de la lire entièrement mais les captures d'écran peuvent être utiles par la suite, notamment lorsqu'on ne sait plus où se trouve telle ou telle option.

Figure 2 : Page d'accueil d'Ucinet

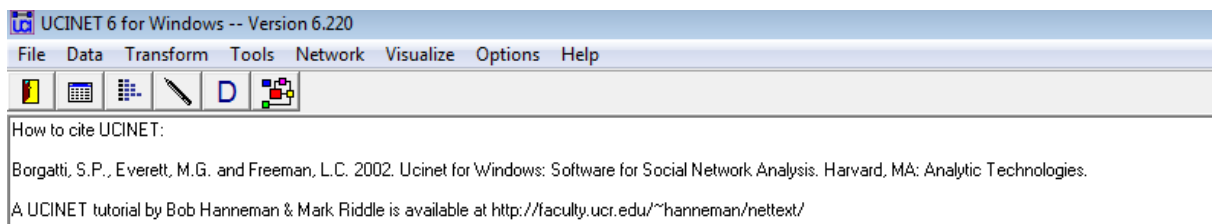
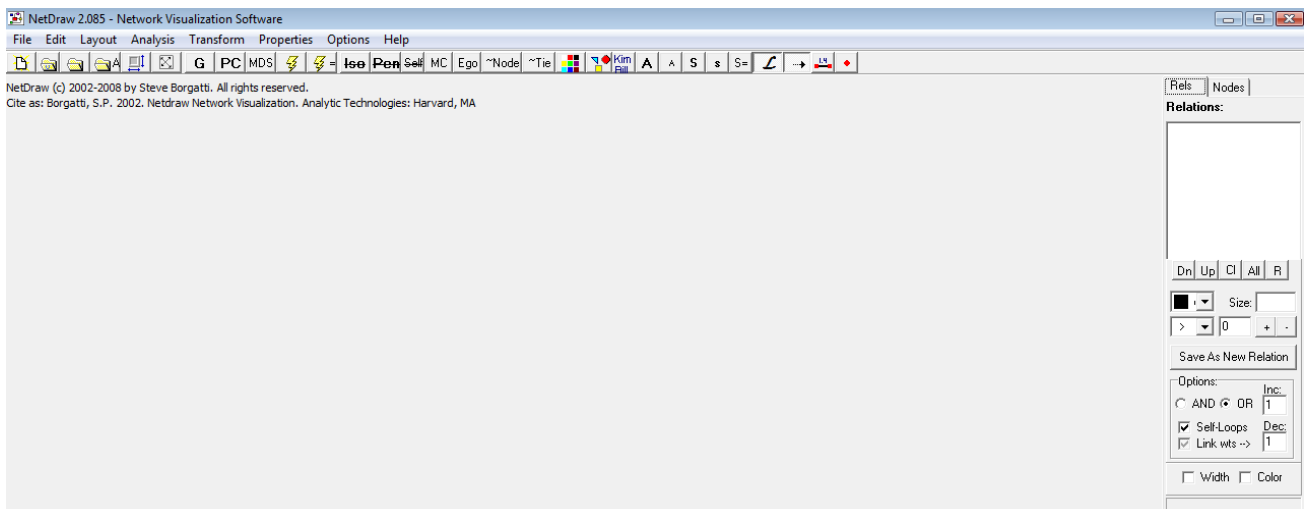


Figure 3 : Page d'accueil de Netdraw



Pourquoi 2 logiciels ? Ucinet permet de réaliser un certain nombre de traitements sur des matrices, Netdraw permet de les représenter graphiquement. Certaines mesures peuvent être effectuées directement dans Netdraw.

Les icônes d'Ucinet

5 icônes sont présentes dans la barre des tâches d'Ucinet. Leur nom apparaît lorsqu'on les survole avec la souris (il est indiqué entre parenthèses dans les lignes ci-dessous).



De gauche à droite :

la porte (Exit) permet de quitter le logiciel.

la feuille (Matrix spreadsheet editor) permet d'ouvrir une matrice vierge

les barres (Excel DL editor) permettent d'importer des fichiers

le stylo (Edit Text File) permet d'ouvrir un fichier .txt

le D (Display Ucinet Dataset) permet d'aller rechercher un fichier Ucinet

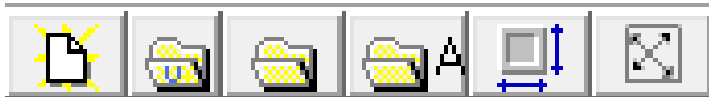
le graphe coloré (Visualize network with Netdraw) permet de lancer Netdraw

L'icône 'Matrix spreadsheet editor' permet d'importer des matrices saisies dans un tableur avec un simple Ctrl+C (dans le tableur) – Ctrl+V (dans Ucinet).

Comme souvent, les actions réalisables en cliquant sur une icône peuvent aussi être sélectionnées via les menus déroulants.

Les icônes de Netdraw sont beaucoup plus nombreuses (30 !). Certaines permettent de représenter le graphe en fonction de diverses propriétés, elles ne seront donc décrites que dans la partie 6. Là encore, le nom apparaît quand on survole l'icône

Les cinq premières de gauche à droite



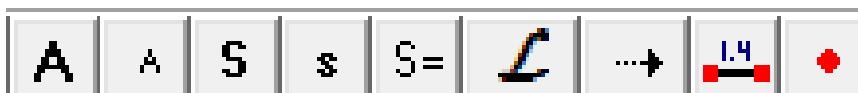
la feuille permet de nettoyer l'écran

les trois dossiers permettent d'ouvrir des fichiers de format variés (Ucinet bien sûr, mais aussi .txt, Pajek etc.).

l'écran aux flèches bleues permet d'élargir le graphe afin que celui-ci occupe tout l'écran.

l'écran aux flèches croisées ouvre une boîte d'outil pour faire pivoter le graphe dans la direction voulue, de zoomer sur une partie etc.

Les neuf dernières de gauche à droite



Les deux A permettent d'agrandir ou de réduire la taille des nœuds, les deux S permettent d'agrandir ou de réduire la taille des noms des nœuds. Le L italique et la flèche permettent de représenter ou non les noms des nœuds et les directions des relations.

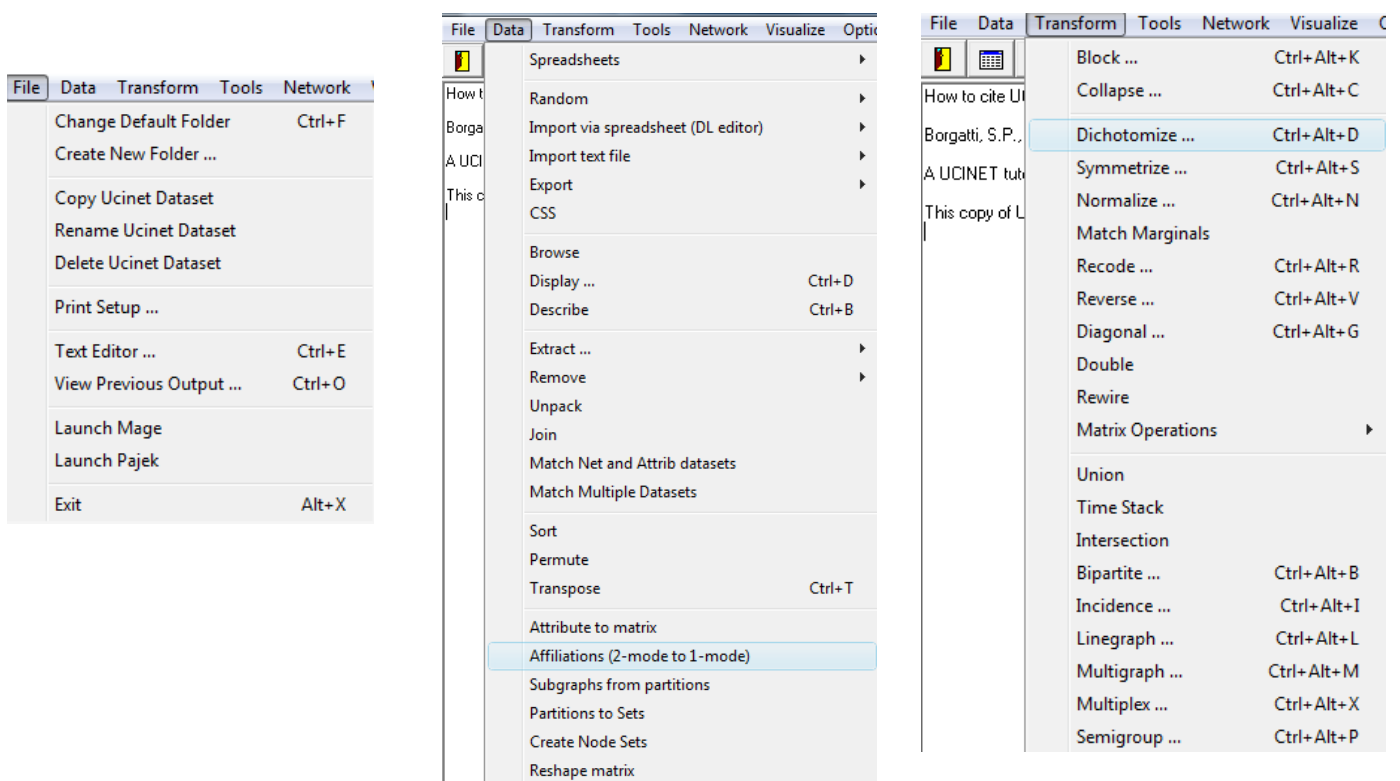
La barre numérotée 1-4 permet de représenter ou non l'intensité des liens en cas de 'valued graph'. Enfin le point rouge permet de dessiner de nouveaux nœuds.

La barre de tâches située à gauche sera abordée elle aussi dans la partie 6 consacrée aux représentations graphiques.

Les menus déroulants

Ucinet propose un choix très riche et il n'est pas question ici de présenter systématiquement toutes les options disponibles et leurs conséquences. Les captures d'écran ci-dessous montrent ce qui se cache derrière les termes de la barre supérieure. Toutes les commandes ou presque ouvrent une boîte de commande permettant de modifier les paramètres de l'action à réaliser.

