PARENTE SUITE USER'S GUIDE

(c) 1994-97

Douglas R. White

Social Science, U. C. Irvine 92697

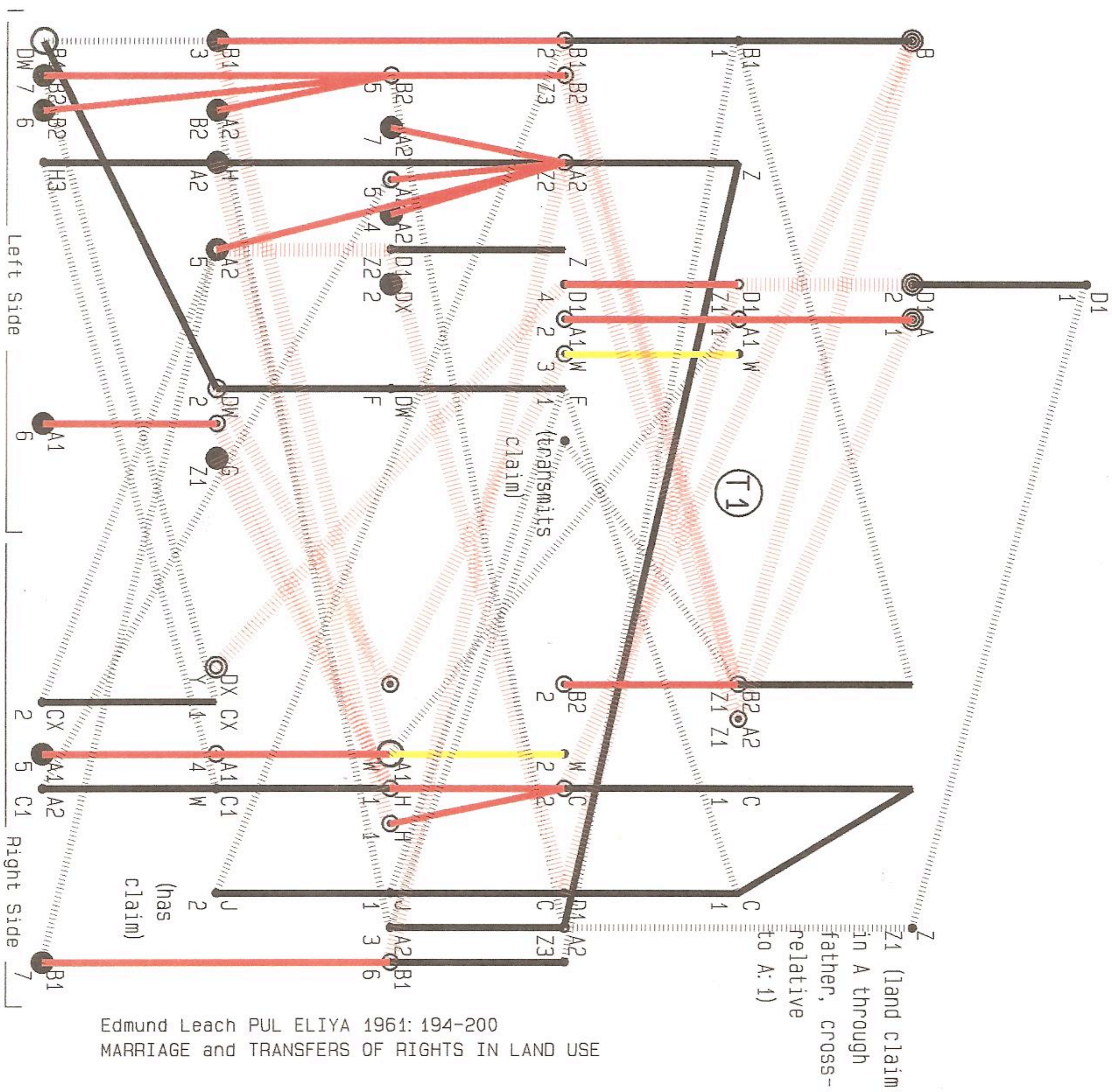
BETA 2.6 Version, May 1997

please report bugs
to drwhite@uci.edu

VOL. 1, Containing

Six Steps for Kinship Graphs:
from Genealogical Data ~~to Graphs~~ to Network Analysis

Edited by Patricia Skyhorse



Edmund Leach PUL ELIYA 1961: 194-200
MARRIAGE and TRANSFERS OF RIGHTS IN LAND USE

PARENTE SUITE USER'S GUIDE

Douglas R. White
Social Science, U. C. Irvine 92697

VOL. 1

Containing

Six Steps for Kinship Graphs: from Genealogical Data to Network Analysis

Edited by Patricia Skyhorse

DOS Version 2.6 May, 1997

PGRAPH (2.6): Representation and Analytic Program for Kinship and Marriage Networks

Shareware Registration: $10US, 15DM or 50FF

Courtesy of the working groups in discrete structures
in the social sciences, sponsored by the
Maison des Sciences de l'Homme and the Maison Suger,
with support of University of Cologne, Institute of Ethnology,
funding from the A. von Humboldt Foundation
and Leibnitz Award to Thomas Schweizer

===

PREFACE: Acknowledgements and Design

===

Inspired by the work of Lévi-Strauss (1949) and Weil (1949), the prototype of the p-graph (having parental couples as nodes rather than individuals) was developed in 1960 by Guilbaud (1970; see Cuisinier 1962, Bertin 1967). Extending this representation from marriage system models to kinship network applications, the program analyzes p-graph representations (White and Jorion 1992 Current Anthropology 33: 454-63) of kin connections among marriages 1 to N, linked through male and female genealogical ties, using vector genealogy as developed in Jorion and Lally (1983). See SPONSORS/PRECURSORS for further bibliography.

Thanks to Patricia Skyhorse for interface suggestions and major revisions of the user manual.

Special thanks to Clemens Heller, Maison des Sciences de l`Homme and the French Ministry of Research and Technology for support of the Discrete Structures in the Social Science working groups that made the programming and larger project possible, and to Paul Jorion, Michael Houseman, Alain Degenne, Thomas Schweizer, Vincent Duquenne, François Héran and Lin Freeman for collegial support.

In the design of this family of programs for network analysis of kinship data, the P-graph framework is central. The P-graph is an ordering of kinship relations between couples by individuals who link their parents coupling to their own parental coupling. It reverses the conventions of ordinary social networks since in P- graphs individuals are the lines or directed arcs and couplings are the points or vertices. It is also a dual intersection graph ordered by generations, where upward chains intersect in ever more inclusive sets of ancestors while downward chains of descent lines intersect in marriages. The P-graph, in contrast to assumptions that structural analyses are ahistorical, provides for inclusion of the percolation of historical time. It is only as a second step that we reduce the transactions represented in the graphs to cognitive transformational structures repeating themselves in time, that may represent indigenous people or observers "thinking about kinship" as social rules and conventions. The mathematics of kinship, in either case, attempts not to define the content of kinship, but takes culturally defined elements as givens and focuses instead on the relations among them.

The P-graph was initially stimulated by the work of Lévi-Strauss and Weil in 1949 and given form by Guilbaud in 1960 at l'Ecole des Hautes Etudes en Sciences Sociales. It was only one of the varieties of kinship graphs developed by Guilbaud (see Cuisinier 1962, Bertin 1967, Guilbaud 1970). Guilbaud's graphs were mainly applied to permutation group models of kinship (Weil 1949). The group-theory approach to kinship, stressing models of a timeless repetition of the transformational structures of kinship, fit Lévi-Strauss's concept of kinship as a "cognized" domain in which the universals of thought could make sense, to quote a phrase he cited from Goldenweiser, out of the "impossible complexity" of actual kinship relations.

My own exposure to this framework (the formalism is discussed in section 2.1), initial ideas for blood-kin marriage computations, and the stimulus for the graphic representation were provided by Paul Jorion (White and Jorion, 1992, "Representing and Analyzing Kinship: A Network Approach," CURRENT ANTHROPOLOGY; and 1996, "Kinship Networks and Discrete Structure Theory: Applications and Implications," SOCIAL NETWORKS). It was his contribution to show how P-graph data are given as vectors for an ordered array of couples. Parental couples are given in one array for males, in another for females. I added optional dates of birth and death, lineage membership, arrays for couples including labels (numeric and text), row and column coordinates, colorings, and so forth. Computations on kinship relations use vector operations wherever possible, which greatly enhances the speed at which they are carried out. More general ideas for analysis of marriage alliance in the absence of blood-kin marriage were developed in collaboration with Michael Houseman (see references). For suggesting our possible collaborations the three of us are indebted to Françoise Héritier. Two years into the project I met François Héran, who was then completing his publication on P- graphs (H,ran 1995), and have benefited since from his suggestions and insights.

In contrast to Andre Weil's group-theory approach to kinship, the mathematical idea utilized here of the P-graph as a network derives from theories of ordered sets (partial orders, Galois lattices). Vincent Duquenne oriented me to the foundations of lattice and ordering theories, on which the current approach is based, and gave helpful stimulus and advice on programming.

Preliminary work on the programs in 1991 was supported by the Maison des Sciences de l'Homme, the Maison Suger, and the Ministère de la Recherche et de la Technologie, within the framework of an international and interdisciplinary working group on discrete structures in the social sciences created around the support and research facilities of the Maison Suger. Support for further programming developments during 1992 was provided by Alain Degenne's LASMAS research group at IRESCO and by a Bourse d'haute niveau from the MinistŠre de la Recherche et de la Technologie. Degenne and I had both done previous work on network homomorphism, some of which finds its place in the current approach to kinship networks.

Parenté Suite: Kinship Network Analysis Software User's Guide

TABLE OF CONTENTS Page

PREFACE: ACKNOWLEDGEMENTS AND DESIGN........................ 3

1.0 GETTING STARTED... 9

1.1 Hardware.. 9

1.2 Installation.. 9

1.3 Technical Support...................................... 10

1.4 Citing the Program..................................... 10

VOL. ONE: TUTORIAL

2.0 INTRODUCTION AND CONVENTIONS........................... 11

2.1 PGRAPH the Program and P-Graphs the Object.......... 11

2.1a What is PGRAPH, anyway?.............................. 11

2.1b So what is a P-Graph?................................ 11

2.2 Notational Conventions............................... 11

2.2a Emphasis... 12

2.2b Screen Output.. 12

2.2c Input Requests....................................... 12

2.2d Wild Cards... 13

2.2e Defaults... 13

2.3 DATA ENTRY & CONVERSION: Standard Genealogical Format: TUTORIAL STEP ONE: PREPARING A KINSHIP DATA SET...... 13

2.3a Data Files... 13

2.3b Demo.txt... 15

2.3c Primary data formats................................. 15

2.3d Fortran FORMAT Statements............................ 16

2.4 DATA CONVERSION EGO2CPL program..................... 16

TUTORIAL STEP TWO: TRANSFORMING THE DATA............. 16

2.4a How to Run the Program............................... 16

2.4b Processing Options................................... 18

2.4c Types of output...................................... 18

2.4d End of File check.................................... 19

2.4e Error Checking-Proxy Parents......................... 19

2.4f Errors file.. 20

2.4g Inconsistencies in the data 21

2.4h Marriage Statistics.................................. 21

2.4i Size or Memory Limits of Ego2Cpl..................... 22

3.0 THE PGRAPH ENVIRONMENT................................. 23

TUTORIAL STEP THREE: PROCESSING A DATA SET WITH 23
PGRAPH (ver. 2.6): Parental or Pair Graph Analysis

3.0 P-GRAPH Menus and Help........................... 23

3.0a Aborting the Program.............................. 23
3.0b Starting the Program.............................. 23

3.1a Basic Tutorial using simplest options 23
3.1b Graphic Defaults and Choices...................... 24
3.1c Data Processing Options for New Slides............ 25
3.1d Data Format Request for New Slides................ 25
3.1e Choosing a Data set............................... 26
3.1f Descent Order Rule for New Slides................. 26
3.1g Generational Ranking Option....................... 27
3.1h Normal Descent Order for New Slides.............. 27
3.1i Vertical Descent Line Continuity.................. 27
3.1j Backup Coordinate Files........................... 27
3.1k Endogamy.. 28
3.1l Disconnected Part Options for New Graphs........... 28
3.1m Draw by Generations............................... 28

