Jacob's Ladder 11.0
Multidimensional Data Animation, Visualization and Intonation Application for Creating Virtual Reality Displays

Benjamin Elbirt[1]
School of Informatics
Department of Communication
SUNY Buffalo
sarq_6669@yahoo.com
elbirt@buffalo.edu

Abstract

This paper introduces Jacob's Ladder 11.0©, software for the visualization, intonation and animation of multidimensional data in a virtual reality (VR) space. Jacob's Ladder 11.0© is targeted toward individuals interested in the examination of social networks and attitude relationships or any research involving the use of multidimensional data matrices. The software provides individuals with the ability to animate and/or overlay real-time multidimensional data of any type or size in a VR space with maximum variability in dimensional representations including, but not limited to, object size, type, position, color, transparency, temporal positioning and audio frequency. This paper then demonstrates Jacob's Ladder 11.0© data describing the migration among Canadian provinces. Results are presented along with implications and future directions for this software development.

Keywords: Data Visualization, Network Analysis, Multi Dimensional Scaling (MDS), Virtual Reality (VR), Process / Temporal Animation

---

[1] Special thanks to Dr. George Barnett and Dr. Joseph Woelfel for their assistance in this project.

Introduction

Visualization of social process is an important analytical tool for understanding social change and the affects of specific events on this process. Understanding these events related to process helps the researcher direct and predict the future within a studied environment.

Examples of visualization can be found throughout social science research tools. Software like ThoughtView (Woelfel, 1998), Multinet (Seary, 2005), and UCINet (Borgatti et. al., 2002) are a few examples of modern visualization techniques that represent complex data structures in multidimensional space. However, the software available to researchers does not provide a true 3D environment and are unable to produce any animated sequences of events in such an environment. They are limited to 2D instant snapshots of time with limited 3D perspectives through 2D displays.

Jacob's Ladder 11.0© is the first publicly available version of a private software research project originally developed to represent coordinate data sets from Galileo™ and CatPac™ software packages in a VRML space. Galileo™ and CatPac™ are software designed to map the beliefs and attitudes of people based on metric distance pair relationships and fuzzy logic neural network analysis (Woelfel, 1998).

Unlike previous versions, Jacob's Ladder 11.0© is a noninvasive binary executable for Windows™ that uses Microsoft Excel™ files for input, thus making data translation from other software packages easier. It can animate and overlay multiple data sets in the same space and currently has the capacity to visually represent nine simultaneous dimensions of data[2].

VRML (Virtual Reality Markup Language) was chosen because it was the best technology available for Internet based virtual reality displays at the time (1997) and continues to be the simplest method of representing and animating 3D spaces. Data distribution, real-time update and cross-platform considerations further support the use of VRML as the visualization technology for multidimensional spaces.

Design

The earlier versions of Jacob's Ladder© were UNIX based Perl scripts that converted Galileo™ and CatPac™ Riemann space coordinate output files into Euclidian VRML space allowing the researcher to transverse the space and look at the data from any perspective along the *x*, *y*, and *z* plane. A semi-transparent *x*, *y* and *z* axis plane were rendered for perspective and each node in the space was represented by a sphere with a radius defined by the user. The first three dimensions of the coordinate data were used to position the nodes within the VRML space and each node sphere was labeled and colored as per user specification.

The representation of coordinate data in a multidimensional space is beneficial to applications other than Galileo™ and CatPac™. Jacob's Ladder 11.0© is specifically designed to allow any multidimensional matrix to be represented visually in a VRML space with as much variance in presentation methods as possible. The Microsoft Excel™ file format was chosen for inputs to accommodate the vast applications available to researchers for analysis. Most delimited and fixed-width data outputs can be read and

[2] Three for coordinates, three for color, and one for radius, transparency, and frequency.

converted into a Microsoft Excel™ file with little effort or data loss. Furthermore, Jacob's Ladder 11.0© can read the Microsoft Excel™ file directly.

The Jacob's Ladder© input format was simplified making most matrix import directly applicable to the final Microsoft Excel™ input file. Any space within Excel™, unused by the Jacob's Ladder© format, is available for content storage relevant to the data, but unnecessary to Jacob's Ladder©. The immediate advantage is keeping data in one location.

Animation of data is achieved by using multiple spreadsheets of data available in the Excel™ file. Each of these sheets is a mirror of the other structurally, however the data used for all the object and world attributes will be representative of the specific spreadsheet when read. There is no limit on animation periods because Jacob's Ladder 11.0© can sequentially read any number of files specified in a stack as though they were all in one file.[3]

Each spreadsheet in the Excel™ file is a data set that defines title, origin, labels, color, radii, transparency, sound frequency, sound instrument, axis origin, axis color and axis transparency. The Jacob's Ladder 11.0© interface allows for the specification of which axis to render, which coordinates to animate, which features to animate, environment controls, dynamic viewpoints, sound controls, normalization options, animation options and file IO related specifications. These definitions can be saved to a Jacob's Ladder© definition file for later use making the reapplication of method across multiple datasets possible.