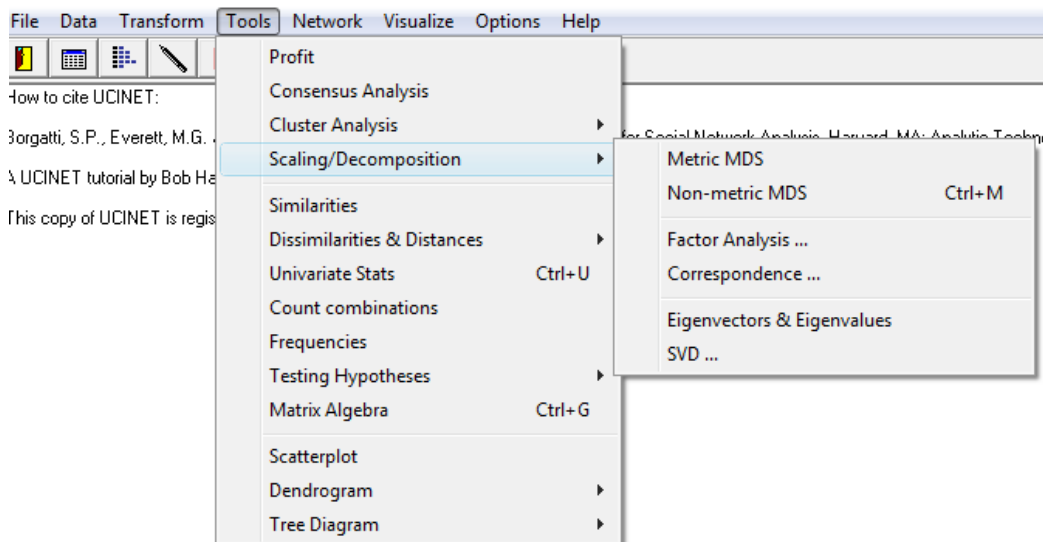
Figure 4 : Menu File, Data et Transform



Le menu 'Transform' est très utile pour modifier des matrices et propose de nombreuses options. Il est par exemple de les symétriser en utilisant le minimum, le maximum, la somme, la moyenne etc.

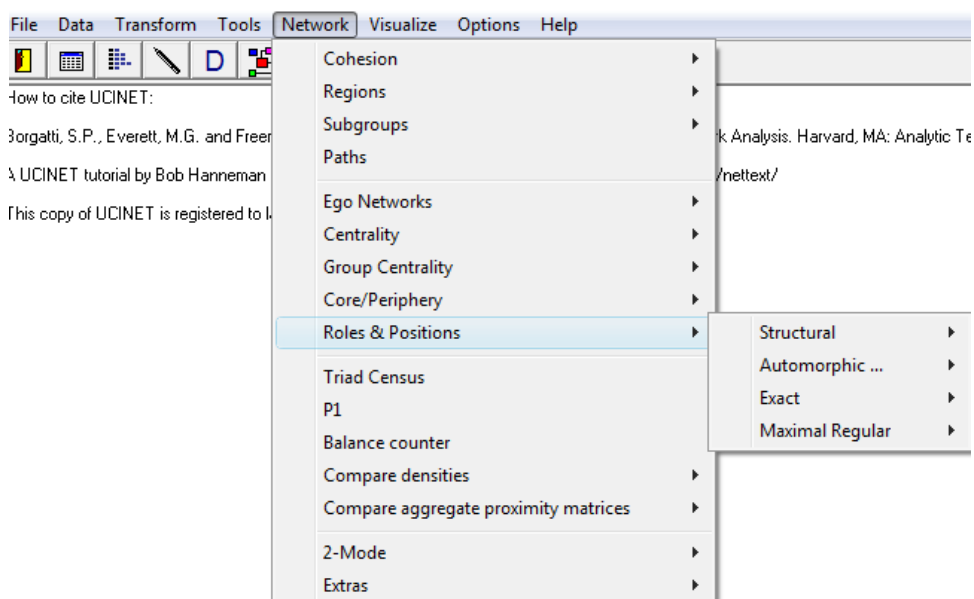
Le menu 'Tools' permet de réaliser des analyses statistiques à partir des matrices étudiées.

Figure 5 : Menu Tools



Le menu 'Network' permet de mener à bien très des analyses concernant le réseau étudié.

Figure 6 : Menu Network



Le menu 'Visualise' permet de lancer au choix Netdraw, Mage (effets 3D) ou Pajek.

Le menu 'Options' permet divers réglages (nombre de décimales, largeur des colonnes etc.).

Enfin le menu 'Help' est très utile, surtout la partie 'Help Topic' d'une grande qualité (raccourci clavier : F1). On trouvera pour toutes les commandes une description détaillée du principe, le détail des algorithmes utilisés, les conditions d'application ainsi que des références bibliographiques.

Les menus déroulants de Netdraw seront évoqués dans la partie 6.

3. Questions de vocabulaire

Pour utiliser Ucinet ou Netdraw, il est nécessaire d'avoir une matrice (k lignes * l colonnes). Lorsque $k = l$, la matrice est dite carrée. Si les intitulés des colonnes et des lignes sont les mêmes, il s'agit d'un '**one-mode network**'. Lorsque $k \neq l$, la matrice est dite rectangulaire ('**2-mode network**'). Il est possible de saisir une matrice avec Ucinet mais il est beaucoup plus pratique de la créer avec son tableur habituel, de la copier et de la coller ensuite dans Ucinet. Trois matrices carrées ('**one mode network**') serviront d'exemple tout au long de ce livret². Les deux premières sont binaires ('**binary**'), c'est-à-dire qu'elle signale si un lien entre deux nœuds est présent (1) ou absent (0) ; l'une est symétrique ('**non directed**'), l'autre ne l'est pas ('**directed**'). La dernière est évaluée ('**valued**'), ce qui signifie que, non seulement l'existence de la relation importe, mais l'intensité de la relation compte également.

Par convention, dans le cas de '**directed graph**', l'origine du lien est en ligne et le destinataire du lien est en colonnes.

La nature des acteurs et des liens n'a pas d'importance. Ayant une formation de géographe, je les avais créées pour étudier des flux commerciaux entre pays, mais ce pourrait très bien être des rapports de parenté entre personnes ou des liens financiers entre des entreprises...

Les points du réseau sont les '**nodes**', les liens existant les '**ties**'. Par la suite, les termes nœuds et acteurs sont équivalents à '**nodes**', les termes liens ou relations équivalents à '**ties**'.

Les trois matrices sont présentés ci-dessous de deux façons différentes et équivalentes ; sous leur forme originale dans Excel, sous leur forme graphique dans Netdraw.

² Les '**signed graphs**', les '**two-mode networks**' et autres '**multigraphs**' seront abordés dans la version 1.1 de ce livret (version prévue pour l'automne 2010)

Figure 7 : Directed binary matrix and graph (GR1)

	A	B	C	D	E	F	G	H	I	J	K	L
A	0	1	0	0	1	1	0	0	0	0	0	0
B	1	0	0	0	0	0	1	1	0	0	0	0
C	0	0	0	0	0	0	0	1	0	0	0	0
D	0	0	0	0	0	0	0	1	0	0	0	0
E	0	0	0	0	0	0	0	0	1	0	0	0
F	0	0	0	0	0	0	0	0	0	1	0	0
G	0	0	0	0	0	0	0	0	0	0	1	0
H	0	0	0	0	0	0	0	0	0	0	0	1
I	0	0	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	0	0	0
K	0	0	0	0	0	0	0	0	0	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	0

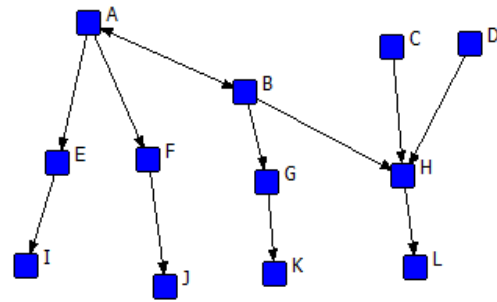


Figure 8 : Undirected binary matrix and graph (GR2)

	A	B	C	D	E	F	G	H	I	J	K	L
A	0	1	1	0	0	0	0	0	0	0	0	0
B	1	0	1	1	0	0	0	0	0	0	0	1
C	1	1	0	1	0	0	0	0	0	0	0	0
D	0	1	1	0	1	0	0	0	0	0	0	0
E	0	0	0	1	0	1	0	0	0	0	0	0
F	0	0	0	0	1	0	1	1	1	0	0	0
G	0	0	0	0	0	1	0	1	0	0	0	0
H	0	0	0	0	0	1	1	0	0	0	1	0
I	0	0	0	0	0	1	0	0	0	1	0	0
J	0	0	0	0	0	0	0	0	1	0	1	0
K	0	0	0	0	0	0	0	1	0	1	0	0
L	0	1	0	0	0	0	0	0	0	0	0	0

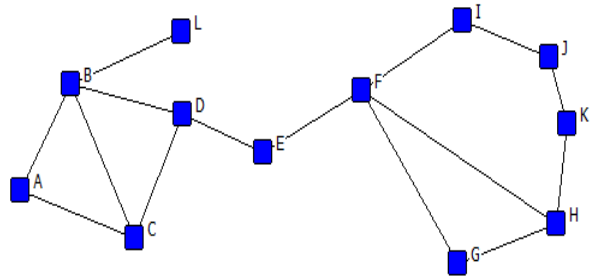
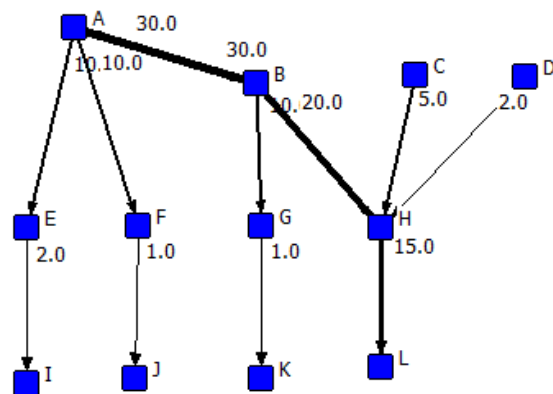


Figure 9: Directed valued matrix and graph (GR3)

	A	B	C	D	E	F	G	H	I	J	K	L
A	0	30	0	0	10	10	0	0	0	0	0	0
B	30	0	0	0	0	0	10	20	0	0	0	0
C	0	0	0	0	0	0	0	5	0	0	0	0
D	0	0	0	0	0	0	0	2	0	0	0	0
E	0	0	0	0	0	0	0	0	2	0	0	0
F	0	0	0	0	0	0	0	0	0	1	0	0
G	0	0	0	0	0	0	0	0	0	0	1	0
H	0	0	0	0	0	0	0	0	0	0	0	15
I	0	0	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	0	0	0
K	0	0	0	0	0	0	0	0	0	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	0



Deux remarques concernant ces graphes; ils sont dits ‘**simple**’ (à l’inverse des ‘**multigraphs**’) pour deux raisons. Premièrement, les liens d’un nœud à lui-même (appelé ‘**loop**’ ou ‘**self-tie**’) n’existent pas, ce qui se vérifie par le fait que la diagonale soit vide dans les trois matrices. Deuxièmement, une relation et une seule existe entre les acteurs.