3.2 Viewing or Editing a Slide that Exists............. 29

3.2a Where Do I GO From Here? 29
3.2b Vector Format: Strings and Cores for New Graphs..... 29
3.2c Demographic Format covered under ADVANCED OPTIONS 29
3.2d Examine, Edit, or analyze Existing Kinship Graphs... 30
4.0 KINSHIP GRAPHS WITH PGRAPH 31
TUTORIAL STEP FOUR: EDITING A PGRAPH SLIDE 31

4.1 Capabilities: P-graphs............................ 31
4.2 Basic Editing of a Slide: PGRAPH Menu Choices...... 31
4.3 The PGRAPH Screen................................. 31
4.4 PGRAPH Ordering of Generations.................... 31
4.5 PGRAPH Slide menu Commands and Subcommands......... 32
4.5a SLIDE MENU Options................................ 32
4.5b Column Editing 32
4.5c Row Editing 32
4.5d Drawing Options................................... 33
4.5e Viewing Options 34
4.5f Slide Options: save, export....................... 34
4.5g Placing a Slide 35
4.5h Quitting a Slide to go to the next 35
4.5i Quitting Slide menu to return to main menu 35
4.6 Quitting the program 35
4.7 Size or Memory Limits of PGRAPH................... 35
5.0 PRINTING YOUR P-graphs
TUTORIAL STEP FIVE: PRINTING A P-graph SLIDE.......... 36
Douglas White
5.1 A Note on Vector versus Raster (Pixel) Images....... 36
5.2 PAR-PLOT does P-graph HPGL VectorGraphic Instructions36
5.3 Running PAR-PLO.BAS to create a *.PLT file.......... 36
5.4 How to get colors by editing your *.PLT files....... 37
5.5 Shareware Program SPLOT to simulate the Pen Plotter..39
5.6 LP and PrintaPlot pen plotter simulation programs....40
5.7 Postscript file for printing with DOT graphics 41

6.0 KINSHIP STATISTICS: THE PAR-CALC PROGRAM...............42
TUTORIAL STEP SIX : Par-Calc - Consanguineal Marriages 42
Douglas White and Paul Jorion

6.1 The Order in which to run the programs.............. 42
6.2 Capabilities of PAR-CALC: Marriage Statistics....... 43
6.3 Running PAR-CALC: Menu Choices and Help............. 43
6.4 PAR-CALC Output Files............................... 45
7.0 WHERE TO GO IN YOUR ANALYSIS............................ 49

7.1 Social Networks analysis............................ 49
7.2 Network Analysis with UCINet........................ 50
7.3 Economic, Social and Cultural Capital............... 50
7.4 Processes of Percolation............................ 51
7.5 Analysis of Relinking: Group, Class, Ethnicity...... 51
7.6 Mapping Exchange and other relations and properties 51
7.7 Simulation and Hypothesis Testing................... 52

===
8.0 OUTPUT FILE STRUCTURES 54

8.1 FILE NAMING CONVENTIONS............................. 54
8.1a The .VE* data files: .VED VES VEC and .VEM.......... 54
8.1b The .DO* text and .DES data files................... 55
8.1c Special Features of the .DES File Format............ 55
8.1d The .CR* graphics files............................. 55
8.2 PREPARING DATA in VECtor format [empty]............ 55
8.2a Data Formats for .VED Vector Files.................. 55
8.2b Data Formats for .CRD Vector Files.................. 55
8.2c Data Formats for .DES Vector Files.................. 55
8.3 Demographic Data in .DEM Format.................... 55
8.4 Time and Social Process............................. 55
9.0 ADVANCED OPTIONS INSIDE PGRAPH 55

9.1 Introduction..
9.2 Advanced Slide and Menu Options..................... 55
9.2a Changing Colors..................................... 55
9.2b Allocating Memory for Generational Depth beyond 20.. 55
9.2c Allocating Memory for Non-Implicit Labels........... 55
9.3 Generational Editing 55
9.4 Simulation: Random Permutations of Marriages........ 55
9.5 Capabilities: M- and C-graphs 55
9.7 Exchange of Siblings M-Graphs .... 56 ** 9.6
Creating P-Graph Homomorphisms 55
9.8 Using Homomorphisms to define new kinship slides.... 59
10.0 AUXILIARY DATA AND LINKS TO OTHER PROGRAMS............ 59

10.1 Visualized Algebra: links to NetGraph Package....... 59
10.2 Links to UCInet [empty] see 7.2 59
10.3 PAR-BLOC for marriage cycles (PAR-LINK obsolete).... 60
10.4 PAR-SIDE.. 60
10.5 PAR-COEF.. 61

10.6 Time-Sorted PAR-CALC OUTPUT........................ 61
10.7 Secondary Analysis from the PAR-CALC Output Files... 62
10.8 PAR-FLIP and Entailment Analysis...................
11.0 UTILITY PROGRAMS.................................... 68

11.1 UTILITY Program: PAR-DOUB for double descent........ 68
11.2 UTILITY Program: PAR-UP 69
11.3 UTILITY Program: PAR-DOWN 69
11.4 UTILITY Program: PAR-NAME 69
11.5 UTILITY Program: PAR-MARR 69
11.6 UTILITY Program: PAR-TERM for Cognized Kinship...... 69
12.0 TROUBLESHOOTING 70

12.1 What if no files are shown on the menu in 3.1g?
[The SORT program and its workings]................ 70
12.2 Extra Memory for Non-Implicit Labels................ 70
13.0 PROGRAMMER'S GUIDE 70
13.1 Connected Components and Generations: By Vectors 90

==

1.0 GETTING STARTED

== There are three versions of PGRAPH: two for DOS, and the other a prototype windows version, PGRAPHW. Instructions for the Tutorial section are similar.

1.1 Hardware

PGRAPH requires a computer running MS-DOS, or other DOS compatible operating systems. All versions of DOS (2.0 or higher) are supported, as well as all CPU chips that DOS runs on (8088, 8086, 80286, 80386, 80486, etc.) and math co-processor chips (8087, 80287, 80387, 80487, etc.). This version of PGRAPH program is 32 bit and needs at least 4 mg of RAM and the presence of DOS extenders DOSXNT.386 and DOSXMSF.EXE in order to run. These programs utilize DOS Graphics which means you must start the programs from the DOS prompt, as windows does not support DOS Graphics. All versions of PGRAPH also require SORTF.COM in order to run and courier.fon, tmsrmn.fon, QBasic.exe, PAR-PLOT.exe, and LP.exe or other printer utilities (section 5) to print slides.

If you have a DOS version of the program and your system is running Windows 95 contact the author for instructions on running the program, which uses DOS Graphics, which are not supported in Windows 95, or obtain the Windows version of the program, PGRAPHW. PGRAPHW requires an IBM compatible computer running MicroSoft Windows 3.1 or higher.

PGRAPH (DOS version) requires a VGA graphics screen but PAR-CALC does not. PGRAPH- is the version which runs on 8088 processors and doesn't require so much memory, but of course, doesn't handle very large data sets.

1.2 Installation

You must have about 3 Megabytes of disk space free to install to a hard drive. To create a directory on your hard disk, named PGRAPH for example, place the PGRAPH diskette in your 3.5" floppy drive (for example B:), and type:

```
MD PGRAPH CD PGRAPH B:pkunzip B:PARDOS.zip
```

The PARDOS.zip file unzips to unlock all of the programs, data files, and fonts when it is executed. If you install to a network the hard disk may be write protected after installation. If so the programs will read data files from the hard drive but will expect to write to files in floppy disk drive A: so you must have a formatted disk installed in A:

An alternative way to install the program is to copy all files on the diskette to your hard drive and then use pkunzip to unzip all the files with a ".zip" file extension.

1.3 Technical Support

Owners of the Professional edition and teachers purchasing an Instructional license are entitled to unlimited technical support. Please feel free to write or call the author during working hours with any questions, suggestions and bug reports at the following address, although the quickest way to get a response is to send email to: drwhite@sdsc.edu

Douglas R. White Tel: (714) 824-5893 (messages) Social Science Plaza
824-5041 (secretary) University of California, Irvine Irvine, Ca 92697 E-mail:
drwhite@uci.edu

1.4 Citing the Program

If you use any of the programs in the Parente Suite to perform any analyses which result in a publication, please cite the program in that publication (this is part of the licensing agreement). The citation should treat the program as though it were a book. For example: White, Douglas R. 1997. PGRAPH Version 2.6. Irvine, CA, University of California, Irvine.

1.5 License and Limited Warranty

This software is protected by both United States copyright law and international treaty provisions. You must treat this software like a book, except that (i) you may copy it to the hard disk of a single computer or to the hard disk of any computer(s) that you have exclusive use of, and (ii) you can make archival copies of the software for the sole purpose of backing it up to protect your investment from loss.

By "treat this software like a book" we mean that the software may be used by any number of people, and may be freely moved from one computer to another, so long as there is no possibility of it being used simultaneously on two different computers. Just as a given copy of a book cannot be read by two different people in different places at the same time, neither can the software be legally used by different people on different computers at the same time.

....Limited Warranty [basically, replacement of faulty original diskettes within first year. This License and Limited Warranty shall be construed, interpreted and governed by the laws of the State of California....

The author and published specifically disclaim all (other) warranties, express or implied, including but not limited to, any implied warranty of merchantability or fitness for a particular purpose. They make no representation or warranty that the software is fit for any particular purpose and any implied warranty of mercantability is limited to the ninety-day duration of the Limited Warranty covering the physical diskette(s) and physical documentation only (and not the software) and is otherwise expressly and specifically disclaimed.

==

2.0 TUTORIAL SECTIONS: INTRODUCTION AND CONVENTIONS

==

2.1 PGRAPH the Program and P-Graphs the Object

2.1a WHAT IS PGRAPH, ANYWAY?

PGRAPH is a program that recodes relationships among individuals -- people and their spouses and parents -- into a different form where we arbitrarily assign successive numbers to couples and the remaining single individuals and then we record the existing parental relationships between them. More precisely, each element -- couple or singleton individual if not coupled -- in the series is assigned an index number in a vector, and two functions are defined, F and G, such that for any element x in this vector, F(x) will return the number of the female's parents and G(x) the number of the male's parents. The numbers of the parents are elements in the original vector.

When only one person in a node has a number from the data set (e.g., is single), the other half of the node has no specific individual number. We use three designations for the numbers of unspecified partners: 0 indicates coupled-but spouse unknown, -1 indicates known to have never married, and -2 indicates not coupled yet, but could marry later.

The advantage of this way of coding relationships is that it is used to draw compact and analytically useful graphs of kinship and marriage networks, of virtually any size, and these graphs can be inspected, edited and analyzed for properties of interest to the researcher.

2.1b SO WHAT IS A P-graph?

The set of points and the directed arcs connecting points that are defined by the ordered pairs <x,F(x)> and <x,G(x)> constitute the digraph that we call a p-graph. In a p-graph drawn from genealogical data, this digraph is asymmetric and acyclic since directed cycles and reciprocal pairs of arcs are impossible (no one is their own ancestor or their parent's parent). Like genealogy, the p-graph is ordered, and can be arranged into generations or levels.

TIME is the first dimension of a p-graph, drawn vertically.

A second dimension, drawn horizontally, is DESCENT. Usually, we put people together in groups that are descended from the same ancestors. Usually we think of people descended from a common ancestor as a kind of ancestral tree. What kinds of trees of descent we use to group people, however, might depend on whether the have a rule that emphasizes a single line of descent -- patrilineal through males or matrilineal through females. The alternative is to keep people grouped together because they are bilateral descendants of one or more common ancestors.

A third dimension that emerges from analysis is the extent to which people are embedded in endogamous relationships in which they are not only linked by blood, but linked and relinked by marriage. The density of these kinds of ties is a feature of the MARRIAGE ALLIANCE patterns of a kinship network

2.2 Notational Conventions

2.2a Emphasis

In this manual, we represent things we want to emphasize by <u>underlining</u>. For example: The word number is usually abbreviated as "<u>No</u>."

2.2b Screen Output

Requests for input from you and other screen output will be in bold or italics, for example: The program asks for: *FILENAME?*

2.2c Input Requests

We represent keys to be pressed in bold-face letters, such as **Esc** or **Enter**. **Enter**, **Return**, **Carriage Return**, and a symbol consisting of an arrow pointing down and then to the left: all represent the same key, which has **<----Ù** different names on different computers.

Information that you must enter is in quotes and double- underlined. If the manual says to type the command

"PGRAPH" and press **Enter**/**Return**

you are supposed to type PGRAPH and press the Enter key, you are not supposed to type the quotes.