Normalization is a process that allows the user to collapse or expand data to any space desired without loss of spatial relationship integrity. Most cases require some form of normalization because data is often densely packed or overly expansive making the visualization difficult to see. Jacob's Ladder 11.0© includes normalizations for coordinate positions, radii, sound frequency, transparency and color

Spatial relationships are limited to the first three dimensions of the data, representing the *x*, *y* and *z* coordinates in Euclidian space. Dimensions above the first three may be represented as sphere radii, color, sound frequency, and transparency where the larger the value, the greater the radius, brighter the color, higher pitched the sound and *less* transparent (opaque) the sphere. The researcher can choose to make these visualizations a function of data dimensions above the first three *or* hard code the value.

Colors in VRML are represented as RGB values. Each value can be between *0* and *255*, where *255* is absolutely on, and *0* is absolutely off. *R* is red, *G* is green and *B* is blue. All three at *255* (*255,255,255*) is white and all three at *0* (*0,0,0*) is black. Each color element can be specified as a function of a spatial dimension or hard coded resulting in $\mathbf{(N-3)^3}$ combinations if **N** is the number of total dimensions in the data matrix and that value is at least four[4].

---

[3] MS Excel™ has spreadsheet limitations of 2048 spreadsheets consisting of no more than 250 columns and 65,536 rows per spreadsheet. This feature breaks the 2048 spreadsheet limitation.

[4] The first three dimensions of space are limited to the X, Y, and Z coordinates and cannot be referenced for other characteristics at this time. Future releases may alter this restriction, however no plan is currently in place to allow the specification of specific dimensions of data as specific coordinates in space freeing up the first three from spatial presentation limitations. Researchers who want to change the dimensions used for the X, Y and Z coordinates should alter the presentation order of the data dimensions in the Excel™ file to put the desired dimensional data in the first three columns.

Colors are an important visualization method used by Jacob's Ladder©. The validity of massive data volume is achieved by assigning a single dimension of space as the color representation. Stronger values will have brighter colors and any odd value appearances can be easily detected by the absence or presence of color. Colors can also be used as factor/characteristic measurements.

Animations can be setup to utilize colors that represent different periods of time or as animated changes in strengths. Hard coding colors in non-animated sets of data, where two or more are presented within the same coordinate space, allows the researcher to differentiate between those sets.

Transparency is similar to Color in representation of dimensional data. However, it controls how opaque the object is. The larger the data value, the more opaque the object. Completely transparent objects *cannot be seen* and completely opaque objects *cannot be seen through*. The concept object labels can be forced opaque or allowed to vanish with completely transparent values **(T = 0)**. This creates a blinking in and out effect and can be used to hide objects that appear later in the animation.

A base transparency is recommended when hard coding the value allowing stacked objects to be seen one within the other. Animation can alter transparency between time periods; however the amount of noticeable difference can vary based on the data and may not be completely visible during the animation process.

Radii can be dimensionally based or hard coded and determine the size of the concept's spherical representation. Radii can represent error ratios, sample sizes or any other characteristic value. A ten-to-one proportion is a reasonable world to radii size relationship.

Jacob's Ladder© provides an automatic radii normalization feature. Defining the world normalization size allows the researcher to continually re-create the VR output until a proper balance is achieved.

Audio provides a new means of data presentation – the ability to hear things (Barnett, 1992) Jacob's Ladder© allows for a frequency specification of **(0 <= N <= 12,191)** and will normalize the data used to a specified value in that range. By default, it will always normalize value sets larger than **12,191** to prevent invalid frequency values. All dimensional values used are taken as an absolute value prior to normalization.

Migration between Canadian Provinces

The data provided for demonstrating Jacob's Ladder 11.0© describes the patterns of migration among the Canadian provinces and territories. It is part of an ongoing study about social cohesion in Canada; the ties that bind the country together given the possible separation of Quebec (Barnett, Sung & Park). The data represents quarterly migration for the first quarter 1987 to the first quarter of 2005 as obtained from Statistics Canada (Belanger, 2000). The data on the number of migrants are based on Revenue Canada Tax and child tax credit files. They consist of 73 matrices, where each cell s_{ij}, indicates the number of people who migrated from province or territory *i* to territory *j*. The matrices are not symmetrical ($s_{ij} \neq s_{ji}$). The first 38 data sets contain the frequency of migration among 12 provinces and territories. The final 33 are among 13, with the addition of Nunavut in 1996 (3^{rd} quarter). Data on the population of the provinces and territories were also obtained from Statistics Canada. Only annual data were available for 1987-1996. Quarterly data was available from 3^{rd} quarter 1996.