Autre précision, tous les nœuds sont reliés par au moins un lien ; le graphe est dit **‘connected’** car il n’y a pas de nœud isolé (**‘isolate’**). Dans les graphes GR1 et GR3, il y a un lien de A et B et un autre de B vers A ; on parle alors de **‘reciprocal tie’**.

Dernier point de vocabulaire, la distance entre deux nœuds se mesure en nombre de liens. La distance la plus utilisée est nommée géodésique (**‘geodesic distance’**), elle désigne la distance la plus courte existant entre deux nœuds. Dans un graphe connecté (**‘connected’**) et symétrique, il existe une distance entre tous les nœuds.

Par exemple, dans le graphe GR2, les distances géodésiques entre nœuds sont les suivantes
AB : 1, AL : 2, AJ : 7 etc.

Par contre, si le graphe est orienté (**‘directed graph’**), seules les distances possibles seront mesurées. Dans le graphe GR1 par exemple, il n’est pas possible de se rendre de L à H.

4. Premiers pas, premières mesures

La première étape consiste à importer les données. La solution la plus simple consiste à copier dans son tableur la matrice et de la coller directement dans Ucinet en ouvrant une feuille (**‘Data> Spreadsheets >Matrix’** ou **‘Ctrl+S’** ou un clic sur l’icône **‘Spreadsheets’**). Les options présentées sont surtout utiles lorsque l’on saisit directement des données dans Ucinet. Une fois la matrice collée, enregistrer-là où vous le souhaitez (**‘Save’** ou **‘Save as’** ou **‘Ctrl+S’**).

Un certain nombre de mesures sont utiles pour caractériser le graphe dans son ensemble ;

- le nombre de nœuds
- le nombre de liens
- la distance géodésique maximale entre deux nœuds du graphe
- la distance géodésique moyenne entre deux nœuds
- la densité du graphe (nombre de liens existants/nombre de liens réels).

Prenons l’exemple du graphe GR1.

Avec Ucinet, plusieurs options sont possibles

Densité et nombre de liens peuvent être obtenus avec la commande

Network > Cohesion > Density > Density overall

Le résultat est un fichier texte.

DENSITY / AVERAGE MATRIX VALUE

```
-----  
Input dataset:      GR1 (C:\Users\laurent\Desktop\LOGICIELS\SOCIAL NETWORKS\UCINET MODE  
D'EMPLOI\DATASETS\GR1)  
Output dataset:    GR1-density (C:\Users\laurent\Desktop\GR1-density)  
      Density  No. of Ties  
-----  
GR1      0.0909   12.0000
```

Systématiquement, les résultats indiquent : l'intitulé de l'opération effectuée, le chemin du fichier source et celui du fichier sortie. Si des opérations ont été effectuées avant le calcul (ainsi certaines procédures nécessitent des matrices symétriques et binaires), elles sont indiquées avant les résultats.

La densité de liens est donc de 9%, c'est-à-dire que 9% des liens possibles sont présents. Pour vérifier le calcul ; le graphe étant orienté, il y a 12×11 liens possibles soit 132. 12 liens existent donc la densité est de $12/132$ soit 0.09 ou 9%.

On obtient la densité et l'écart-type (mais pas le nombre de liens) avec la commande

Network > Cohesion > Density > Old Density Procedure

Les distances entre nœuds sont calculées à l'aide de la commande

Network > Cohesion > Distance

Les résultats obtenus avec le graphe GR1 sont les suivants

GEODESIC DISTANCE

```

Type of data:          ADJACENCY
Nearness transform:   NONE
Input dataset:        GR1 (C:\Users\laurent\Desktop\LOGICIELS\SOCIAL NET
Output distance:      C:\Users\laurent\Desktop\LOGICIELS\SOCIAL NETWORKS
  
```

For each pair of nodes, the algorithm finds the # of edges in the shortest path between them.

```

Average distance (among reachable pairs) = 1.692
Distance-based cohesion ("Compactness") = 0.139
(range 0 to 1; larger values indicate greater cohesiveness)
Distance-weighted fragmentation ("Breadth") = 0.861
  
```

Frequencies of Geodesic Distances

	Freque	Propor
1	12.000	0.462
2	10.000	0.385
3	4.000	0.154

Geodesic Distances

		1	2	3	4	5	6	7	8	9	0	1	1	1
	A B C D E F G H I J K L													
1	A	0	1			1	1	2	2	2	2	3	3	
2	B	1	0			2	2	1	1	3	3	2	2	
3	C			0					1				2	
4	D				0			1					2	
5	E					0			1					
6	F						0			1				
7	G							0				1		
8	H								0				1	
9	I									0				
10	J										0			
11	K											0		
12	L												0	

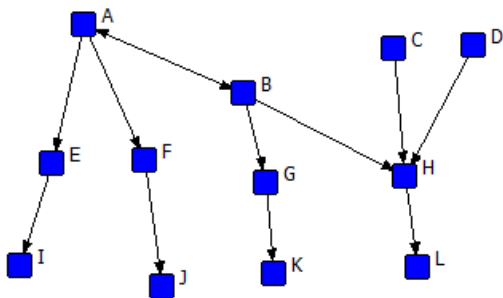
Qu'est-ce que ça veut dire ? La distance moyenne entre nœuds est de 1.692. La fréquence des distances (effectif brut et pourcentage) apparaît ensuite. Enfin, une matrice donne les distances entre chaque nœud (il faut par exemple 3 liens pour relier A à L : AB BH HL).

Dans le premier tableau, la ligne '3 4.000 0.154' signifie qu'il y a une distance de 3 liens présente 4 fois dans le graphe, ce qui représente 15% du total de nombre de distances entre nœuds. La dernière ligne du tableau correspond à la distance la plus courte possible entre les nœuds les plus éloignés les uns des autres du graphe. Cette mesure est aussi appelé diamètre ('**diameter**'). Toutes les cases laissées en blanc dans la matrice des '**geodesic distances**' signalent des distances impossibles (A ne peut pas joindre C et D).

Il est possible de s'intéresser aux propriétés des nœuds en tant que tels.

L'un des indicateurs les plus communs est celui de '**degree**', il correspond au nombre de liens adjacents à un nœud. Dans le cas des '**directed graphs**', on distingue les '**in-degree**' (nombre de liens qui atteignent le nœud) et les '**out-degree**' (nombre de liens qui partent du nœud).

Examinons le graphe GR1



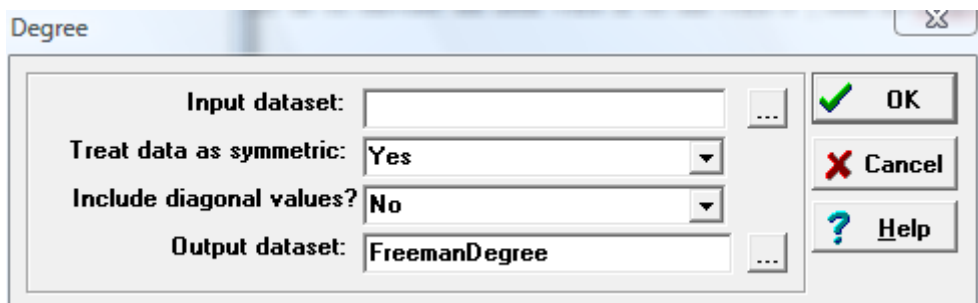
A reçoit un lien issu de B et envoie un lien vers E et un vers F ; son **'in-degree'** est de 1, son **'out-degree'** est de 2.

Il est possible de les calculer sous Ucinet comme sous Netdraw.

Sous Ucinet,

Network > Centrality > Degree

Puis régler les paramètres (matrice symétrique ou non, diagonale valide ou non)



Sous Netdraw

Analysis > Centrality measures

Il faut préciser si le graphe est **'directed'** ou non.

Les résultats du graphe GR1 avec Ucinet sont les suivants (symmetric NO, diagonal values NO)

FREEMAN'S DEGREE CENTRALITY MEASURES

Diagonal valid? NO
 Model: ASYMMETRIC
 Input dataset: GR1 (C:\Users\laurent\Desktop\GR1)

		1	2	3	4
		outDegree	InDegree	NrmOutDeg	NrmInDeg
1	A	3.000	1.000	27.273	9.091
2	B	3.000	1.000	27.273	9.091
3	C	1.000	0.000	9.091	0.000
4	D	1.000	0.000	9.091	0.000
5	E	1.000	1.000	9.091	9.091
6	F	1.000	1.000	9.091	9.091
7	G	1.000	1.000	9.091	9.091
8	H	1.000	3.000	9.091	27.273
9	I	0.000	1.000	0.000	9.091
10	J	0.000	1.000	0.000	9.091
11	K	0.000	1.000	0.000	9.091
12	L	0.000	1.000	0.000	9.091

DESCRIPTIVE STATISTICS

		1	2	3	4
		outDegree	InDegree	NrmOutDeg	NrmInDeg
1	Mean	1.000	1.000	9.091	9.091
2	Std Dev	1.000	0.707	9.091	6.428
3	Sum	12.000	12.000	109.091	109.091
4	Variance	1.000	0.500	82.645	41.322
5	SSQ	24.000	18.000	1983.471	1487.603
6	MCSSQ	12.000	6.000	991.736	495.868
7	Euc Norm	4.899	4.243	44.536	38.569
8	Minimum	0.000	0.000	0.000	0.000
9	Maximum	3.000	3.000	27.273	27.273

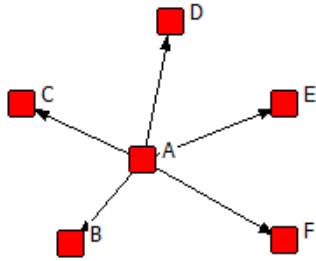
Network Centralization (Outdegree) = 19.835%
 Network Centralization (Indegree) = 19.835%

On retrouve les 'out' et 'in degree' bruts (premier tableau, colonne 1 et 2) et normalisés (tableau 1, colonne 3 et 4) pour chaque nœud. Ainsi A envoie 3 liens et pourrait en envoyer au maximum 11 (pas de 'self-tie'), il envoie donc 3/11 soit 27% des liens possibles.