2.2d Wild Cards

In DOS, the character *, or asterisk, is called a wild card. It can represent any character. So if we refer to a *.ve* file, the filename could be any initial set of alphanumerics, followed by a period, ve and some other final extension. P-Graph files usually have .vec, .ved, .vem, or .ves as a file extension. We will not continue repeat all these file names ad infinitum, instead we will use .ve*. We will cover filenames more extensively later.

2.2e Defaults

Whenever you see the word default, it means that if you press **Enter**/**Return**, that this option will automatically be engaged.

2.3 DATA ENTRY AND CONVERSION: Standard Genealogical Format

TUTORIAL STEP ONE: PREPARING A KINSHIP DATA SET

This chapter is part of the Tutorial and is intended primarily as a walk-through of the EGO2CPL program. It gives a brief overview of the program, covering topics such as formats, menus, and errors. Topics which need more than a simple explanation will be covered in greater depth in later chapters.

2.3a Data files

A standard genealogical data file has the following minimal information: personal id numbers, characteristics and personal names, gender, spouse, father and mother. The data file containing the genealogical trees from "THE HERITAGE OF THE CONQUISTADORS" by Samuel Z. Stone, as coded by Fabio J. Flouty, is set up as follows: The first line is the format number of the data set, followed by the title of the data set. The second line is a Fortran FORMAT statement; this will be explained in section 4.3d. Individual entries are next, with data laid out as follows: (Please note that these column assignments may vary with each data set)

```
columns  1-5  are personal id numbers
column     6  is  gender
columns  7-41 are personal names
column  42 is personal characteristics (P-President/S-Senator)
```

```
 columns 43-48 are spouse's id numbers
 columns 49-54 are father's id numbers
 columns 55-61 are mother's id numbers
 columns 61+   could contain clan, land rights, political
 office, religion, or any other attributes you have arranged in
 columns or any pertinent marginal notes which you care to include.

4   Conquistadores
  (i5, a1, 1x, a20, 14x, a1, 3i6)
  1 M cristobal de alfaro                2     0     0
  2 F catalina gutierrez xaramillo       1     0     0
  3 F maria de alfaro                    4     1     2
  4 M cristobal de chaves                3     0     0
  5 F ana chaves de alfaro               6     4     3
        ...
985 M gabr. cardenal Caldera S-16       S0     868   869
986 F ma. josefa alvarado guevara        987   53    54
987 M          ?                         986   0     0
990 F esmeralda martinez urtecho         991   854   853
991 M rodolfo cardenal arguello          990   993   0
992 M parent of 848 and 849              0     0     0     fict.
993 M parent of 991 and 861 and 868      0     0     0     fict.
994 M parent of 866 and 867              0     0     0     fict.
995 M parent of 896,953,954,955,957      0     0     0     fict.
996 M parent of 958 and 959 and 960      0     0     0     fict.
END
```

All files should begin with an ego number; then there are four regularly used choices of order for: Sex, Name, Spouse Number, Father Number, and Mother Number. The word number is usually abbreviated as "No." Ego numbers must be unique for individuals, but do not have to be consecutive and can be in any order. Gender MUST be in capital letters, F or M. The numbers for Fa, Mo, and Sp correspond to their unique ego numbers. You will notice that proxy parents have been created, such as person number 992. It is necessary to assign numbers to at least one parent of people in the data set when parents aren't known and a child or several children are in more than one couple, or the graph drawing algorithm cannot show these couples as having any connection, since there will be no older generation to tie them to. It is also necessary to assign a number to a parent of people in the data set when parents aren't known and you want a separate line to be drawn for these people in the graph.

We strongly recommend that these numbers be several hundreds or thousands larger than the rest of the numbers so they are easily differentiated. If you have people in the data set who are known to have never married (bachelors or spinsters), we designate their spouse and spouse's parents with the number -1; for people who aren't yet married we use the number -2 in the same manner. Similarly, for those who have married someone unknown or whose spouse's parents are unknown we may use the number 0 or, if it is important to show the unknown spouse in the graph, we may create a proxy spouse and/or spouse's parent to stand for a real but unknown person. We often give proxies a large ego number in a special range reserved for proxy spouses or parents. Keep careful track of numbers you use for proxy parents. There is error checking in the Ego2Cpl routine that will check for multiple uses with distinct children of the same proxy parent number, which is helpful if you have only created one proxy parent for each of those who are multiply married with no siblings. There is also a procedure described below for the demo.txt that helps keep track of numbers.

2.3b Demo.txt

For your convenience we have included a demonstration data set named Demo.txt to illustrate how to best number individuals and proxy parents. If female proxy parents are given even numbers and male proxy parents are given odd numbers, it makes error checking simpler. For small data sets, you may simply type in names, clan etc. into the demo.txt file and then rename the file.

2.3c Primary Data Formats

The 4 primary formats are:

1. EgoN, name, sex, FaNo, MoNo, SpoNo
2. EgoN, sex, name, FaNo, MoNo, SpoNo
4. EgoN, sex, name, SpoN, FaNo, MoNo
5. EgoN, sex, SpoN, FaNo, MoNo, name

(The *.Dem file format is explained later. There are also a several options (#3, #6, #7) that contain special features or programming for specific data sets. You will want to check with the author before attempting to use these options on your data)

The following is an example of format 2:

2 Keanae - Maui
(I3,1X,A1,1x,a9,i3,i4,i4)

```
 1 M Billy     29   0   5
 2 F Vickie    29   0  22
 3 F Akuna     29   0   4
 4 M dead       0   0   3
 5 F Pearl     30   0   1
 5 F Pearl     30   0   6
 6 M Henry      0   0   5
...
23 M Abraham   33   0   0
24 M Harry     33   0   0
25 M Alfred    33   0   0
26 M Phillip   25   0   0
27 M #1        23   0  31
28 F #2         1   5  32
29 M dead       0   0   0
30 M Pearl's Fa 0   0   0 (Proxy)
31 F #1         1   5  27
32 M #2        23   0  28
33 M            0   0   0
34 F Roslyn    33   0   0  end
```

Ego numbers must be unique for individuals, but do not have to be consecutive and can be in any order. The numbers for Fa, Mo, and Sp correspond to their unique ego numbers. Thus, Pearl is Billy's wife, and her father is unknown (name not given). Pearl is listed twice, as she has two husbands, which necessitated creating a proxy father. In this data set, there was no need to give her mother a number. The common relations of Billy, Vickie and Akuna to their father, however, are the only indicators that they are siblings.

The first line has the format type (in this case: 2) and title. If no format number is given in the beginning of the first line of the file, EGO2CPL will request the number of the file format when you run this program. The second line is the FORMAT statement.

2.3d Fortran FORMAT Statements

The second line contains a Fortran FORMAT statement:

(I2,1X,A1,1x,a9,i3,i4,i4)

In Fortran: I2 = read a two column integer (the ego number)
1X = skip one column (2X=two columns, etc.)
A1 = read one letter of the alphabet (gender F or M)
A number which occurs in an alphanumeric field
is not a problem. The inverse is not true.
(A2=two columns, A15=fifteen columns, etc.)

Numbers may be placed anywhere so long as they fall strictly within the columns specified by the FORMAT statement. The way to correctly END a list is to type "END" after the data or to cause an error in reading a numeric variable. When the program reaches END or a line in error there is a letter instead of a number, and the program stops. Errors may occur if a letter of the alphabet mistakenly occurs in a numeric field. Where there is such an error in the placement of the columns, however, program will treat it as an END statement in the list, and the program will have read the data only to this point. For this reason, the program always shows the data fields of the last line read.

2.4 DATA CONVERSION EGO2CPL program Ego2Cpl.for

(c) 1997 Douglas R. White, UC Irvine Ver. 2.6 Again, this chapter is part of the Tutorial and is intended primarily as a walk-through the program. It gives a brief overview of the program, covering topics such as formats, menus, and errors. Topics that need more than a simple explanation will be covered in greater depth in the advanced sections.

2.4a How to Run the Program

This program reads an input file that contains kinship data and converts the data to a format that the P-Graph can read. It also contains error checking routines that will detect problems with conflicting codes such as people coded male in one place and female in another, people who have married both a male and a female, people who are their own parent, etc.

Prior to starting Ego2cpl, it is a good idea to use the DOS editor to look at the last line in the data set and to write down the number of lines in the data set and to also copy down or print the last line. Then when the program tells you how which is the END or invalid input line in your data set, you are able to tell if there is an error in the file which is stopping the program prematurely.

In order to start the program type:

"EGO2CPL"

The program asks for:

filename?

At this point you type in the name of the file containing your data, and press **Enter/Return**

Then, if there is a number given at the beginning of the file, the program will tell you:

Your data is in format # __ (and it will give the number) If the format number is not present, it will still show the screen below but will then request the format number. So, the screen will list the formats it can process:

FORMATS FOR GENERAL USE

1. *EgoN, name, sex, FaNo, MoNo, SpoN*
2. *EgoN, sex, name, FaNo, MoNo, SpoN*
3. *EgoN, name, sex, SpoN, FaNo, MoNo, DECades*
4. *EgoN, sex, name, SpoN, FaNo, MoNo*
5. *EgoN, sex, SpoN, FaNo, MoNo, name*

CASE SPECIFIC FORMATS FOR SPECIALIZED USE (see D. White)

6. *EgoN, sex, SpoN, FaNo, MoNo, House of Birth, House of Post-Marital Res., name (e.g., Feistritz data)*
7. *EgoN, sex, SpoN, FaNo, MoNo, Pre-M Res, Inher, Decade of Birth, Post-Marital Res.,Purchase, Name (e.g., Feistritz data)*
8. *EgoN, sex, SpoN, FaNo, MoNo, Name, Decade of Birth, Post-Marital Res1, Post-Marital Res2(e.g., Trukese data)*

Then, if no format number is given in the beginning of the first line of the file, EGO2CPL will request the number of the file format The screen will then respond:

Please enter the appropriate format number and press Enter/Return

In either case, the screen will then respond:

Hopefully you remembered to create "proxy parents" for those individuals who have multiple spouses and no siblings

press **Enter**/**Return**

press **Enter**/**Return**

2.4b Processing options

Next the screen will give you three options:

Choose an option: *and then press **Enter/Return***

D)ata-no eliminations (default) *S)trip off nodes with no parents or no children ("strings")* *and C)ut off all layers of nodes with no multiple links ("core")*

You enter the first letter of the option you wish to use, that is: "D", "S", or "C", and press **Enter**/**Return**

D)ata is the first option, and the one most commonly used. It preserves all the detail of the original data set and eliminates nothing.

S)trings, the second option, will eliminate couples with only one link to the rest of the data set

C)ore is the last option, and will eliminate all couples with less than two links to the rest of the data set.

The screen will then say:

number of records to list? (default:0)

You should then enter a number between 1 and 12, which will be the number of records you wish to see on the screen. The program will then display the requisite number of records from the beginning of the data set. This gives you the opportunity to see what the program is seeing. Or, you may use the default, and see no records.

2.4c Types of Output

Next, the program will give you three options:

1. output p-xxx. pgraph files*
2. output p-xxx.ged files
3. *both types*
?

Number 1 is the usual choice and the default. press **Enter**/**Return**

GED files go into a Genealogical Event Database. These files are most often used by commercial genealogy programs and by members of the Church of Latter-Day Saints, but today are sometimes used by anthropologists.

2.4d End of File Check

The screen will display a copy of the Fortran FORMAT statement at the beginning of the file, the first however many records you asked it to list and:

END or invalid input line = , followed by a number. The number given is the number of the last line the program could parse in the data file, and not the ego number on that line. You can then check your notes to see if this is really the end of the file, as part of the text of this line will be displayed. Be aware that the information is not necessarily displayed in the correct order, so check for content and not order.

The screen will then respond:

press Enter/Return

press **Enter**/**Return**

and then again: *press Enter/Return*

press **Enter**/**Return**

The next thing that will happen is that unless the data set is extremely small, the marriage statistics will quickly scroll past. They will be saved in the <u>errors</u> file, and you can print them later. Also see "2.4h" Marriage statistics."

The screen will then respond:

press Enter/Return press **Enter**/**Return**

2.4e Error Checking-Proxy Parents/Parents with Multiple Children

Then the program responds with:

Error checking routine: This is helpful in eliminating duplication of "proxy parents" for those individuals who have multiple spouses and no siblings.

Do you want error checking (yes/no)? (default=no)?

If no error checking is desired:

press **Enter**/**Return**

If, on the other hand, you wish to check for duplicate numbers,

type <u>y</u> and press **Enter**/**Return**

If you press y, the screen responds:

Lowest number used for "proxy parents" for those individuals who have multiple spouses and no siblings.