File Contents

The Jacob's Ladder© data set contains 13 nodes (the Canadian provinces and territories) (**N**) by 13 dimensions (**M**) for 10½ year quarters or 42 spreadsheets plus an additional 13 nodes (**N**) by 15 dimensions (**M**) for 7¾ year quarters or 31 spreadsheets. This numbers 73 spreadsheets and 13,143 unique cells of data. With color, frequency, transparency and radii definitions included (2,847 cells) and the axis definitions (511 cells) the total is 16,501 unique cells of data.

The multidimensional inputs for Jacob's Ladder© were generated using the Galileo program (Woelfel & Gilham, 1977). The coordinate for each point in time were created through a metric multi-dimensional scaling of the socio-matrices. Then the coordinates were rotated to the previous point in time such that departure from congruence was minimized (Woelfel, Holmes & Kincaid, 1988). The rotated coordinates were entered into Excel for input into Jacob's Ladder©.

Jacob's Ladder© can process the Excel™ in less than 30 seconds on an average laptop computer including all normalization procedures[5]. Tests have shown significant improvement in processing speeds when using flat files vs. Excel™[6] and the use of databases like MySQL or Oracle can substantially improve performance when aggregating large volumes of data into matrixes and performing real-time updates.

Final output content varies based on display type. The animation of 73 points in time results in a 335K VRML file and the non-animated output files containing all objects averaged 716K. A 28.8k modem will require roughly 12 seconds downloading the animated files and roughly 25 seconds for the non-animated file.

Display Contents

Three viewpoints were created with Jacob's Ladder© to review the Canadian migration data – complete cluster, seasonal cluster and time animated. Seasons/quarters are represented by color where autumn is yellow, winter is blue, spring is green and summer is red. Node radii are the population measures, transparency is set to opaque and sound frequencies are the mean link strength of each socio-matrix. All world sizes are normalized to +/- 500 meters with radii of 50 meters[7].

---

[5] The processing time is primarily reading the Excel™ file and writing the Jacob's Ladder© output file.
[6] Excel is a more manageable format for the data and it is the preferred method at this time.
[7] VRML uses meters as the unit of measure for linear distance, radians for angles and seconds for time (Carey & Bell, 1997).