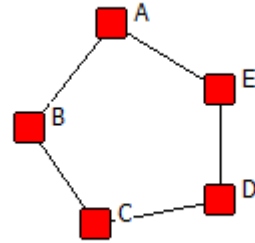
Le deuxième tableau présente des statistiques sur les deux 'degree'.

L'indicateur de centralisation présent aux deux dernières lignes compare la situation étudiée avec deux situations théoriques ; celle d'un graphe circulaire (l'égalité entre nœuds est parfaite) et celle du graphe en étoile (un nœud est lié à tous les autres et aucun autre lien n'existe, l'inégalité est maximale). L'indice de centralisation calculé par Ucinet apparaît en dessous de chaque graphe.

Figure 10: Star network and circle graph



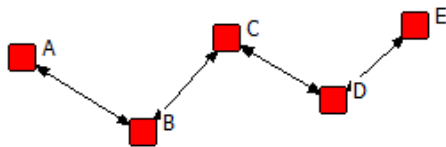
Network centralization (out-degree) 100%



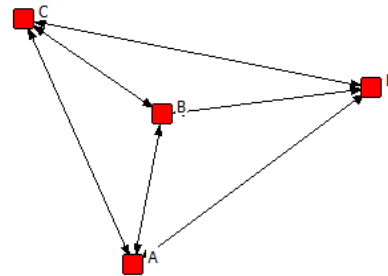
0%

Il est utile de connaître deux autres cas de figure théorique; la ligne et le graphe complet. Même si on ne les rencontre jamais, ils fournissent des valeurs de référence pour comparer entre eux des graphes de différentes tailles, densités etc.

Figure 11: Line network and complete network



Network centralization 16.67%



0%

Des mesures plus sophistiquées sont disponibles pour caractériser la situation relationnelle d'un nœud dans un graphe. Pour se limiter aux plus employés, citons '**closeness**', '**betweenness**' et '**eigenvector**'. Ces mesures sont à manier avec précaution car elles ont été créées pour étudier des faits sociaux et ne peuvent être appliquées à tous les flux (économiques notamment).

Attention : Ucinet rend la matrice symétrique et binaire avant de calculer les valeurs

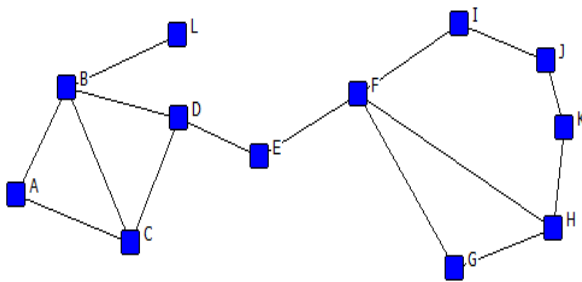
'**Closeness**' mesure la proximité avec les autres nœuds du réseau. Plus un nœud est proche de tous les autres, plus son score est élevé.

'Betweenness' permet de saisir la position d'intermédiaire de certains nœuds. Cette position d'intermédiaire peut en effet être cruciale dans certains réseaux (possibilité de faire circuler une information, de transmettre une maladie etc.)

'Eigenvector' recherche les acteurs les plus centraux en étudiant à la fois la position du nœud et la structure de l'ensemble du réseau via une analyse factorielle.

Calculons ces indicateurs sur le graphe GR2

Network > Centrality > Multiple mesures



Dans le cas d'un graphe de cette taille, il est prudent d'observer sa structure avant de jouer à la DS pour sortir x résultats plus ou moins contradictoires impossible à interpréter (je sais de quoi je parle !). Au moins un nœud apparaît clairement périphérique (L). E a très peu de liens mais il assure la jonction entre les groupes ABCD et FGHIJK.

Les résultats sont les suivants :

Important note:
This routine automatically symmetrizes and binarizes.

Normalized Centrality Measures

	1	2	3	4
	Degree	Closeness	Betweenness	Eigenvector
1 A	18.182	27.500	0.000	49.219
2 B	36.364	35.484	25.455	72.182
3 C	27.273	34.375	7.273	65.376
4 D	27.273	42.308	50.909	61.315
5 E	18.182	45.833	54.545	33.808
6 F	36.364	45.833	59.091	33.172
7 G	18.182	34.375	0.000	20.219
8 H	27.273	36.667	15.455	23.336
9 I	18.182	35.484	13.636	15.348
10 J	18.182	28.947	1.818	9.724
11 K	18.182	29.730	2.727	11.829
12 L	9.091	26.829	0.000	25.827

DESCRIPTIVE STATISTICS FOR EACH MEASURE

	1	2	3	4
	Degree	Closeness	Betweenness	Eigenvector
1 Mean	22.727	35.280	19.242	35.113
2 Std Dev	7.873	6.316	21.893	20.827
3 Sum	272.727	423.365	230.909	421.353
4 Variance	61.983	39.886	479.316	433.759
5 SSQ	6942.149	15415.132	10195.041	20000.000
6 MCSSQ	743.802	478.629	5751.791	5205.109
7 Euc Norm	83.320	124.158	100.970	141.421
8 Minimum	9.091	26.829	0.000	9.724
9 Maximum	36.364	45.833	59.091	72.182

Les mesures sont normalisées, ce qui n'aide pas à comprendre et interpréter les résultats.

La colonne '**degree**' a déjà été vue ; pourcentage de liens réels par rapport au total des liens possibles par nœud. Plus la '**closeness**' est élevée, plus le nœud est proche de tous les autres (comparez le score de L avec celui de E et F). La '**betweenness**' est très dispersée avec trois nœuds ayant une valeur de zéro. Pourquoi ? L n'est sur le passage d'aucun chemin dans le graphe (c'est un cul de sac) ; A et G sont connectés davantage mais nullement indispensables, un lien direct existe entre BC et entre GH. A l'inverse, les nœuds E et F doivent être franchis pour joindre de nombreuses paires de nœuds.

Enfin le '**eigenvector**' donne l'avantage aux nœuds fortement connectés (B et D) de la partie gauche du graphe qui est plus dense que la partie droite.

Aucune valeur n'est meilleure qu'une autre, elles mesurent des phénomènes différents en utilisant des méthodes différentes.

D'autres indices existent et j'avoue une faiblesse pour un indice très malin mis au point par Bonacich. Ce dernier a eu une intuition brillante ; avoir de l'influence suppose d'avoir de nombreux liens certes, mais l'influence sera d'autant plus forte que les contacts seront eux isolés. Si je donne une information à quelqu'un qui connaît énormément de monde, je suis une voix parmi d'autres. Si je donne la même information à quatre ermites, ma voix compte davantage.

L'indice est calculée via

Network > Centrality > Alpha centrality

Le résultat est intéressant pour les graphes de grande taille.

5. A la recherche des structures

L'une des grandes forces d'Ucinet est le nombre de méthodes disponibles pour créer des partitions de graphes – ce qu'un géographe soucieux de régionalisation ne peut qu'apprécier.

La première étape consiste à déterminer la méthode appropriée. Dans le cas de '**non directed graph**', les principales méthodes disponibles sont les cliques (et tous les dérivés possibles). Les différentes formes d'**'equivalence'** sont indiquées pour les '**directed graphs**'.

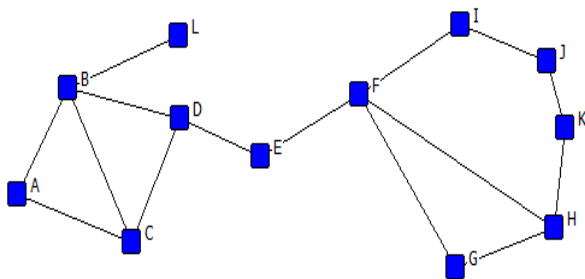
5.1 Non directed graph

La définition d'une clique est simple ; un sous graphe complet d'au moins trois nœuds. Dit autrement, un sous ensemble de trois nœuds minimum à l'intérieur duquel tous les liens possibles sont présents.

La commande pour les obtenir est la suivante

Network > Subgroups > Cliques

Le graphe GR2 sera utilisé comme exemple dans toute cette partie.



Les résultats de la commande ci-dessus se présentent ainsi;

3 cliques found.

- 1: A B C
- 2: B C D
- 3: F G H

Clique Proximities: Prop. of clique members that each node is adjacent to

	1	2	3
A	1.000	0.667	0.000
B	1.000	1.000	0.000
C	1.000	1.000	0.000
D	0.667	1.000	0.000
E	0.000	0.333	0.333
F	0.000	0.000	1.000
G	0.000	0.000	1.000
H	0.000	0.000	1.000
I	0.000	0.000	0.333
J	0.000	0.000	0.000
K	0.000	0.000	0.333
L	0.333	0.333	0.000

La liste des cliques trouvées apparaît en premier lieu. Puis, un tableau indique la distance entre chaque nœud en ligne et chaque clique en colonne. B étant membre des cliques 1 et 2, la distance est de 1. Le nœud L est situé à une distance de 1 d'un élément d'une clique de taille 3, la valeur indiquée est donc de 1/3 soit 0.33. Le nœud A est lié à deux des 3 éléments de la clique BCD ; la valeur indiquée est donc de 2/3 soit 0.667 etc.