You must enter a number here, for the program to continue. This is an error checking routine. If you have people in your data set who have married more than once, have no siblings and you have created a special series of parent numbers for these parents, then you are able to check and see if you have inadvertently used the same numbers for different individuals. Type in the number of the first proxy parent in the data set. If you haven't created any proxy parents, but you wish a list of parents with more than 1 child, as an error check, type in the 1st ego number of the proxy parent in the data set, and

press **Enter**/**Return**

The screen will then respond:

Highest number used for "proxy parents" for those individuals who have multiple spouses and no siblings?

Again, you must enter a number here, for the program to continue. Type in the largest number you used for a proxy parent, and

press **Enter**/**Return**

The screen will then respond with a list of all parents who have more than 1 child, and you then

press **Enter**/**Return**

2.4f Errors File

At this point, the output of the program begins to scroll by this display culminates in a screen which lists the files created by ego2cpl and gives a message about the errors file, as follows:

p-xxx.do is the graphic labels file p-xxx.ve* is vector output errors file contains marriage statistics and a list of errors your processing is complete! Stop - Program terminated*

To be able to see a copy of the errors file, type:

"errors | more" and press **Enter**/**Return**

This will enable you to see the file, one page at a time. If you wish to print the file, type:

print errors and **Enter**/**Return**.

The errors file will contain a list of inconsistencies in the data, followed by the marriage statistics.

2.4g Inconsistencies in the data

If you get messages that say, for example, "number 174, a male married a male and a female" it probably means there is an error to correct and you should check the text file for errors. It is easiest to use the DOS editor, go to the search command, type in "174" to find the first occurrence and then press the search key to find all other occurrences of the number 174.

Also, when you have an error which indicates a reciprocal relationship is not reported by both individuals, you will get messages of this type:

No Matching Husband report for Husband 58 and Wife 59 or: *No Matching Hu report for Fa 58 and Mo 59*

2.4h Marriage Statistics

After a report of the inconsistencies in the data, the marriage statistics will appear. Since the original impetus for this work is alliance theory, we take particular notice of alliances with siblings and they are noted at the end of the report. Screen output should look something like this:

```
99  males, 103 females
42  coupled males 55 coupled females
98  nodes
```

```
1203 highest number used for any individual
Statistics for males
 57  males have no      alliance  57*0= 0
 38  males each have  1 alliance  38*1=38
  4  males each have  2 alliances  4*2= 8
46  alliances by 42 males =1.10 alliances per person for males

Statistics for females
 47  females have no alliance       47*0= 0
 53  females each have  1 alliance  53*1=53
  2  females each have  2 alliances  2*2= 4
  1  female       has   3 alliances  1*3= 3
60  alliances by 56 females =  1.07 alliances per person for  females
 38 has leviratic husbands
 47 has sororal wives
```

The program creates files P-xxx.ved, P-xxx.doc, P-xxx.nad, the file called "errors" and sometimes p-xxx-1.crd and p-xxx-2.crd file (certain formats have an option for decades, which makes it possible for Ego2Cpl to do the generational ordering in a strict temporal fashion). Every time the program is run, it will create these files and overwrite the files created in the previous run. This makes it imperative that when you have a copy of these files that you want to keep, you must rename them. The easiest way to do this is to type
"ren p-xxx*.* p-" and then type 3 letters (e.g., "ABC" of your choosing which will identify the data set for you, continuing with
"*.*" and only then press **Enter**/**Return**.

It is important that you don't forget the initial star/asterisk else the .crd (coordinate) file will not be renamed, and P-Graph won't run properly without it. Of course, this doesn't rename the errors file; some people rename it in order to keep the marriage statistics. Others prefer to simply correct the errors, cut the marriage statistics to a clipboard and insert them in another file and always just overwrite the errors file.

Let us say you renamed your p-xxx*.* files p-yrs*.* and that you had chosen the D)ata file output option. The files p-YRS.VED p-YRS.DOD p-YRS.NAD contain the makings of a p-graph in the .VED file and all of the labels needed to go with it in the .DOD and .NAD files (the DOD file is for the computer, and the NAD file is for you as a reference file as to how the p-graph was constructed).

2.4i Size or Memory Limits of Ego2Cpl

The maximal number permitted is 9999.

3.0 THE P-GRAPH ENVIRONMENT
TUTORIAL STEP TWO:
Elementary Data set Processing

===

This section -- 3.0, 3.1, and 3.2 -- is still part of the Tutorial and is intended primarily as a walk-through the program. It gives a brief overview of the program, covering topics such as menus, editing options, and how to print a slide. The nature of PGRAPH data sets, analyses, and outputs will be covered in depth in the next chapter.

3.0 PGRAPH Menus and Help

PGRAPH is menu-driven. You choose what you want to do by selecting items from a menu. Menus may be nested, so that to get to certain choices you may have to step through many menus along the way.

3.0a Aborting the Program

To exit the program, at any time: push the **Control** key and the **C** key together or else the **Control-Break** keys together.

3.0b Starting the Program

To start PGRAPH, you must type the following:

"PGRAPH" and press **Enter**/**Return**

3.1a Basic Tutorial Using Most Simple Options

Other basic options are covered in section 3.2.
The first screen reads:

Marriage Networks 5/97 (c) 1991-97 Douglas R. White.

Freeware *http://creativecommons.org/licenses/by-nc-nd/3.0/*

You are granted a limited license to use this working version of Parente Suite which includes PGraph, Ego2Cpl, and Par-Calc
Douglas R. White, Social Science Plaza,
University of California, Irvine, Irvine, CA 92697
Please report debug error messages to drwhite@uci.edu
Inspired by the work of Lévi-Strauss (1949) and Weil (1949), the prototype of the p-graph (having parental couples as nodes rather than individuals) was developed in 1960 by Guilbaud (1970; see Cuisenier 1962, Bertin 1967). Extending this representation from marriage system models to kinship network applications, the program analyzes p-graph representations (White and Jorion 1992 Current Anthropology 33: 454-63) of kin connections among marriages 1 to N, linked through male and female genealogical ties, using vector genealogy as developed in Jorion and Lally (1983, unpublished ms). See SPONSORS/PRECURSORS for further bibliography.
Thanks to Patricia for interface suggestions and major revisions of the user manual.
:
Special thanks to Clemens Heller, Maison des Sciences de l`Homme and the French Ministry of Research and Technology for support of the Discrete Structures in the Social Science working groups that made the programming and larger project possible, and to Paul Jorion, Michael Houseman, Alain Degenne, Thomas Schweizer, Vincent Duquenne, François Héran and Lin Freeman for colleagial support.

3.1b Graphic Defaults

If you have a matrilineal descent group, the marked sex will be female, and therefore white. If this doesn't suit you, you may type M and then you will go to a submenu (see 9.2a). This is true of all four options, but for most people and purposes, press Enter/Return works very well.

GRAPHIC defaults:

M)arked Sex: solid white lines

U)nmarked Sex: dotted red lines

C)o-parents : solid gray

B)ackground : solid black

To change, press letter indicated by)
To use the default settings press Enter/Return
press **Enter**/**Return**
the screen will scroll up, and will add at the bottom
MEMORY defaults:

G)eneration depth (default = 20)

E)xtra memory for non-implicit label (default: 1)

To change, press letter indicated by)
To use these settings press Enter/Return
press **Enter**/**Return** (or see sections 9.2b and 9.2c):

3.1c Data Processing Options for New Slides

the screen will scroll up, and will take you to the data processing screen. The screen reads:

Choose an option to: create graphs FROM NEW DATA in one of two initial formats

V)ector format (CHOOSE THIS FOR NEW DATA IF IN DOUBT)

D)emographic format (individuals and birth/death/marriage dates)
Or:
E)xamine, E)dit, or analyze E)xisting kinship graphs (default)
R)ead explanations about DATA SETS/BIBLIOGR/ANALYSES/SOFTWARE/FORMATS/ SPONSORS/PROJECTS/PRECURSORS/or HELP and return to this menu

Optional: The R option will give you many choices among the documentation files which are included in the Parente Suite package. This option catalogs a collection of data sets, the bibliography of these data sets, previously analyzed slides, sponsors of the program, and related projects. If you choose this option, you may RETURN to the menu above by pressing "-."
optional: press R and **Enter**/**Return**
optional: press - and **Enter**/**Return** to return to the menu above.
Since we are this point assuming you are processing your first set type V and press **Enter**/**Return**

3.1d Data Format Request for New Slides

The format that you used to process the data in Ego2cpl is also the format you should use here. The final letter of the .Ve* file indicates how the data was originally processed and how it needs to be processed by P-graph.

The screen will scroll up, and will add at the bottom this menu:

```
Data format:
  D)ata, complete format (default)
  S)tripped/String format (uses P-_?_-1.CRS P-_?_-1.VES file)
  C)ore                   (uses P-_?_-1.CRC P-_?_-1.VEC file)
  B)lood kinship network  (uses P-_?_-1.CRC P-_?_-1.VEC file)
  M)odel data             (uses P-_?_-1.CRM P-_?_-1.VEM file

  E)XPRESS run with complete data format: create slides non-stop

  K)onvert: *.VE* to DEMographic format:
            p)revious menu Choice?
```

If you used the D)ata option, press **Enter/Enter** If you used the S)tripped/String option, press S If you used the C)ut or Core option, press C If you used the B)lood Core option, press B (see section 6.1) Other options will be covered under their respective headings.

The screen will then scroll up and you will be at the data set menu.

3.1e Choosing a Data set

Once you have chosen a data set format, all the available data sets that are available will be numbered alphabetically by their three-letter alphanumerics that identify particular data sets. For example, the screen will look like something like this:

```
  1 ANU 2 BEN 3 *** 4 *** 5 **** 6 MENU
 NUMBER of the data set?      and  press Enter/Return:
```

Enter a number from the set **1 2 3 4 5 6** and press **Enter/Return**

press **Enter/Return**

press **Enter/Return**

The number of times this message occurs varies in accordance with the size of the file, but eventually you will encounter a screens which say:

```
LABELS READ FROM p-***.DO* file
Minimum Difference to print in Generation levels (default =2)?
If in doubt, press Enter/Return
```

press **Enter/Return**

3.1f Descent Order Rule for New Slides

The screen will scroll up, and will add at the bottom:

```
the options affecting how couples are clustered when
the PARENTAL ORDER GRAPH is drawn are:
1.  Patrilines
```

2. *Matrilines*
3. *Bilateral*
4. *Optimize 1,2, or 3 according to the density of ties*

What is your choice?
type the number of your choice only
Author's Note: at this time the bilateral option is disabled;
we are working on getting it back in operation.

3.1g Generational Ranking Option --

For a slower and more sculptured ranking press N)o To rank generations as simply as possible (default)
press Enter/Return

If in doubt, press Enter/Return
...computing...
press **Enter**/**Enter**

3.1h Normal Descent Order for New Slides

Normal Descent will be used unless you specify
D)escendants of unmarked sex are ordered next to their respective descent groups (D may be useful only after running the PAR-MARR routine)

If in doubt, press Enter/Return
press **Enter**/**Enter**
the screen will scroll up, and will add at the bottom:

3.1i Vertical Descent Line Continuity

Realign vertical descent lines so that if there are
no sons in the patriline a daughter appears next,
no daughters in the matriline a son appears next
use Y)es to realign or the default
to keep your descent lines separate?
If at all in doubt press Enter/Return
press **Enter**/**Enter**

3.1j Backup Coordinate Files

Coordinates of every run are saved in an auto-save file, along with a back-up copy, just is case you make a mistake. The next screen lists the name of the auto-saved files:
*A backup copy of P-***o.crd is named P-***-p.crd* or *A backup copy of P-***-p.crd is named P-***-O.crd* or *A backup copy of P-***- m.crd is named P-***-O.crd*
where "***" is the mnemonic that you have given for your data, and the last letter before the extension (0, p, m) depends on which of the descent order rule you have chosen. You needn't worry about these files; they primarily serve as backups for the first slide that is produced, and they are used if you choose the option to R)epeat previous analysis for data set in section 3.1d.
press **Enter**/**Enter** press **Enter**/**Enter**

3.1k Endogamy

The screen will scroll up and give figures for the number of individuals in the data set whose parents were not known.

Percentage of spouse's parents unknown or exogamous: ___ or __ %
[this excludes individuals coded as unmarried with -1
press Enter/Return
press **Enter**/**Enter** TWICE
You should now see your P-graph on the graphics screen except when the kinship and marriage graph has disconnected parts. You will then be asked the following questions.

3.1l Disconnected Part Options for New Graphs

Please make a choice between:
A)dd A)ll remaining (including unconnected) links to the current graph
S)kip this segment and go to the next separate component
T)his segment computed for separate slide, then next segment?

This menu tells you that PGRAPH has connected the largest component of the network into a graph, and is willing to stop and make a slide of this component if you choose option **T**, and then make a separate slide of the next component, until you want to stop this process of making separate slides for separate components.

The components chosen by PGRAPH in this manner are successively less deep generationally (often but not necessary having fewer couples). If you want to include all the remaining links in the next slide, at any time in this process, choose option **A**. Or, if you want to skip a segment, choose **S**. Hence

Press a letter from the set **A S T** and press **Enter**/**Enter**

3.1m Draw by Generations

The next screen reads:

Program requires CGA, EGA, or VGA CARD
use "S" to step through colorings? reply: "S"

If you want to watch the generations in the graph being drawn one layer at a time,
type S and press **Enter**/**Enter**
and keep doing this, each time the keys are pressed, another generation is drawn
press **Enter**/**Enter** TWICE, and the entire graph appears immediately
You should now see your P-graph on the graphics screen.
Go to section 4 to learn how to edit your graph.

3.2 Viewing or Editing a Slide that Exists

3.2a Where Do I Go From Here?

Remember the screen from 3.1c? The screen reads:

Choose an option to: create graphs FROM NEW DATA in one of two initial formats

V)ector format (CHOOSE THIS FOR NEW DATA IF IN DOUBT)

D)emographic format (individuals and birth/death/marriage dates)
Or:
E)xamine, E)dit, or analyze E)xisting kinship graphs (default)
R)ead explanations about DATA SETS/BIBLIOGR/ANALYSES/SOFTWARE/FORMATS/ SPONSORS/PROJECTS/PRECURSORS/or HELP and return to this menu

We are about to go through the other options, one at a time

3.2b Vector format - Strings and Core Options for New Graphs

The options and steps are essentially the same as given above in section 3.1 for complete data sets. The resulting graphs are smaller as a result of cutting off nodes not connected to more than 1 node in the case of the Stripped or String option and not connected to more than 2 nodes in the case of Cut or Core data sets

3.2c Demographic format - covered under Advanced Options

3.2d Examine, Edit or Analyze Existing Kinship Graphs

type E at the screen from 3.1c and you will get this menu:

1. *Return to previous menu to process new .VE* or DEMographic file*
2. *Data, Complete Network (.CRD & VED files already processed)*
3. *Stripped Network Data (.CRS & VES files already processed)*
4. *Cut Network Data (.CRC & VEC files already processed)*
5. *Network Blood Cores (.CRB & VEB files already processed)*
6. *Kinship network models (.CRS .VEM files already processed)*
7. *Quit*

0. Read DATA SETS/BIBLIOGR/ANALYSES/SOFTWARE/FORMATS/SPONSORS/PROJECTS/ PRECURSORS/or HELP and return to this menu

If you choose Option #1 you will go back one menu.

If you choose Option #2,3, or 4 go to step 3.1e, follow the instructions there and the program will read the slides you have already created and show them to you as you request them by number.

OPTION #5 Network blood Cores is covered in the Advanced section.

OPTION #6 Kinship Network Models is covered in the Advanced section.

OPTION #7 will terminate the program

OPTION #0 will give you many choices among the documentation files which are included in the Parente Suite package. This option catalogs a collection of data sets, the bibliography of these data sets, previously analyzed slides, sponsors of the program, and related projects.

4.0 KINSHIP GRAPHS WITH PGRAPH

==

4.1 Capabilities: P-graphs

The primary purpose of PGRAPH is to create graphic slides of complete genealogical networks or of their separate connected components; to allow you to edit the positions of couples in the slide, save the slide as a set of graphic coordinates; and to retrieve and display previous slides. The PRIMARY GRAPH used for this purpose is the P-graph, with two kinds of vertical lines representing males and females respectively: the tops of the lines meet as sibling sets produced at a parental node, and the bottoms of lines may themselves intersect as couples or parental nodes.

4.2 Basic Editing of a Slide: PGRAPH Menu Choices

PGRAPH Menus, as we have seen above (3.1), give the user a number of options available by pressing letters or numbers from the keyboard. If only a single number of letter is required, the program responds immediately by executing the option. If a number might require one or two digits, the option is executed only after you enter the number and press the RETURN key. After the introductory menus (section 3.1 and 3.2) come the slide menu options discussed below in section 4.5.

4.3 The PGRAPH Screen

After the introductory choices are made to select an existing set of kinship data (see sections 3.1/2 for selecting data), the main PGRAPH menu switches to high-resolution (VGA) graphics mode.

4.4 PGRAPH Ordering of Generations

PGRAPH orders generations of couples, not individuals, so that an individual's marriages may belong to more than one generation. There is no unique generational assignment of individuals. For kinship that conforms to biological principles, however, there is a unique minimal assignment of couples to generations such that a parental couple (from which the individual's birth is traced) always precedes their child's marriage.

PGRAPH generations are assigned to minimize the generational distance between ancestors and descendants. The number of marriages in each generation depends on how the data were collected, such as number of founders, and on demographics. The lowest generation may have few marriages and many unmarried children.

4.5 PGRAPH Slide menu Commands and Subcommands

Letter or number options for the slide menu are typically indicated on the STATUS LINE at the bottom of the screen. Options are usually activated by pressing single keys indicated by the letters that are preceded by a closed parenthesis on the status line.

4.5a Slide menu Options

C)ols E)ditRows D)raw V)iew S)lide P)lace A)dvanced -) Q)uit
SUBMENUS under these choices are as follows:

4.5b Column Editing

C)ols:
A)ddLeftCol M)oveCol T)rans2col U)nilines S)ides O)vr C)los Q)
(note: there are other options listed that are nonoperational)

A) blanks leftmost column and moves others 1 column to right.
M) moves all elements of a column to another column location.
T) transposes two columns by interchange of locations.
U) reorders graph by descent group of marked gender
S) draws unilineal sides assuming you have segregated an initial cluster at the left. When you give two column fields, such as m and n, the 1-m columns are the left seed, and the m-n columns are the right seed. The algorithm then iterates (as many times as you choose) by moving new segments (beyond column n) to the left or right side, opposite where they have marriage links (in the unmarked sex). Section 10.4 discusses statistical tests for sidedness. Use of the sides option creates a COLORING of nodes into two colors, differentiating the lines emanating form both studies of the structure.
O) checks and moves apart points that overlay one another.
C) closes up empty columns by moving columns together.

4.5c Row Editing

E)dit:

Editing of rows consists of changing column positions of elements keeping the row constant. Four lines of input are requested: the number of the row whose elements may be relocated, the first and last elements of a set of elements to be moved, and u starting column position to which the set of elements is moved. Elements in the set will be moved to this and succeeding positions depending on the number of elements moved. To move only one element the 1st and Last elements must be the same. For example, the response 5, 6, 5 as responses to the following questions

Row to change:
Elements 1 2 3 4 5 6
Col. Nos. 3 0 0 6 7 8
1st No to Move: 5
Last No to Move: 6
MoveSet to Col: 4 will produce a new row of elements 1-4 in cols 3 4 5 6. If the MoveSet to Col: response had been 8 the new alignment would be
Elements 1 2 3 4 5 6 7
Col. Nos. 3 0 0 6 0 8 9 The editor does not identify elements by name but only by their position as elements at regular intervals on the graph, slots numbered from 1 to N. Further, if you responded with 6 as the target column above (at an index number lower than the first element

to be moved), a column already occupied the editor would oblige by moving the existing element at location 6 two columns to the right to make room for the set of two elements to be moved.

4.5d Drawing Options

The menu for drawing options is:

```
D)raw:
  F)ull-then s?ingle generation   P)oints only   R)females W)males
  S)witch   C)olor
```

These are the choices:

F) full graph (use of s? option "s" steps through graphic
adding generation by generation
P) lines removed, only points remain
R) red lines: only females if patrilineal, males if matrilineal
W) white lines: males if patrilineal, females if matrilineal
S) switches the colors for the two genders
C) provides a menu for adjusting color options:

Choice C) has a submenu for change of colors:

```
M)arked sex U)nmarked C)o-par B)kgrnd P)tn.colors+1 ?)cuts I)n
```

M) marked sex, e.g., change "red" male LINES to blue
U) unmarked sex, e.g., change "white" female LINES to violet
C) color of links between the two sets of parents whose
children have married (advanced option: section 9.2)
B) background, e.g., black to white
P) iterate colorings of POINTS if activated
?) cut 1, 1 to:
I) initialize colors

When you specify color choices under options in the set **M U B C**, the color selection menu is as follows (see 9.2a):

```
 1 = dk blue
 2 =    green
 3 =    cyan
 4 = dp red
 5 =    purple
 6 =    brown
 7 = offwhite
 8 =    gray
 9 = lt blue
10 = lt green
11 = lt cyan
12 = lt red
13 = lt purple
14 =    yellow
15 = bt white
16 =    black
```

4.5e Viewing Options

V)iew:
B)order L)abels

B) add border, delete menu (for graphics screen-dump printing)
L) add labels for couples with the following options
SUBMENU: N) numeric labels for couples
  I) implicit numeric couples (vector index)
  C) column number labels
  M) male ID labels
  F) female ID labels
  L) text labels -- you will be asked to give two 2- digit numbers between 1 and 15, the first for the starting column and the next for the final column from which to extract a maximum of 15-column date from the p-*.DO* file for your data set.

4.5f Slide Options: save, export

S)lide:
N)ew slide number S)ave slide save to U)CINET P)ostscript file for DOT

N) new slide number
S) save slide operates directly to save an edited network to a .CRD file. The name of the .CRD file depends on the previous option used to generate the current slide:

Option	S)AVE .CRD File
------	-----------
Normal: Patrilineal	P-***-P.CRD
Normal: Matrilineal	P-***-M.CRD
Repeat:	P-***-O.CRD
Slide: e.g., 2	P-***-2.CRD

U) save to UCINET in DL linked list format with couples as
the unit; parental links between couples; two relations,
one for parents of males, one for parents of females.
[we are working on an individual-to-individual output].
P) save to a Postscript file for printing with the DOT
graphics program

4.5g Placing a Slide

P) allows you to move to a new slide or create a new slide from the existing screen. You are asked for a single digit slide number, and given one of two responses:
-- slide does not exist: new slide can be saved for this number
-- slide exists: can retrieve from *.crd get new slide: Y/N?

4.5h Quitting a Slide to go to the next

Slide viewing begins with slide 1. To quit a slide and go to the next,
Type "Q"

4.5i Quitting Slides to return to main menu

After the last slide, "Q" will take you to the file selection menu.
To return to the file selection menu at any time.
Type "-"

4.6 Quitting the program

Type, in the graphics menu, for -)Q)uit:
-) returns to format and file selection
Q) to next slide; if none, to file selection

Once you return to the section of the program where files are selected by number, choose the last number to get the main menu, and then choose option 6 to quit the program (see section 3.2e).

4.7 Size or Memory Limits

Arrays are dynamically dimensioned and so depend on memory limits, but there is a limit (which can be changed by the programmer if necessary) on the number of descent groups and permissible colorings.

5.0 PRINTING YOUR P-graphs

TUTORIAL STEP FIVE: PRINTING A P-graph SLIDE

==

5.1 A Note on Vector versus Raster (Pixel) Images

The image you see on the screen in PGRAPH consists of a graph or network (graph in which the arcs are valued) of lines and points. To translate these graphics to a printed copy, vector graphics are usually incomparably sharper than raster or pixed-based graphics. Vector graphics are produced by a set of instructions such as the Hewlett-Packard Graphics Language, called HPGL in version 1 and HP-GL/2 in version 2. Upwardly compatible and virtually indistinguishable for our purposes, HPGL files (often called *.PLT files generically, or *.HPG if special version 2 commands are used) allow pen assignments to take different thicknesses and colors.

HPGL languages were originally designed to provide instructions to plotters. To send HPGL instructions to a printer, a plotter simulation program is used. A .PLT file can be used with color printers provided you have a HP-GL/2 plotter simulation program such as described in section 5.5. If your printer is monochrome, you may use HPGL instructions and the LP program from Insight Corporation described in 5.6 (Insight's color version of this program is also described). The HPGL Plotter language works not only with Hewlett-Packard laserjet, deskjet and inkjet printers but with other printers, depending on the capabilities of the plotter simulation software that you use.