Actor-by-Actor Clique Co-Membership Matrix

		1	2	3	4	5	6	7	8	9	0	1	2
	A B C D E F G H I J K L												
1 A	1	1	1	0	0	0	0	0	0	0	0	0	0
2 B	1	2	2	1	0	0	0	0	0	0	0	0	0
3 C	1	2	2	1	0	0	0	0	0	0	0	0	0
4 D	0	1	1	1	0	0	0	0	0	0	0	0	0
5 E	0	0	0	0	0	0	0	0	0	0	0	0	0
6 F	0	0	0	0	0	1	1	1	0	0	0	0	0
7 G	0	0	0	0	0	1	1	1	0	0	0	0	0
8 H	0	0	0	0	0	1	1	1	0	0	0	0	0
9 I	0	0	0	0	0	0	0	0	0	0	0	0	0
10 J	0	0	0	0	0	0	0	0	0	0	0	0	0
11 K	0	0	0	0	0	0	0	0	0	0	0	0	0
12 L	0	0	0	0	0	0	0	0	0	0	0	0	0

HIERARCHICAL CLUSTERING OF OVERLAP MATRIX

	A	B	C	D	E	F	G	H	I	J	K	L
Level	1	2	3	4	5	6	7	8	9	0	1	2
2.000	.	XXX
1.000	XXXXX	.	.	XXXXX
0.750	XXXXXXXX	.	XXXXX
0.000	XXXXXXXXXXXXXXXXXXXXXXXXXXXX											

Les tableaux fournis ensuite par le logiciel montre le degré de coappartenance de chaque nœud aux différentes cliques. B et C sont dans les cliques ABC et BCD ; les croisements ligne/colonne dans la matrice indiquent 2.

La coappartenance est enfin représentée sous forme d'histogramme qui permet une visualisation rapide des principaux centres du graphe. L'interprétation des valeurs inférieures

à 1 est problématique. Il peut paraître étrange que A soit inclus dans le groupe ABC au niveau 1 et que D soit inclus au niveau 0.75 seulement. Ucinet considère les nœuds dans l'ordre où ils sont dans la matrice ; c'est pourquoi A est considéré en premier. D n'est examiné qu'ensuite ; mais pourquoi 0.75 ? N'ayant pas encore trouvé la réponse, je conseillerai volontiers de découper l'histogramme seulement à partir de 1....

La définition de la clique a été assouplie de diverses façons pour mettre en évidence des structures denses mais non complètement interconnectées.

La '**N-clique**' groupe les nœuds connectés séparés par une distance maximale N (2 est généralement utilisé, Ucinet le propose par défaut).

Les résultats de la commande 'Network > Subgroups > N-cliques' sont les suivants

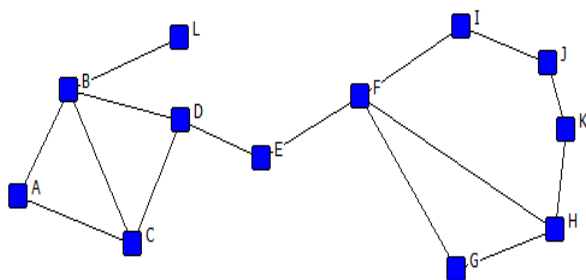
6 2-cliques found.

- 1: E F G H I
- 2: D E F
- 3: B C D E
- 4: A B C D L
- 5: F H I J K
- 6: F G H I K

HIERARCHICAL CLUSTERING OF OVERLAP MATRIX

	J	G	H	I	F	K	B	C	D	E	A	L
Level	1				1							1
	0	7	8	9	6	1	2	3	4	5	1	2
3.000	.	.	XXXXX
2.000	.	XXXXXXXX	.	XXXXX
1.800	.	XXXXXXXXXX	XXXXX
1.600	.	XXXXXXXXXX	XXXXXXXX
1.000	.	XXXXXXXXXX	XXXXXXXX	XXX
0.911	XXXXXXXXXXXX	XXXXXXXX	XXXXX	XXX
0.429	XXXXXXXXXXXX	XXXXXXXXXXXX
0.101	XXXXXXXXXXXXXXXXXXXXXXXXXXXX

Le plus simple pour comprendre le principe est de revenir au graphe de départ



Au départ, il y a 3 cliques (ABC, BCD, FGH). Le '**N-clique**' autorise l'intégration des nœuds situés à deux pas de chacun des nœuds de ces cliques. D et L sont à distance 1 ou 2 de la clique ABC, d'où la '**n-clique**' ABCDL. L'apparition d'une 2-cliques DEF est étrange mais logique ; ces nœuds sont connectés les uns aux autres et la distance maximale entre ses éléments est bien de 2... L'inconvénient de la méthode est évident ; cela favorise l'étirement des sous-groupes créés.

La recherche des ‘N-clan’ vise à limiter cet inconvénient en ajoutant une contrainte liée à la distance maximale autorisée entre les nœuds des groupes formés (diamètre). Sur un graphe de cette taille, les résultats ne sont pas très différents (‘Network > Subgroups > N-clan’ – diamètre maximal 2)

HIERARCHICAL CLUSTERING OF OVERLAP MATRIX

		G	E	F	H	I	J	K	A	C	D	B	L
5 2-clans found.								1	1				1
	Level	7	5	6	8	9	0	1	1	3	4	2	2
1:	E F G H I	-	-	-	-	-	-	-	-	-	-	-	-
2:	D E F	.	xxx	xxx	xxxxx	.	.	.
3:	B C D E	.	xxxxxxxx	xxxxx	.	.	.
4:	A B C D L	.	xxxxxxxxxx	xxx	xxxxxxxxxx	xxxxxxxxxx	.	.	.
5:	F H I J K	.	xxxxxxxxxxxxxxxx	xxxxxxxxxx	xxxxxxxxxx	xxxxxxxxxx	xxxxxxxxxx	xxxxxxxxxx	xxxxxxxxxx	xxxxxxxxxx	xxxxxxxxxx	xxxxxxxxxx	xxxxxxxxxx
	0.038												

Seule différence, la N-clique numéro 6 (FGHIK) n’est pas un(e) N-clan. Pourquoi ? La distance entre I et K, dans ce sous-groupe, est de 3, ce qui dépasse le diamètre maximal autorisé.

Une autre approche possible est de recherche les quasi-cliques, c’est-à-dire les sous groupes fortement connectés auxquels manquent quelques liens pour former une clique.

Les deux méthodes permettant cette recherche sont les ‘k-plexes’ et les ‘k-cores’. Pour les ‘k-plexes’, un nœud est membre d’une clique de taille n si il a des liens directs avec k-n membres de cette clique. Un ‘k-core’ est un groupe dont tous les nœuds sont liés à k membres du groupe.

Les résultats des deux méthodes sont les suivants (commandes ‘Network > Subgroups > K-plex’ et ‘Network > Regions > K-core’)

Les paramètres proposés par défaut sont gardés (K-plex : value of K 2, minimum size 3).

18 k-plexes found.

```

1: A B C D
2: A B L
3: B C L
4: B D E
5: B D L
6: C D E
7: D E F
8: E F G
9: E F H
10: E F I
11: F G H
12: F G I
13: F H I
14: F H K
15: F I J
16: G H K
17: H J K
18: I J K

```

HIERARCHICAL CLUSTERING OF OVERLAP MATRIX

	I	J	G	E	F	H	K	A	C	B	D	L
Level	9	0	7	5	6	8	1	1	3	2	4	2
4.000	.	.	.	XXX
3.000	.	.	.	XXXXX	XXX	.	.	.
2.133	.	.	.	XXXXXXXX	XXX	.	.	.
2.000	XXX	XXXXXXXX	XXXXX	.	.	.
1.853	XXX	XXXXXXXXXX	XXXXX	.	.	.
1.500	XXX	XXXXXXXXXX	XXXXXXXX	.	.	.
1.374	XXXXXXXXXXXXXXXX	XXXXXXXX	.	.	.
1.107	XXXXXXXXXXXXXXXX	XXXXXXXXXXXX	XXXXXXXX	.	.	.
0.040	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXXXXXXXX	.	.	.

Garder la taille minimale par défaut (3) donne un nombre très élevé de ‘k-plexes’ pour une raison logique. Il suffit d’avoir 2 liens entre 3 nœuds pour en créer un. La méthode permet la mise en évidence du groupe ABCD.

Les résultats des ‘k-cores’ se présentent sous une forme différente

K-CORE PARTITIONING

	A	B	C	D	E	F	G	H	I	J	K	L
Degree	1	2	3	4	5	6	7	8	9	0	1	2
2	XXXXXXXXXXXXXXXXXXXXXXXXXXXX
1	XXXXXXXXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXXXXXXXX

Partition Metrics - each column is a partition

		1	2
1	nClusters	1.000	2.000
2	CL1	1.000	0.917
3	CL2		0.083

LEGEND:

- (1)nClusters = number of clusters in the partition
- (2)CL1 = proportion of items that fall in cluster 1
- ...
 - (n)CLn = proportion of items that fall in cluster n

Coreness values for each node

```

1
-
A 2
B 2
C 2
D 2
E 2
F 2
G 2
H 2
I 2
J 2
K 2
L 1

```

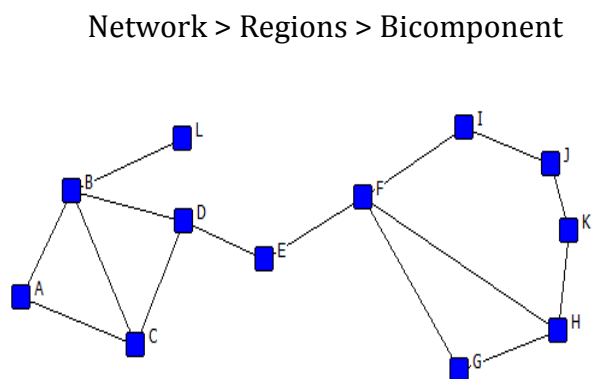
La position périphérique du nœud L apparaît, et c'est à peu près tout. Les résultats sont beaucoup plus significatifs avec des graphes de grande taille. Ils ont également un avantage non négligeable pour des géographes ; ils fournissent une image centre/périphérie nette et le codage des nœuds pour une future cartographie est fourni.

Les méthodes vues précédemment étudient les propriétés des nœuds pour déterminer les structures. Les deux méthodes suivantes partent du graphe complet.

Le premier indicateur est la présence ou non de '**component**' ('Network > Regions > Components') ; si le graphe est connecté, la valeur est de 1, si le graphe comporte deux sous-graphes non liés, la valeur est de 2 ; si le graphe comprend cinq nœuds isolés, la valeur est de 5 etc. Cela peut être utile avec des graphes de grande taille éclatés en nombreux sous-graphes (Ucinet fournit la taille et la composition des divers components).