Thanks is due to Linton C. Freeman who first introduced me to *.PLT files and the LP program, and wrote the prototype program in BASIC that would translate from PGRAPH formats and coordinate files to *.PLT files.

5.2 PAR-PLOT does P-graph HPGL VectorGraphic Instructions

Once you have run the PGRAPH program and so generated a .CRD-type file with graphic coordinates for making visual slides of the kinship network, you may create vector graph instructions that will reproduce this slide by running the PAR-PLOT program. PAR- PLOT will ask the user for names of data files to use (including an identifier for a specific .CRD file where the coordinates for one of your slides are stored) and other options, and generate a plot file with the extension .PLT. This file will contain the HPGL graphics language instructions needed to reproduce your graph as a vector graphic.

HPGL or .PLT files (including HP-GL/2) can be edited as ordinary ASCII files so as to perfect the graphic image to your liking. These images may have high resolution (up to 600dpi) and can be produced in publication quality. The minimum quality required is 300dpi, since 150dpi vector graphics produce squiggly lines.

5.3 Running PAR-PLO.BAS to create a *.PLT file

To run PAR-PLO or PAR-PLOT (which differ in memory allocation), you must have QBasic.exe in your path, and the PAR-PLO.BAS or PAR- PLOT.BAS programs, plus your PGRAPH .VE* and .CR* files on your directory. A PAR-PLO.BAT or PAR-PLOT.BAT batch file is provided that contains an instruction to look in the /DOS directory for QBasic and to execute a command such as

```
QBasic par-plot.bas
```

You give this command in DOS. QBasic is a BASIC language interpreter widely distributed with other MS-DOS utilities. You will know you are successful if QBasic loads and you see the QBasic Menu line at the top of the screen with the rest of the screen loaded with a basic program.

Press **Alt**-R to run the basic program.

Press **Enter**/**Return** to start to execute the program.

The program asks for the following:

3 LETTER CASE NAME?

You reply with the 3-letter mnemonic for your data. Ex: ABR

M=Model B=Blood C=Core S=String D=Data?

Reply with a single letter from the set **M B C S D**

P M 1 2 ... Slide COORDINATES FILE?

Reply with a single letter from the set **1 2 ... 9 a b P M ...**, corresponding to a number or letter that you used to save a slide while in PGRAPH.

The program will open a .CRD file with this designation. For example, if you chose ABR, D, 2, the program will open the .crD file p-ABR-2.CRD. Next, the program will search for the corresponding .DO* (.DOD) file, as indicated by the following message:

Doc file=p-ABR.DOD

LABELS from .DO file: C)ouples I)ndividuals rtn=other or none?*

If you choose C or I the program will open your .DOD file, in this example p-ABR.DOD, and will retrieve data contained in this file which pertains either to individuals (columns 1-8) or couples (columns 9-forward). You will only be able to use a maximum of 15 columns worth of data that was part of the NAME FIELD when you compiled Ego2Cpl (section 2.4).

If you do not choose option C or I above, the following option will appear:

Label by .VE number H)usb W)ife or .CRD number N)Des M)arr# I)mplicit ?*

Enter H for the husband's identifier number Enter W for the wife's identifier number Enter N for the descent identifier number Enter M for the marriage identifier number Enter I for the implicit couple identifier number

D)raw or O)mit circles ?

Enter D to include large circles to identify points Enter O to omit circles to identify points

A)ll or S)ame D)ifferent group lines?

Enter A to draw all lines Enter S to draw lines within the same descent group Enter D to draw only lines between descent groups

The program will then list the name of the coordinate file, the title, and a header for POINT LABELS, followed by a request for the COLUMN NUMBERS of the .CRD file from which you wish to extract labels (if any), as follows:

p-ABR-2.CRD Genesis Genealogy of Canaan POINT LABELS Couple LABELS from .DOC file from column: ? Couple LABELS from .DOC file to column: ?

You may respond with two numbers, the second exceeding the first by no more than 14, which reaches the total size of the labels field. Normal placement in the .CRD file of data output of Ego2Cpl is as follows:

Cols. 1-4 Husband's Identifier Number Cols. 5-8 Wife's Identifier Number Cols. 9-37 Husband's Text Data (e.g., name) Cols.38-66 Wife's Text Data (e.g., name)

The *.DO* file may be edited to place whatever text data one wants in certain columns so as to retrieve these data as labels.

You will then see a list of the labels that you have selected for points.

The program will then list

Past End: Colors: p-ABR-2.CRD

CHECK DUPLICATES IN EACH LOCATION ?

Enter Y to check whether one or more points are on top of each other, or press **Enter**/**Return** to skip the checking.

HORIZONTAL LABELS ?

Enter Y to have your labels printed horizontally on the page. Enter N or press **Enter**/**Return** otherwise

VERTICAL LABELS ?

Enter Y to have your labels printed vertically on the page. Enter N or press **Enter**/**Return** otherwise

The program will list points and their labels.

Press any key to continue?

press **Enter**/**Return**

and you will return to the QBasic menu page. To exit,

press **ALT**-F followed by X to exit.

You may now open your plot file (in ASCII) in a word processor to inspect the HPGL instructions that have been generated. The file in this example will be named as follows:

```
      / the three letter mnemonic: abr
     / / the slide number asked for: 1
    / // the type of file: d
   / /// the labels option chosen: c p-abr1dc.PLT
```

The *.PLT extension is needed to indicate an HPGL file.

5.4 How to get colors by editing your *.PLT files

Colors are assigned in HPGL and HPGL/2 only by pen number and not by direct assignment within the *.PLT file. Different HPGL plotter simulation programs use different default color assignments and allow you to change the color assignment of pens. Default SPLOT (5.5) and AutoCAD color assignments are, for example:

SPLOT AutoCad 1 black red 2 blue yellow 3 green green 4 red blue 5 yellow purple 6 aqua pink 7 olive black 8 gray gray In addition, 255 color pen assignments can be distinguished, and new colors can be mixed from the three primary colors, depending on your plotter simulation program that actually prints the graphs.

The PAR-PLO(T) programs equipped with PGRAPH make a *.PLT file with three pen types distinguished. Open your file with an ASCII word processor and search for SP. In *.PLT files made with the new version of PAR-PLO.BAS (dated after May 1997) you will find: *SP1;PT.2;PW.4* above text labels: Set Pen to 1, Thickness =.2mm. *SP1;PT.2;PW.4* above solid lines: Set Pen to 1, Thickness =.2mm. *SP4;PT.4;PW.4* above dotted lines: Set Pen to 2, Thickness =.4mm. The colors are set for the following colors in SPLOT: black, black, and red, so as to reproduce the DOS graphics computer screen with solid lines and text in black, and red dotted lines in red. These numeric assignments can be changed, and commands to set pen number (SP#) or width (PT.#, PW.#) can be inserted elsewhere in the *.PLT file to precisely define color effects. Although most HPGL files recognize PT as pen thickness, SPLOT (5.5) recognizes only the command PW, so Par-PLO.BAS inserts both commands, one of which will generated a harmless plotter error message. To eliminate this message, and check for other errors, edit out the incorrect commands.

5.5 Shareware Program SPLOT to simulate the Pen Plotter

SPLOTW32.EXE is "The Plotter Simulator V3.0 for Win32" (Windows 32 bit, but runs as a DOS application in Windows) written by Alexandr Novy, Havlickova 2209, 390 02 Tabor, Czech Republic, anovy@bohem-net.cz. It is a 30-day trial unregistered version available as shareware (not free software) at

http://www.clavis.tel.cz/AlexandrNovy/index.htm

It comes as a zipped file SPLOTW32.zip that contains about 30 pages of documentation, drivers and function libraries, test *.PLT files from the following plotter or programming manuals:

Plotter	# Pens	Language	Varieties	Pen Width
HP7470A	2	HP-GL	Vector .025 mm	PT command
HP7475A	6	HP-GL	Vector .025 mm	PT command
HP7440A Color	8	HP-GL	Vector .025 mm	PT command
Draftmaster	256	HP-GL	Vector	PT command
		HP-GL/2	Vector, Raster	
		DXY-GL	Vector	

Registration forms are included, and single users can register and receive manuals for $65 plus $5 postage outside Europe. This package is well worth the investment. A multisystem single- network installation costs $140 (plus $5), and the UC Irvine lab has been asked to purchase a network installation copy.

The program loads your *.plt file, lets you choose among 16 plotter emulation types, allows you to preview the slide (WYSIWYG), change colors and pen widths, crop or transform the slide, and print the slide on a variety of different printers and page sizes (A0-A4; ANSI, ISO), with stick fonts for English, French and Czech.

I have added upper left and lower right guideline points to the HPGL files made by PAR-PLO(T) to make it possible get ALMOST EXACT matches in placements of a graph when exiting and reentering the program (these programs do not SAVE a new *.PLT file according to your transformations of the graph). This overcomes a slightly troublesome feature that the plot does not automatically fill the entire screen and needs to be resized. The way to resize with the mouse, then, is to click my upper left guide-point on the image and drag the mouse to the lower right guide-point, thereby defining the area to TRIM. Then click TRIM on the SPLOT menu bar and the figure will resize to fill the entire page. Now click ROTATE and click again ROTATE 180 degrees, and the graph will be properly situated for viewing and printing.

When printing, mouse commands follow WORD conventions, and it is best to use the LANDSCAPE option in PRINTER SETUP.

5.6 LP and PrintaPlot pen plotter simulation programs

LP.EXE (Laser Plotter 2.4 for HPGL 1998) and PPP.EXE (1997 $149 academic price) are pen plotter simulation programs produced by the Insight Corporation (1-800-825-4115; 2420 Camino Ramon Suit 205, San Ramon, CA 94583). Their web address is:

http://www.insight.dev.com

LP produces black and white graphics, but allows resolution to be set (150 draft and 300 dpi final but menu driven pen thickness commands vary accordingly) in the program and line thicknesses to be set in the *.PLT file. One good feature of these programs is that the plot automatically fills the entire screen without cropping (see 5.5), but like SPLOT, the graph in PPP must be rotated 180 degrees.

PPP produces color graphics, and is comparable in quality to SPLOT (5.5). The physical difference in the graphs is that the FILL PATTERNS (e.g., for dotted lines) will vary. All three programs, LP, PPP and SPLOT allow setting sheets to be used to adjust parameters for individual graphs.

5.7 Postscript file for printing with DOT graphics

PGRAPH has a S)lide option to save to a P)ostscript file for printing with the DOT graphics program (section 4.5f). Information on DOT graphics is available through the UC Irvine program in social networks (Lin Freeman, Jeff Stern, Doug White, Ben Jester are knowledgeable about DOT graphics.

===

6.0 KINSHIP STATISTICS: THE PAR-CALC PROGRAM

TUTORIAL STEP SIX : Par-Calc - Consanguineal Marriages

(Douglas White and Paul Jorion)

===

6.1 The Order in which to run the programs

For analysis of a new network, data preparation and programs are best executed in the following order:

A. Prepare your data in ascii files (section 2.3: *.TXT files)

B. Run Ego2Cpl.exe (section 2.4) to generate .VED, .DOD, and .NAD files or their equivalents (differing in the final letter).

C. Run PGRAPH (chapter 3) to see and edit the graphic and make slides, or use advanced options (VOL. TWO) to analyze network structure or export your data to UCInet network analysis formats.

D. RUN PAR-PLOT (section 5.1) for VECTOR GRAPHICS, and view the .PLT file with a plotter simulation program (section 5), but don't edit the .PLT commands to perfect your image until you have done step five and begun to study the structure of your data, including advanced options if needed.

E. RUN PAR-CALC (this section) to analyze blood kin marriages and prepare advanced options. Although PAR-CALC can be run directly after Ego2Cpl (step B in the sequence above), it is always best to run the PGRAPH program before PAR-CALC since it provides a number of data checks as well as a visual display of the network.

F. PAR-CALC produces several auxiliary output files besides the main statistical analysis of blood marriages. Some of these are discussed in the advanced options section (10.6-10.7). There is one file, p-xxx.VEB, however, of special importance, since it contains a reduced marriage network consisting of only those marriages where spouses are consanguineals, plus their ancestral couples. Thus, after running PAR-CALC with a file designated p- yyy.ved you will want to use the following command

rename p-xxx.veb p-yyy.veb

where "yyy" is your own mnemonic file designator. You can then return to run PGRAPH (section 3) using the B)lood option (3.1.6), retrieve this as a "new" data set and produce graphs of this structure.

6.2 Capabilities of PAR-CALC: Marriage Statistics

Blood-kinship links between spouses using the parental vectors by which P-graphs are constructed. In vector representation of a kinship structure marriages are numbered sequentially and functions F and G are defined to assign to each the respectively marriage numbers of the wife's and husband's parents. A Canaanite example from White and Jorion (1992) is shown in Table 1, where marriages are numbered from 1 to 20.

TABLE 1 Vector Genealogy Input Data for Canaan Ruling Houses

j	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
G(j)	0	1	1	2	2	2	2	4	6	7	4	4	5	10	13	13	13	13	13	19
F(j)	**0**	**0**	**0**	**0**	**4**	**0**	**4**	**0**	**0**	**0**	**8**	**8**	**10**	**0**	**0**	**0**	**9**	**14**	**14**	**0**
GSib	0	1a	1a	1	2	1	3	1b	0	0	1b	1b	0	0	1c	1c	1c	2d	2d	2
FSib	0	0	0	1	1	2	2	0	0	0	1	2	0	0	0	0	0	1e	2e	0
ID*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

a-e, identical individuals in multiple marriages.

* needs not be the same as the index variable, j.

Ancestors are designated by functions such as GGF for a man's father's mother's parents, where GGF(k)=F(G(G(k))). To determine in marriage k whether husband and wife are related by common descent, there must be equality of some pair of functions, G...(k)=F...(k). For example, since GF(18)=FG(18)=10, couple 18 is related by common descent from couple 10 and the prior kin connection is MBD/FZS.

The heart of the PAR-CALC algorithm, given G and F as input vectors, consists of tracing convergences by vector genealogy:

```
BEGIN trace husbands' genealogical trees to a nonempty vector
  BEGIN trace wives' genealogical trees to a nonempty vector
    save convergences
  continue wives' trees until done (no nonempty vectors)
  END
continue husbands' trees until done (no nonempty vectors)
END
```

Results of applying this algorithm to the Canaan kinship vectors in Table 1 are discussed in section 6.4.

6.3 Running PAR-CALC: Menu Choices and Help

To run PAR-CALC, you need only your p-???.VE* file, such as the one produced by Ego2Cpl. Enter as a DOS command

```
PAR-CALC
```

The first screen reads:

The initial screen gives the following introduction:

```
      PROGRAM PAR-CALC (Parent-Graph Genealogical Calculation)
                                        Version 2.4 May 1996
  computes PRIOR KIN CONNECTIONS among marriages numbered 1 to N
  given as input marriage numbers of husband Js parents in G(J)
              and marriage numbers of wife Js parents in F(J).

    (c) 1992-96 D.R.White and described by White and Jorion 1992
   "Representing and Computing Kinship: A Network Approach" in
    CURRENT ANTHROPOLOGY
   based on the idea of vector genealogy developed in Jorion
   and Lalle 1983, "An Algorithm for the Analysis of Genealogies
   as to prior kin connection between spouses."

      Option for files of type P-___.VEC (will be listed below)
      that have the following data format:
 NMAR    (title optional)             = number of couples in list
  H1P H2P H3P H4P H5P H6P H7P H8P H9P.. numbers of hu pa in list
  W1P W2P W3P W4P W5P W6P W7P W8P W9P.. numbers of wi pa in list
  ... ... ... order of husband in his sibling set (optional)
  ... ... ... order of wife      in her sibling set (optional)
  1    2    3    4 ... consecutive OR relabel numbers (optional)
Vector format D/S/C/B/M/K   ^^^ No=Demographic?
```

Enter a letter from the set *D S C B M K* to specify the type of file. The remaining menu choices are extremely simple:

```
 1 ANU 2 BEN 3 *** 4 *** 5 **** 6 MENU

  Which P-_?_.VED (or VEM) data file to use ?
```

Just type the number of the data set and **Enter**/**Enter**

```
  Choose a Depth of Genealogical Search
  that will identify blood marriages up to
  1=sibling marriages
  2=cousin marriages
  3=2nd cousin marriages
  4=3rd cousin marriages
  5=4th cousin marriages
  6=5th cousin marriages
```

7=6th cousin marriages
<Press Enter or Return for default=7>:
Press a number and/or **Enter**/**Enter**

Is there some marriage path to test (0=no, or give number)?
Press a number and/or **Enter**/**Enter**

After these questions are answered, the program shows a running tally of the number of blood kin marriage relations which are being checked by the program. When the program is done computing, which may take a few minutes at a depth of 7 generations and several hundred marriages (computing time varies linearly with number of marriages or network size), output such as in Table 2 (section 6.4) appears on the screen, but all output is also saved to the files listed in the following section.

6.4 PAR-CALC Output Files

The first four lines of output are as follows:

Genesis Genealogy of Canaan
Depth of Genealogical Search: 7
8(100%) of 8 endogamous couples in blood marriages
8 blood marriages 40% of all marriages

This means that the program was instructed to trace ancestries back 7 generations, that 100% of the couples both of whose parents are contained in the graph married blood kin. Blood kin marriages constitute 40% of all marriages. In the remaining 60% of the marriages we lack knowledge of at least one set of parents.

Blood marriages are listed at the END OF THE FILE, as shown for this example in Table 2. Table 2 is a PAR-CALC output that gives each type of prior consanguineal kin connection and information on their occurrence and frequency.

TABLE 2 Prior Kin Marriage Output of the P Graph Algorithm

```
Blood Kinship      No.    Frequency
between Spouses    of         & Percent
Kinship    Kin     Tie        Poss Per- Vector of Couples= List of  Equations
Types   F M Gn Frq ible cent                    Couples
G=FG       D       1 2 0   2/    8  50% ....1.1...110...000. 11 12
G=FG       BD      1 2 0   2/    8  50% ....1.1...110...000.  5  7
GF=FG      MBD     2 2 0   2/    4  50% ....-.-...--0...011. 18 19
GGF=FGGF   FMZSSD  3 4 0   2/    3  66% ....-.-...---...a11. 18 19
GGFG=FGGG  FMFBSSD 2 6 0   2/    2 100% ....-.-...---...-11. 18 19
GGG=FGGG   FFBSSD  1 6 0   2/    2 100% ....-.-...---...-11. 18 19
GGG=FGGFG  FFBDSSD 2 6 0   2/    2 100% ....-.-...---...-11. 18 19
GG=FGG     FBSD    1 4 0   1/    4  25% ....-.-...--1...000. 13
GFG=FGG    MFBSD   2 4 0   1/    2  50% ....-.-...--1...0--. 13
GF=FGF     MZSD    3 2 0   1/    2  50% ....-.-...--1...a--. 13
GG=FGFG    FBDSD   2 4 0   1/    1 100% ....-.-...--1...... 13
GFGFG=FGG  MFMFBSD 3 5 0   1/    1 100% ....-.-...---...1--. 17
GFGG=FGG   MFFBSD  2 5 0   1/    1 100% ....-.-...---...1--. 17
GGG=FGG    Fp-BD   1 5 0   1/    3  33% ....-.-...---...100. 17
GGFG=FGG   FMFBSD  2 5 0   1/    3  33% ....-.-...---...100. 17
```

```
G=F         Z       1 1 0   0/  14
GG=F        FZ      1 2 0   0/  12
GF=F        MZ      2 1 0   0/   9
```

Reading the first line of table 2, in p-graph notation, G=FG means that for couples 5, 7, 11, and 12 (listed at the right), the male's parents (G) are the same as the female's father's parents, but in two of these cases (11 and 12 - Lot and his daughters) it is also true that the male IS the female's father, hence under Kin Types we have the listing "D" for "marriage with Daughter." Note that the same percentage is listed under each of these two types of marriage, which are similar in p-graph notation. The percentage really reflects that of their <u>combined</u> frequencies.

The Kin Types in this table are composed of the following symbols p- paternal half m- maternal half Z sister B brother F father M mother S son D daughter.

Vector genealogy using F and G does not distinguish D from BD marriage or half-siblings from cousins. Referring back to TABLE 2, the program also considers the individual identities involved in these links. This, for example, a paternal half-sibling is distinguished from a FB child (or the maternal half-sister from a MZ child). The difference is shown in the first two subgraphs of figure 5.1 below. The other two subgraphs show two cross-cousin marriage graphs. One can verify from TABLE 2 (the Canaan genealogy) that D, BD and MBD occur but p-Z, FBD and FZD do not.

```
     o       o         o         o          o           o
     |\      +\       / \       /_\        / .         / .
     | \     | \     /   \     /   \      /   .       /   .
     |  o    |  o   o     o   o     o    o     o     o     o
     | /     | /     .   /     .   /      .   /       \   .
     |/      |/        . /       . /        . /         \ .
     o       o         o         o          o           o
     BD      D        FBD        p-Z       MBD         FZD
Marriages:
     5,7     11,12   None       None      18,19       None
```

Figure 5.1: P-graphs of some marriage types

The program recognizes sibling differentiation in multiple marriages as well as the half-sibling relation by two additional vectors (section 3.4) that give the sibling order or individual identification numbers of children. Using these vectors, the program can also recognize different types of polygyny.

The next two columns indicate that the numbers of paternal or male ties and maternal or female ties in this blood relation. In this example -- that of the biblical figure of LOT lying with his daughter, which is a couple, and produces a son, but is not a marriage -- this one male figure enters into TWO kinship links, one as his mate's father and one as her mate. Since the number of maternal links is even in all but couple 17's marriages, this could be a matri-moiety system (and the program advises the user of percentage of cases which satisfy this possibility).

The number of couples with this prior kin relation, the percentage of couples for which relevant information exists which have this prior kin connection, a list of such couples, and a vector in which the entry for each couple indicates wherever possible the presence or absence of this type of marriage. Any couple (such as 18 and 19) related in multiple ways will recur on two or more of the lists of couples. The percentage of marriages for blood kin types is computed by dividing the number of couples with this type of marriage by the number which have sufficient data to determine whether the marriage is present or absent, and excluding those coded '-' in the marriage vector, who are related by some closer marriage.

The next column (all 0s) are time slices, which are ordinarily all 0s if they are undifferentiated (see 8.4). The following three columns give the number of actual, possible and percentage of marriage of this type. Following this is a vector of (.-01) elements, that for each marriage:

1 present;
0 absent;
a inferred absent;
- absent because there is a closer relation within the kin type;
. insufficient data to determine.

```
                                                 11111111112
Thus, the third line in TABLE 2,        12345678901234567890
GF=FG       MBD     2 2 0   2/   8  50% ....-.-...--0...011. 18 19
```

can be interpreted as two marriages out of 8 possible (25%) with a mother's brother's daughter (where GF=FG).

Overall, TABLE 2 shows a tendency for patrilineal descendants of patriarchs to marry female patrilineal descendants and for the type of realization of endogamous patrilineal preference to change across the generations: Abram and his brother Nahor (5,7) marry nieces (BD); Esau marries a first cousin (FF^SD) and Isaac (13) a first cousin once removed (FBSD); Jacob marries a patrilateral second cousin once removed.

The second body of output in output file p-xxx.frq classifies all the blood marriages (^^^ except those coded '-') in terms of distance to the common ancestor for the husband and for the wife. **Table 3 shows the results for the Canaanite data set.**

Generational Degree (Distance to Common Ancestor): on the wife's side on Hu's
Side 1 2 3 4 5 6 7 20

```
      Freq Freq Freq Freq Freq Freq FreqTotal    %
    1    0    4    0    0    0    0    0    4   18
    2    0    2    2    1    0    0    0    5   23
    3    0    0    2    4    2    0    0    8   36
    4    0    0    2    2    0    0    0    4   18
    5    0    0    1    0    0    0    0    1    5
    6    0    0    0    0    0    0    0    0    0
    7    0    0    0    0    0    0    0    0    0
 Tot    0    6    7    7    2    0    0   22  100
  %     0   27   32   32    9    0    0  100  100
```

TABLE 3 Prior Kin Marriage Output of the P Graph Algorithm

Next, the program assesses the marriage data and addresses six questions in its output, as shown below for the Canaanite data:

Question 1: are there patterns of exogamy? 100% bilateral exogamy of degree 1
0 exceptions over 8 endogamous marriages
__% patri- exogamy of degree _ (compute by hand from qn3 data below)
__% matri- exogamy of degree _ (compute by hand from qn3 data below)

Question 2: Are blood marriages generationally skewed within various degrees?

degree : percent skewed
2: 66%

3: 50%
4: 77%
5: 100%
6: 0%
7: 0%

Question 3: In blood marriages, are there linking relatives of certain sex or generation? (F,M = ancestral Fa/Mo certain generations back)

frequency of linking relative on the husbands side by gender

```
                F vs M   F vs M   F vs M   F vs M   F vs M   F vs M   F vs M
at various
generations                1        2        3        4        5        6        7
removal for  RELATIVES     87  51   55  34   19  15    5   0    0   0    0   0    0   0
versus blood  MARRIAGES    12   6    8   5    4   1    1   0    0   0    0   0    0   0
GAMMA=                     -.079     .005    -.519
```

frequency of linking relative on the wifes side by gender

```
removal for  RELATIVES    114   0   53  25   15  18    4   0    0   0    0   0    0   0
versus blood  MARRIAGES    22   0   14   2    5   4    2   0    0   0    0   0    0   0
GAMMA=                      -.535    -.200
```

(+) gamma female link more likely for spouse than for relative (-) gamma male link more likely for spouse than for relative

```
        number of male links:
female
links    1     2     3     4     5     6
    1     0     4     0     1     1     2     8
    2     0     2     0     2     2     4    10
    3     0     1     0     2     1     0     4
Tot      0     7     0     5     4     6
```

Question 4: Is there evidence of dual organization (sidedness) in the marriages?