Il peut être intéressant également de rechercher les nœuds qui, s'ils étaient supprimés, créeraient des sous-graphes non connectés entre eux.

L'examen du graphe GR2 montre que les zones de fragilité du graphe sont nombreuses. Supprimer B isolerait L, Supprimer D, E ou F entraînerait la création de deux sous-graphes etc. Ces nœuds sont appelés '**cutpoints**' dans Ucinet et sont donnés par la commande




```

Block 1: F G H I J K
Block 2: E F
Block 3: D E
Block 4: B L
Block 5: A B C D

```

Articulation points

```

          1
      CutPoint
      -----
1 A      0
2 B      1
3 C      0
4 D      1
5 E      1
6 F      1
7 G      0
8 H      0
9 I      0
10 J     0
11 K     0
12 L     0

```

Ucinet donne la liste des ‘cutpoints’, les ‘blocks’ peuvent être considérés comme les groupements de base du graphe.

D’autres méthodes sont également disponibles, fondées sur des propriétés différentes.

Les ‘**Lambda sets**’ sont les sous-groupes où le trafic potentiel est le plus important. Ucinet cherche combien de liens doivent être supprimés pour empêcher le lien entre deux nœuds. Ainsi, dans le graphe GR2, supprimer le lien BL rend L isolé, et l’empêche de joindre tous les autres nœuds (d’où la présence de 1 sur toute la ligne L – supprimer 1 lien suffit). A l’inverse, il existe 3 moyens de joindre B et C (BC, BAC, BDC). Pour les séparer, il faudrait donc supprimer 3 liens. L’intérêt est la mise en évidence des deux axes majeurs BC et FH.

Network > Subgroups > Lambda set

```

                                aximum Flow Between Pairs of Nodes
                                1  2  3  4  5  7  6  8  9 10 11 12
                                A  B  C  D  E  G  F  H  I  J  K  L
Input dataset:          GR2 (C  1  A  12  2  2  2  1  1  1  1  1  1  1
                                2  B  2 12  3  2  1  1  1  1  1  1  1
HIERARCHICAL LAMBDA SET PARTITIONS
                                3  C  2  3 12  2  1  1  1  1  1  1  1
                                4  D  2  2  2 12  1  1  1  1  1  1  1
                                A B C D E G F H I J K L
                                5  E  1  1  1  1 12  1  1  1  1  1  1
                                7  G  1  1  1  1  1 12  2  2  2  2  1
                                6  F  1  1  1  1  1  2 12  3  2  2  2  1
                                8  H  1  1  1  1  1  2  3 12  2  2  2  1
                                9  I  1  1  1  1  1  2  2  2 12  2  2  1
Lambda
-----
                                10 J  1  1  1  1  1  2  2  2  2 12  2  1
                                3    .  .  .  .  .  .  .  .  .  .  .  .
                                2    .  .  .  .  .  .  .  .  .  .  .  .
                                1    .  .  .  .  .  .  .  .  .  .  .  .
                                11 K  1  1  1  1  1  2  2  2  2  2 12  1
                                12 L  1  1  1  1  1  1  1  1  1  1  1 12

```

5.2 Directed graph

Dans le cas de ‘**directed graph**’, trois familles d’**equivalence**’ sont classiquement distingués

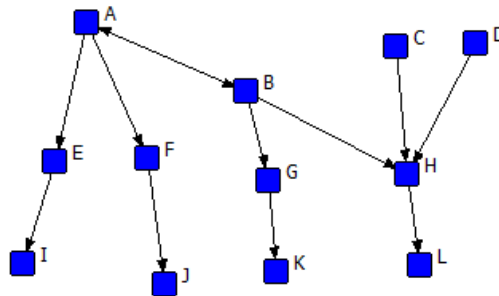
- ‘**structural equivalence**’ : mêmes relations avec les mêmes acteurs
- ‘**automorphic equivalence**’ : mêmes nombres d’**in**’ et de ‘**out-degree**’³
- ‘**regular equivalence**’ : mêmes types de relations avec le même type d’acteurs.

Les paramètres spécifiés par défaut sont souvent corrects mais il est peut être nécessaire de les modifier en fonction de ses besoins, notamment lorsque le logiciel propose par défaut un nombre donné de classes.

Pour comprendre le fonctionnement et surtout les différences entre les méthodes (et elles sont nombreuses), le plus simple est :

- de lire la rubrique d’aide correspondante
- de lancer une analyse et de regarder les fichiers en sortie
- mais surtout de garder ces résultats et de dessiner les partitions effectuées étape par étape.

Les deux exemples ci-dessous concernent le graphe GR1 qu’il convient de décrire.



On distingue au moins 4 types de nœuds. A et B échantent entre eux et envoient deux liens chacun. C et D ne font qu’envoyer. EFGH reçoivent et envoient mais H reçoit beaucoup plus. Enfin IJKL ne font que recevoir un lien.

STRUCTURAL EQUIVALENCE

Deux méthodes sont disponibles dans Ucinet, ‘**CONCOR**’ et ‘**Optimization**’.

³ Pour plus de détails, voir Wasserman et Faust pp.469-473

‘CONCOR’ est un algorithme itératif utilisant la corrélation⁴. Chaque étape donne une partition en deux sous ensembles. De l’aveu des spécialistes, les propriétés mathématiques de l’algorithme restent obscures, mais il est censé donner de bons résultats. Attention, les seules structures susceptibles d’apparaître sont des multiples de 2. Une structure du type centre / périphérie intégrée / périphérie délaissée ne peut donc pas être mise en évidence par cette méthode.

La commande ‘Network > Roles and Positions > Structural > CONCOR’ donne les résultats suivants (tous les paramètres proposés par défaut ont été gardés)

```

Relation 1
Blocked Matrix
      1 8 7 9 1 1 1
      A H G I K L J E F B D C
1 A | | | | | | | | | |
-----
8 H | | | | | | | | | |
7 G | | | | 1 | | | | |
-----
9 I | | | | | | | | | |
-----
11 K | | | | | | | | | |
12 L | | | | | | | | | |
10 J | | | | | | | | | |
-----
5 E | | | | 1 | | | | | |
6 F | | | | | | 1 | | | | |
-----
2 B | 1 | 1 1 | | | | | | |
-----
4 D | | 1 | | | | | | | |
3 C | | 1 | | | | | | | |
-----
Density Matrix
      1 2 3 4 5 6 7
1 -----
2 0.000 0.000 0.000 0.333 1.000 1.000 0.000
3 0.000 0.000 0.000 0.000 0.000 0.000 0.000
4 0.000 0.000 0.000 0.000 0.000 0.000 0.000
5 0.000 0.000 0.500 0.167 0.000 0.000 0.000
6 1.000 1.000 0.000 0.000 0.000 0.000 0.000
7 0.000 0.500 0.000 0.000 0.000 0.000 0.000
R-squared = 0.664

```

Cette méthode, comme les suivantes, donne des résultats sous la forme d’une matrice dont l’ordre a été modifié. Les permutations de lignes et de colonnes visent à mettre en évidence les blocs révélés par l’analyse. Le passage de 12 nœuds à 7 sous ensembles permet de simplifier l’image donnée par le graphe.

La matrice de densité fournit le nombre de liens réalisés sur le nombre de liens possibles entre les groupes formés. Que la diagonale soit tantôt vide, tantôt pleine de zéro est une des énigmes d’Ucinet. Dernier élément fourni, un dendrogramme qui permet parfois de comprendre certains détails de la partition. Cette méthode est présentée ici car il s’agit d’une méthode classique, mais l’expérience montre que les partitions obtenues sont souvent peu pertinentes.

Le choix de la méthode ‘**Optimization**’ suppose de réfléchir au préalable au nombre de blocs désirés ; il peut être utile de tester différentes valeurs pour garder la partition la plus pertinente. L’algorithme vise à créer des groupes minimisant la variance intra-groupe.

⁴ Pour plus de détails, voir Degenne et Forsé pp.113-116 ou Wasserman et Faust, pp.376-381

Voici les résultats obtenus avec les valeurs 2 à 4

2 : ABEF – CDGHIJKL

3 : A – BEF – CDGHIJKL

4 : AB – CD – EFGHIJKL – H

5 : A – B – CDIJKL - EF

AUTOMORPHIC EQUIVALENCE

Dans un graphe, deux nœuds présentant une ‘**automorphic equivalence**’ peuvent être permutés sans modifier la structure générale du graphe. Les regroupements successifs joignent les paires de nœuds ayant la même distance avec tous les autres nœuds du graphe. Dit plus simplement, il s’agit de regrouper les nœuds qui ont le même nombre d’**in** et de **out degree**.

Dans le graphe GR1, si C et D prennent la place l’un de l’autre, rien n’est modifié.

Les deux méthodes disponibles dans Ucinet donnent les résultats suivants

MaxSim (Network > Roles and Positions > Automorphic > MaxSim)

HIERARCHICAL CLUSTERING OF (NON-)EQUIVALENCE MATRIX

```
      A B C D H F G E I J K L
Level 1 2 3 4 8 6 7 5 9 0 1 2
-----
0.000  xxx xxx . xxxxx xxxxx .
5.651  xxx xxx . xxxxx xxxxxxxx
11.507  xxx xxx . xxxxxxxxxxxxxxxx
13.234  xxx xxx xxxxxxxxxxxxxxxx
15.614  xxx xxxxxxxxxxxxxxxxxxxxxx
20.241  xxxxxxxxxxxxxxxxxxxxxxxx
```

La position originale de H est soulignée, tout comme la particularité de L (un seul ‘**in degree**’ mais venant de H qui en a 3)

La deuxième méthode est la suivante ;

All permutations (Network > Roles and Positions > Automorphic > All permutations)

ORBITS:

```
Orbit #1: 1
Orbit #2: 2
Orbit #3: 3 4
Orbit #4: 5 6
Orbit #5: 7
Orbit #6: 8
Orbit #7: 9 10
Orbit #8: 11
Orbit #9: 12
```

Deux commentaires personnels sur cette méthode:

- elle prend un temps fou
- elle produit une division tellement fine qu'elle perd tout intérêt (à quoi bon créer 9 groupes à partir de 12 nœuds ?).

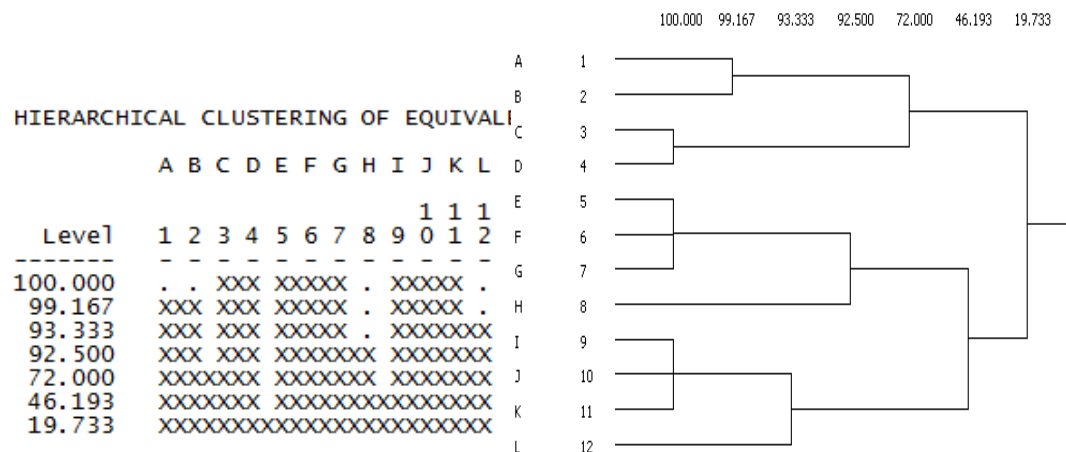
REGULAR EQUIVALENCE

Ces méthodes permettent de créer des groupes dont les structures relationnelles d'ensemble sont équivalentes. Certaines fonctionnent avec des matrices valuées, et donnent des résultats interprétables, même avec des écarts de valeur de 1 à 1000.

Une seule méthode est présentée ici REGE

Network > Roles and Positions > Maximal Regular > REGE

Voici les résultats obtenus avec le graphe GR1



Quand vous analysez de grands graphes, les dendogrammes sont rapidement illisibles. Avant de les enregistrer dans le format désiré, faites varier les valeurs des x et des y pour obtenir un résultat lisible.

Il arrive, même sur un ordinateur aussi récent que puissant, qu'Ucinet mouline longtemps. Le mieux à faire dans ces cas là est de ne surtout toucher à rien. Parfois, un résultat finit par apparaitre au bout de 5 à 10 minutes. Sinon un 'Esc' fera l'affaire. Cela est relativement fréquent avec des matrices lourdes et valuées (ce qui paraît logique).

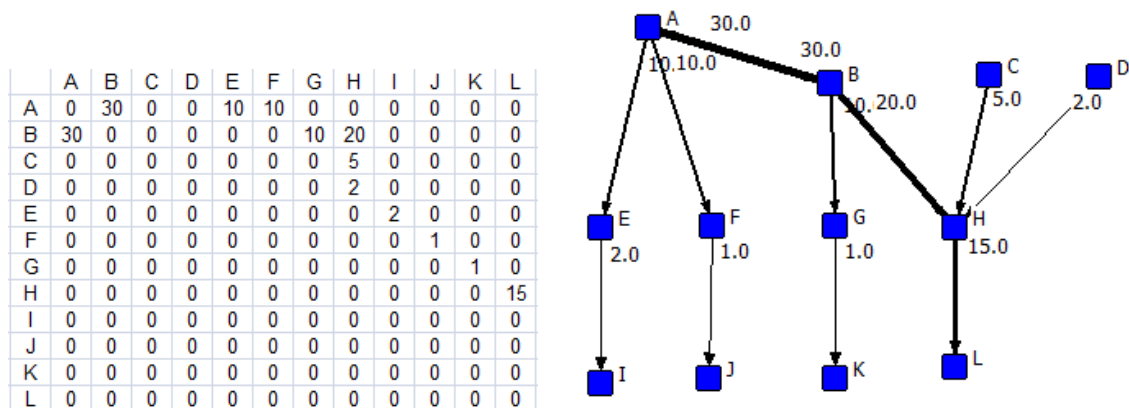
Dernier point sur les méthodes de partition : d'autres méthodes de partition (souvent voisines) existent, et il est très souvent possible de modifier les paramètres par défaut. La procédure la plus prudente consiste à :

- examiner attentivement le graphe
- choisir des méthodes en fonction des particularités du graphe ('binary' ou 'valued', 'directed' ou non)
- lire la documentation accessible via la touche F1
- observer graphiquement les résultats de la partition étape par étape. Mais pour ce dernier aspect, il faut lire le chapitre 6...

6. La représentation des résultats

La grande force de Netdraw est sa simplicité d'utilisation.

Le graphe utilisé pour ce chapitre est le graphe orienté et valué suivant (GR3)



Avant de jouer avec, un mot des icônes non encore explorées.



De gauche à droite

G, PC, MDS modifient l'aspect du graphe en fonction de propriétés mathématiques.

Eclair et Eclair = permettent de repositionner les différents nœuds.

Iso supprime les nœuds isolés

Pen supprime les nœuds qui ont un seul lien (degré 1). Cela sert à explorer graphiquement les graphes touffus pour déceler les nœuds principaux.

Self supprime les 'self-tie' (lien d'un nœud vers lui-même)

MC ne garde que les composants principaux du graphe

Ego permet de montrer l'ego-network d'un ou plusieurs nœuds

~Node permet de supprimer certains nœuds

~Tie permet de supprimer certains liens

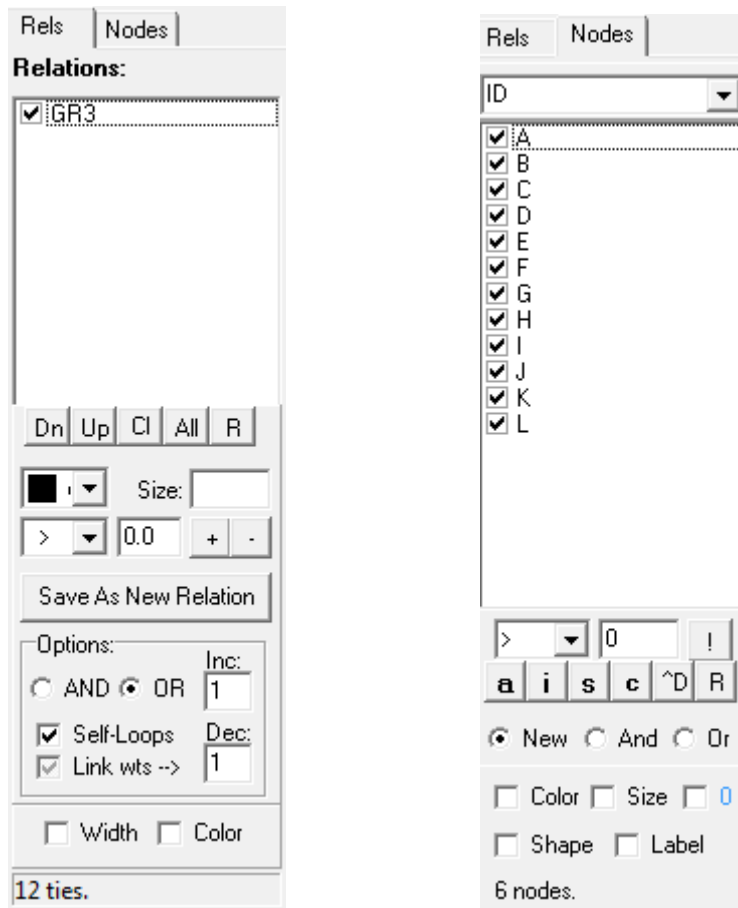
Grille de couleurs permet de colorer les nœuds en fonction de propriétés ('**attributes**')

Formes permet de modifier les formes en fonction d'attributs

Kim Bill permet de sélectionner l'étiquette des nœuds

La plupart des ces fonctions sont accessibles via les menus déroulants.

La barre d'outils situé sur la droite est très utile. Les deux onglets sont les suivants



Ces deux onglets permettent de modifier soit les relations, soit les nœuds.

Une fonctionnalité intéressante est qu'il est possible d'ouvrir, en plus du fichier principal, un fichier '**attribute**' (File > Open > Ucinet dataset > Attribute data). La seule condition est que les nœuds soient intitulés de la même façon – un message d'erreur apparaîtra si ce n'est pas le cas. Il est possible également d'ouvrir un fichier '**coordinates**' donnant des positions en x,y pour chaque nœud.

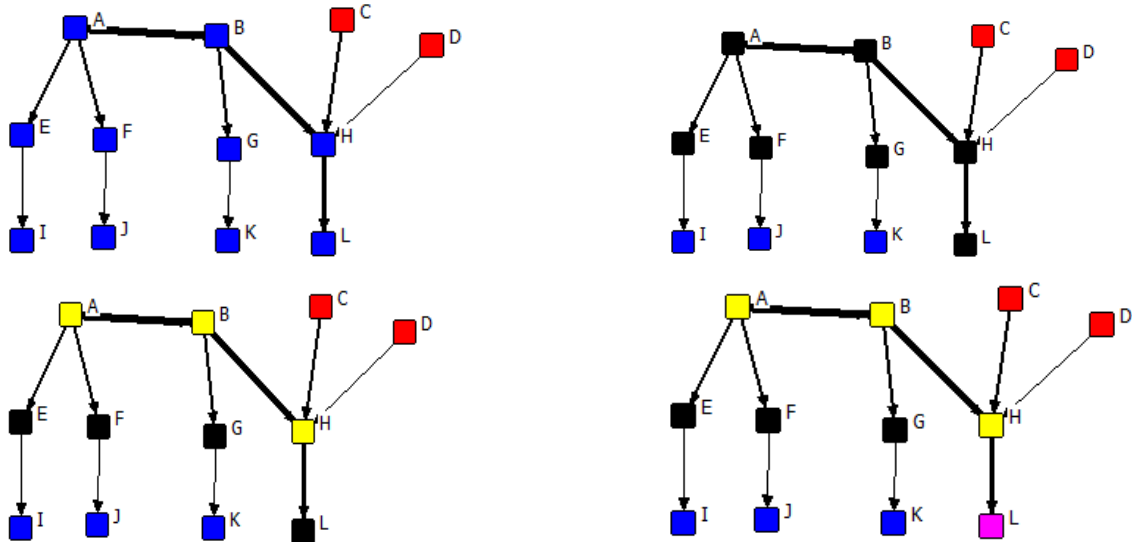
J'ai utilisé l'algorithme REGE sous Ucinet et nommé le fichier de sortie REGE_GR3.