10 22 45% Viri-Sided
18 22 81% UxoriSided

Question 5: Is there a pattern of marriage with Iroquois-cross relatives? that is, where the linking parents of the bride and groom are opposite sex?

In terms of Iroquois (not Dravidian) crossness:
0% of blood marriages are patri-cross (Fa of husband related to Mo of wife)
27% of blood marriages are matri-cross (Mo of husband related to Fa of wife)

Question 6: What are the particular types and frequencies of blood marriage compared to relatives not married?

The answer to this question has been discussed in TABLE 2

==

7.0 WHERE TO GO IN YOUR ANALYSIS

==

7.1 Social Network Analysis

Genealogical graphs, like genealogies, seem to create the opinion in the viewer that what they are viewing is descriptive, but not a real network, since in a network one knows the relation of each person to every other. Since the p-graph shows only parent-child and marriage links, this reasoning goes, we do not really know all of the other kinship links between pairs or groups of people. Is this really true?

As Jorion and Lally (1983) showed, however, once the p-graph exists in the computer, it is easy to compute the relation between any pair (or all pairs) of persons. To see if they are related by descent, we compute the couples or individuals in their ancestral trees, and compare them for overlap. Using vector genealogy, this is done very quickly. To see if they are more indirectly related, including links by marriage, is more difficult. White (1997) shows that this could be done by the use of algorithms that find all the maximal blocks containing cycles of direct (descent) and indirect (marriage) links. This algorithm requires no more time to compute than the question of co-ancestral relations: in fact, it includes them as a special case of a cycle of relations, closing back on itself through a blood marriage. Hence it is simple to compute, on the fly, all the possible compound relations among individuals in the kinship and marriage network.

We could add, then, any computed relations that we want to a p- graph: close relations (like siblings, siblings-in-law), or selected distant ones. We could also compute all the relations of persons of the same lineage, as for example, everyone within 4 generations by descent, either through males or females, from a given ancestor. The p-graph represents the intersections between all such descent groups. The p-graph becomes a scaffolding for representing whatever relations we choose to be of interest.

Many of the types of network analysis for marriage structures are especially adapted to the problem (see White and Jorion 1992, 1996), White (1997), Brudner and White (1997), White and Schweizer (1997), Houseman and White (1996, 1997a, 1997b). Analysis of marriage and kinship networks are related to the formation of institutional social structures, to the study of changing social organization, the formation of communities, class, ethnicities and self-reproducing groups. (see section 7.5).

Other types of analysis are generic to social network analysis. Hage and Harary (1983, 1991, 1996) also give a variety of means for thinking structurally about social and cultural networks and how economic processes are bound up in them.

7.2 Network Analysis with UCINet

One should think carefully about what you will do with your kinship and marriage data after exporting to UCInet (see section 4.5f for automatic export). The graph theoretic block routine is appropriate for identifying maximal clusters of marital relinking (White 1997). Centrality analyses may be revealing. So too, blockmodeling. See section 10.2 for more information on data formats.

7.3 Economic, Social and Cultural Capital

Marriage and kinship networks, particularly through the process of relinking into marriage blocks (7.5), are sites of accumulation, in Bourdieu's terms, of economic, social and cultural capital. Rather that study such processes in the abstract, the network and p-graph approach to the transmission of property, symbolic goods, office, and social or role relationships is to map out these processes concretely on the scaffolding of kinship and marriage relations. The fundamentals of this conception were laid out by W.H.R. Rivers and reviewed by Langham (1981). Appadurai (1988), Brudner and White (1997), Houseman and White (1997a), and White and Schweizer (1997) have written extensively about this approach.

7.4 Processes of Percolation

If we look at longitudinal processes in social networks, resources -- social, cultural, economic -- do not flow, they percolate unevenly, depending of the features of the social web in which they are enmeshed. There is a whole new open field of study of these percolation processes. Rather than treating social processes as uniformly distributed or subject to cultural and normative orders, the various kinds of network densities and interlocks can be studied directly (see frontispiece).

7.5 Analysis of Relinking: Group, Class, Ethnicity

One of the key processes of interlock and group formation is that of relinking. The p-graph is the perfect vehicle for the study of relinking structures and processes, since it contains no cycles other than those formed by various types of relinking:

1. marital relinking through common ancestors (blood marriage)
2. direct marital relinking between two families
3. indirect marital relinking between three or more families. Relinkings as conscious and/or nonrandom strategies are documented through p-graph studies as occurring within 1,2,3,4,5 and perhaps higher number of generations. White (1996, 1997) and Richard (1995) have developed different methods for the statistical study of relinking. Brudner and White (1997) have shown how class formation/differentiation and marital relinking/nonlinking may be intimately bound up together.

7.6 Mapping Exchange and other relations and properties

Hopefully, the data set that you have processed is not simply genealogical, but contains a host of rich data about individuals and their relationships: dates and places of residence, occupations and officeholding, property and usufruct, educational attainment, friendships, exchange relationships, things exchanged, etc. A P-graph, and the PGRAPH program, can help you to label your graphs with the attributes or properties of individuals or couples, keep track of inherited properties, and -- with the aid of other graphics programs -- to superimpose exchange and other kinds of relations on the scaffolding of the kinship and marriage graph.

7.7 Simulation and Hypothesis Testing

One of the key developments in p-graph analysis is the testing of hypotheses, not only for what accounts for the frequency of different patterns, but as to whether the patterns are simply what would be expected given random permutations of the data: say you randomly reassigned people's marriages in each generation, holding constant their sibling-groups of origin. This option is discussed in section 9.4.

Simulation, or marriage permutation tests, make it possible to test a wide variety of null hypothesis models about kinship and marriage structures.

==

Appadurai, Arjun, ed. 1988. The Social Life of Things. Cambridge: University of Cambridge Press.

Bertin, Jacques. 1967. La Semiologie Graphique. Paris: EHESS.

Borgatti, Steve, Martin Everett and Linton Freeman. 1992-97. UCInet. Natick, MA: Analytic Technologies.

Brudner, Lilyan A., and Douglas R. White. 1997. Class, Property and Structural Endogamy: Visualizing Networked Histories, Theory and Society 25:1-48. Reprinting under review for Cambridge University Press, volume on New Directions for Formalization and Historical Analysis, John Mohr and Roberto Fransozi, eds.

Cuisenier, Jean. 1962. Endogamie et exogamie dans le mariage Arabe. L'homme 2(2):80-105.

Degenne and Flament. 1986.

Denham, Woodrow W., Chad K. McDaniel, and John R. Atkins. 1979. Aranda and Alyawara Kinship: A quantitative argument for a double helix model. American Ethnologist 6:1-24.

Duquenne. 1991.

Duquenne. 1996. Social Networks 18:.

Freeman, Linton, and Douglas R. White. 1993. Using Galois Lattices to Represent Network Data. Sociological Methodology 1993: 127-46.

Freeman, Linton, Douglas R. White, and A. Kimball Romney, eds. 1989[1991]. Research Methods in Social Network Analysis. George Mason Press. Reprinted, Transaction Press.

Guilbaud, G. Th. 1970. Syst/,me parental et matrimonial au Nord Ambrym. Journal de la Soci,t, des Oc,anistes 26:9-32.

Hage and Harary (1983

Hage, Per and Frank Harary. 1983. Structural Models in Anthropology. Cambridge: Cambridge University Press.

--- 1991. Exchange in Oceania: A Graph Theoretic Analysis. Oxford: Oxford University Press.

--- 1996. Island Networks. Cambridge: Cambridge University Press.

H,ran, Fran‡ois. 1995. Figures et L,gendes de la Parent,. Paris: Institute National d',tudes d,mograpiques.

Houseman, Michael. 1990.

Houseman, Michael and Douglas R. White. 1996. Structures r,ticulaires de la pratique matrimoniale. L'Homme 139:59-85.

Houseman, Michael and Douglas R. White. 1997a. Network Mediation of Exchange Structures: Ambilateral Sidedness and Property Flows in Pul Eliya, Sri Lanka. In, Thomas Schweizer and drw, eds. Kinship, Networks, and Exchange. Cambridge University Press.

Houseman, Michael and Douglas R. White. 1997b. Taking Sides: Marriage Networks and Dravidian Kinship in South America, In, Maurice Godelier and Thomas Trautmann, eds., 1997. Transformations of Kinship. Washington: Smithsonian Press.

Jorion, Paul. 1982

Jorion, Paul. 1984

Jorion, Paul, and Gisele de Meur. 1980.

Jorion, Paul, Gisele de Meur, and Vuyk. 1982

Jorion, Paul, and Lally. An Algorithm for the Analysis of Genealogies as to prior kin connections between spouses. 1983. unpublished ms.

Langham, Ian. 1981. The Building of British Social Anthropology, W.H.R. Rivers and this Cambridge Disciples in the Development of Kinship Studies, 1898-1931. Dordrecht (Holland): D. Reidel.

Lévi-Strauss, Claude. 1949. Les Structures Elementaires de la Parent. Paris: Presses Universitaires de France. [1969]. English transl. by James H. Bell, The Elementary Structures of Kinship. Boston: Beacon Press.

Lévi-Strauss. 1962 La pensee sauvage. 1962 Structural anthropology. 1963

McConnel, Ursula. 1951.

Romney, A. Kimball. 1971.

Romney, A. Kimball, and David Strauss.

Rose, Frederick G. G. 1960. Classification of kin, age structure, and marriage amongst the Groote Eylandt aborigines; a study in method and a theory of Australian kinship. Berlin, Akademie- Verlag.

Stern, Jeffrey. 1992. Proposal: A Network Graphics Package. University of California, Irvine.

Warner, W. Lloyd. 1931. A black civilization; a social study of an Australian tribe. New York: Harper & Brothers.

Weil, Andre. 1949

White. 1985. Entailment Analysis Package. Irvine, CA: University of California.

White. 1996. Statistical Entailments and the Galois Lattice. Social Networks 18:201-215.

White Douglas R. 1996. Enfoque de redes al estudio de comunidades urbanas, Estudios Demográphicas y Urbanas 9:303-326. Mexico City: Colegio de Mexico.

White Douglas R. 1997. Structural Endogamy and the graphe de parent,. Math,matique, Informatique et sciences humaines 137. Paris: Ecole des Hautes Etudes en Sciences Sociales.

White Douglas R., and Paul Jorion. 1992. Representing and Analyzing Kinship: A Network Approach., Current Anthropology 33:454-462.

White Douglas R., and Paul Jorion. 1996. Kinship Networks and Discrete Structure Theory: Applications and Implications. Social Networks 18:267-314.

White, Murdock and Scaglion. 1971

White, Douglas R., and Karl Reitz. 1983. Graph and Semigroup Homomorphisms. Social Networks 5:193-234.

White, Douglas R., and Karl Reitz. 1989 Rethinking the Role Concept: Homomorphisms on Social Network. In, L.C. Freeman, D.R. White, A.K.Romney, eds., Research Methods in Social Network Analysis, 429-88.

White, Douglas R., and Thomas Schweizer. 1997. Kinship, Property and Stratification in Rural Java: A Network Analysis, In, Thomas Schweizer and drw, eds. Kinship, Networks, and Exchange. Cambridge University Press.

==