J'ouvre le fichier GR3 dans Netdraw (Ucinet dataset > Network) puis le fichier REGE_GR3 (Ucinet dataset > Attributes data)

A l'aide de l'onglet ID, je peux choisir le niveau de la partition à représenter pour mieux comprendre les différentes étapes (il est possible également de passer par le menu déroulant Properties > Nodes > Symbols > Color > Attribute based). Il est théoriquement possible de transmettre une information avec la taille du nœud, une seconde avec la forme, une troisième avec la couleur... et la clarté du graphe sera compromise.

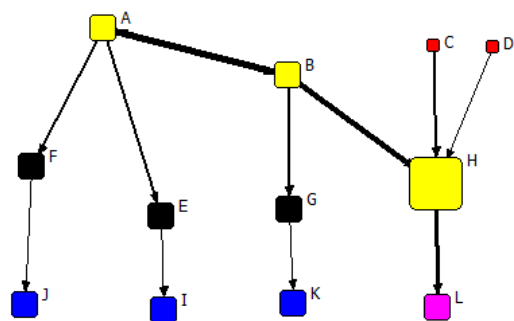
Les graphes ci-dessous montrent les étapes de la partition due à l'algorithme REGE. La couleur des nœuds varie en fonction du groupe d'appartenance. L'épaisseur des lignes varie en fonction de l'intensité du lien.

Figure 12 : Les quatre premières étapes de REGE sur le graphe GR1 (de gauche à droite et de haut en bas)



A l'étape 4, j'ai calculé les mesures de centralité (Analysis > centrality, directed versions), décoché l'onglet 'color' dans l'onglet 'nodes' situé à gauche, coché l'onglet 'size' puis, dans le menu déroulant de l'onglet, choisi 'in-degree'.

Le résultat est le suivant.



Attention, chaque fois que vous modifiez un graphe (ajout ou suppression de liens ou de nœuds), les indicateurs doivent être recalculés.

En guise de conclusion provisoire

Ucinet et Netdraw sont d'excellents logiciels proposant de très nombreuses méthodes d'analyse. Il convient cependant de souligner un inconvénient ; ils ne permettent de traiter que de petites matrices (256*256). Or, pour quelqu'un travaillant sur des réseaux de villes par exemple, cette limite sera vite dépassée. Il deviendra dans ce cas nécessaire de se former à des logiciels permettant de traiter ces grosses matrices (Pajek ou Tulip par exemple).

Reste que pour se former à l'analyse des réseaux sociaux, ces deux logiciels sont tout à fait indiqués.

Comme indiqué page 1, vos remarques sont bienvenues. La version améliorée et augmentée prévue pour l'automne 2010 intégrera les éléments d'analyse pour les types de graphes non abordés ici ('**signed**', complexe et '**multigraph**').

Eléments de bibliographie

La bibliographie sur le sujet est abondante. Toute personne désireuse d'approfondir tel ou tel aspect trouvera dans les ouvrages indiqués les indications complémentaires.

Il existe un bon mode d'emploi pour Ucinet en anglais téléchargeable à l'adresse suivante (<http://faculty.ucr.edu/~hanneman/>). Le logiciel a évolué depuis et il faut fouiller un peu pour retrouver certaines commandes, mais il peut utilement compléter ce livret.

Les non anglophones trouveront une introduction solide dans le Degenne et Forsé (notamment aux chapitres 3,4 et 6).

Le classique de Wassermann et Faust commence à dater et une édition actualisée serait nécessaire, mais il reste mon avis l'ouvrage le plus complet et le plus clair sur le sujet.

Parcourir la revue *Social networks* est un bon moyen de connaître les pistes de recherche actuellement explorées, notamment en ce qui concerne l'analyse des graphes valués et des multigraphes.

Les travaux de L. Freeman sont tous disponibles sur son site Internet à l'adresse suivante (<http://moreno.ss.uci.edu/pubs.html>), ceux concernant la représentation des résultats sont particulièrement intéressants.

DEGENNE A., FORSE M., 1994, *Les réseaux sociaux*, Paris, Armand Colin, collection « U sociologie ».

HANNEMAN R., RIDDLE M., 2005, *Introduction to social network methods*, Riverside, CA, University of California, Riverside

WASSERMAN S., FAUST K., 1994, *Social Network Analysis. Methods and applications*, Cambridge, Cambridge University Press, coll. Structural analysis in the social sciences".

Les sites suivants sont riches en documentation diverse et variée, tant sur les méthodes que sur les logiciels disponibles

<http://www.insna.org/>

<http://supernet.som.umass.edu/> (la homepage est atroce, mais le site riche)

Annexe I : Petit lexique anglo-français

Ce lexique définit brièvement les termes utilisés dans Ucinet, Netdraw et dans les pages précédentes. La traduction proposée dans le manuel de Degenne et Forsé est indiquée en italiques et entre parenthèses.

Adjacent : un nœud est dit adjacent à un autre lorsqu'il lui est relié par un lien

Automorphic equivalence (*équivalence forte ou automorphique*) : deux nœuds partagent une 'automorphic equivalence' s'ils possèdent le même nombre d'**in** et de **out degree**.

Clique (*clique*) : sous graphe complet de 3 nœuds minimum

Complete [graph] (*graphe complet*) : graphe où tous les liens possibles entre les nœuds sont présents. Sa densité est donc égale à 1.

Connected [graph] (*graphe connexe*) : désigne un graphe où tous les nœuds sont accessibles.

Component (*composants*): parties du graphe non connectées les unes aux autres. Dans un 'connected graph', le nombre de 'component' est 1.

Cutpoint (*point d'articulation*) : nœuds du graphe qui, s'ils étaient supprimés, créeraient des 'components'

Degree (*degré*): nombre de liens adjacent à un nœud. Dans le cas de 'directed graph', on distingue les '**in**' (*demi-degré intérieur*) et les '**out degree**' (*demi-degré extérieur*).

Density (*densité*): nombre de liens réels divisé par le nombre de liens théorique.

Diameter : le diamètre d'un graphe est la plus courte distance entre les deux points les plus éloignés d'un graphe connecté

Directed [graph] (*graphe orienté*): graphe dont le sens de la relation importe (AB n'est pas obligatoirement égal à BA). La matrice correspondante est asymétrique.

Ego-network (*réseau personnel*): un ego-network est constitué par un nœud, les nœuds qui lui sont adjacents, et les liens existants entre ces derniers.

Geodesic distance (*distance géodésique*): distance la plus courte, mesurée en nombre de liens, entre deux nœuds.

Graph (*graphe*): équivalent de network. Au sens strict, un graphe désigne un ensemble de points et de liens dont on étudie les propriétés mathématiques. Dès que des informations sont ajoutées (ex. nom des nœuds), on parle de réseau.

Isolate (*isolé*): nœud qui n'est adjacent avec aucun autre nœud.

Loop = self-tie (*boucle*): lien d'un nœud vers lui-même.

Multigraph (*multigraphe*) : graphe représentant simultanément plusieurs types de liens entre les nœuds.

Network (*réseau*): ensemble de nœuds et de lignes portant des informations.

Node (*sommet*): synonyme d'**actors**, traduit ici par nœud.

One-mode [network]: graphe qui comporte une seule catégorie d'acteurs, donc dans lequel potentiellement n'importe quel acteur peut être relié à n'importe quel autre.

Reciprocal tie : dans un graphe orienté, signifie que les liens AB et BA sont présents.

Regular equivalence (*équivalence régulière*) : deux nœuds présentent une '**regular equivalence**' s'ils partagent le même type de liens avec le même type de nœuds.

Self-tie : voir **loop**

Signed [network] (*graphe signé*): les liens entre nœuds peuvent prendre 3 modalités (-/0/+). Ils peuvent traduire des sentiments (n'apprécie pas/sans opinion/ apprécie) soit des soldes (déficit/pas de relation/excédent).

Simple [network] : graphe qui représente un seul type de liens entre les acteurs et qui ne comprend pas de liens d'un nœud vers lui-même.

Size (*ordre*): la taille du graphe correspond au nombre de nœuds qu'il contient

Structural equivalence (*équivalence structurale*): deux nœuds sont structurellement équivalents s'ils ont exactement les mêmes relations avec les mêmes acteurs ;

Tie (*liaison*): lien entre deux nœuds. Traduit ici par lien, relation. On distingue généralement les '**arcs**' (liens orientés - *arcs*) et les '**edges**' (liens non orientés - *arêtes*)

Two-modes [network] (*matrice acteurs-événements*): graphe qui comporte deux catégories d'acteurs (ex. relations entre un groupe d'enseignants et un groupe d'élèves) ou une catégorie d'acteurs et une catégorie d'événements (ex. un groupe de chercheurs, un groupe de revue scientifiques). Un nœud appartenant à une catégorie ne peut être relié qu'à un nœud appartenant à l'autre catégorie.

Valued [network] (*graphe valué*): graphe où l'intensité de la relation importe (ex. matrice de flux commerciaux entre pays)

Table des figures

Figure 1 : Lancer l'algorithme CONCOR.....	3
Figure 2 : Page d'accueil d'Ucinet.....	4
Figure 3 : Page d'accueil de Netdraw	4
Figure 4 : Menu File, Data et Transform	6
Figure 5 : Menu Tools.....	7
Figure 6 : Menu Network.....	7
Figure 7 : Directed binary matrix and graph (GR1).....	10
Figure 8 : Undirected binary matrix and graph (GR2).....	10
Figure 9: Directed valued matrix and graph (GR3)	10
Figure 10: Star network and circle graph	16
Figure 11: Line network and complete network	16
Figure 12 : Les quatre premières étapes de REGE sur le graphe GR1 (de gauche à droite et de haut en bas).....	32

Table des matières

1. Présentation des logiciels	2
2. Les interfaces.....	4
3. Questions de vocabulaire.....	9
4. Premiers pas, premières mesures	11
5. A la recherche des structures.....	19
<u>5.1</u> Non directed graph	19
<u>5.2</u> Directed graph	26
6. La représentation des résultats	30
En guise de conclusion provisoire.....	33
Eléments de bibliographie	34
Annexe I : Petit lexique anglo-français	35
Table des figures	37
Table des matières	38