Complete Cluster

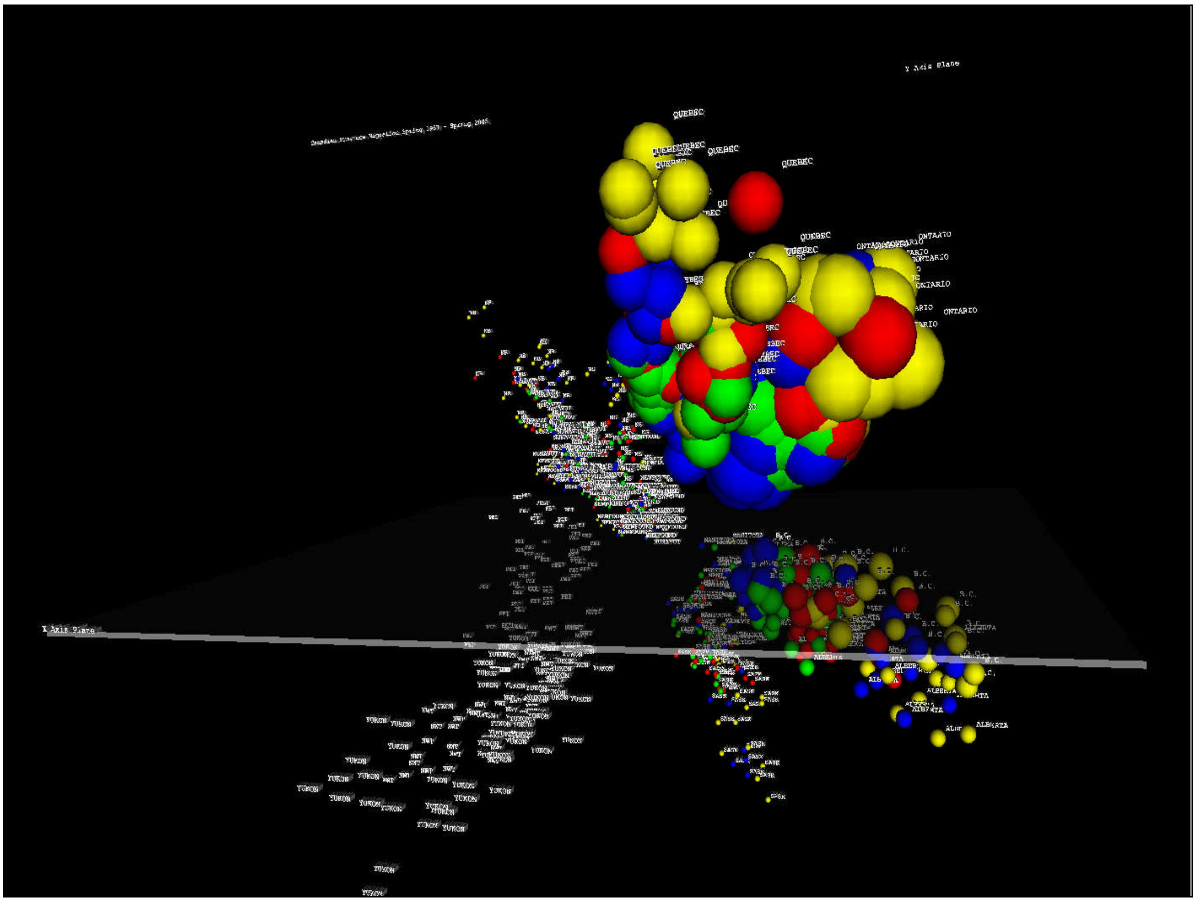


Figure 1 – Complete Cluster

Figure 1 displays clusters based on the overall period. There were consistent positions within a relatively small area for each of the more populous (larger) provinces and the same consistency over a slightly larger area for each of the less populous provinces. Provinces with higher migration exchange remained closer together across all seasons (Barnett, Sung, Lin & Hung, 2003).

The clusters shown in Figure 1 can be broken down into three primary clusters described here as *OQ*, *RR* and *LR*. *OQ* is the closest cluster consisting of Ontario and Quebec. *RR* is the rear right cluster consisting of Alberta, B.C., Manitoba, and Sask. *LR* is the left rear cluster consisting of Yukon, NWT, PEI, NUNAVUT, NEWFOUND, NB, and NS.

OQ displays the spring and winter cluster folding into summer and fall cluster indicating a common migration relation among seasons. That Ontario and Quebec are the only two nodes in the cluster suggests a very high correlation of migration patterns between the provinces, irrelevant of season.

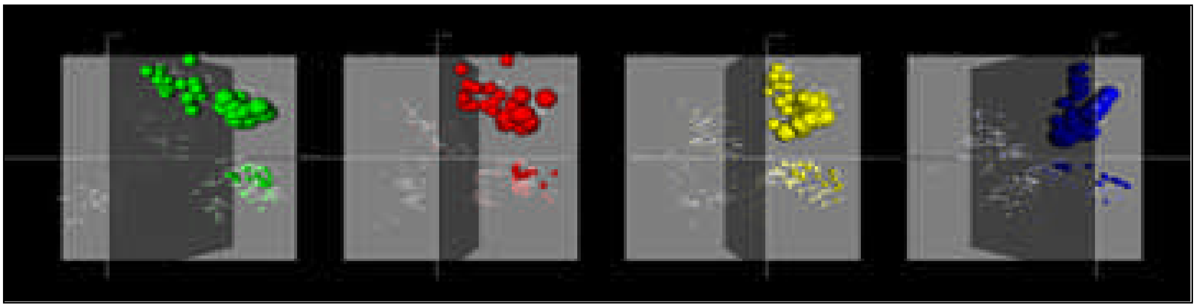
RR forms two sub-clusters that are related by Manitoba. The farther cluster consists of Sask. and Manitoba; the closer cluster containing Alberta and B.C. Manitoba's proximity to Alberta and B.C. suggests it as a bridge for Sask. The seasonality is more overlapped with a high concentration of winter and fall migration for

B.C. and Alberta. B.C. and Alberta also share a high winter and spring correlation with summer overlapping the two.

LR forms two sub-clusters with PEI as a central node. The lower cluster consist of NWT and Yukon; the upper cluster of PEI, NUNAVUT, NEWFOUND, NB and NS. NB positions toward the upper area of the cluster with PEI toward the bottom where Yukon and NWT cluster. Seasonal interrelations are too complex to suggest meaning in this display; however the Seasonal Cluster section will show the relationships to be consistent with those described here for other clusters.

It should also be noted that the three main clusters all seem to point toward the center suggesting a migration level between all provinces.

Seasonal Cluster



Spring ***Summer*** ***Autumn*** ***Winter***

Figure 2 – Seasonal Cluster

This display was used to determine clusters based on season. There are consistent positions within a relatively small area for each of the more populous (larger) provinces and the same consistency over a slightly larger area for each of the less populous provinces.

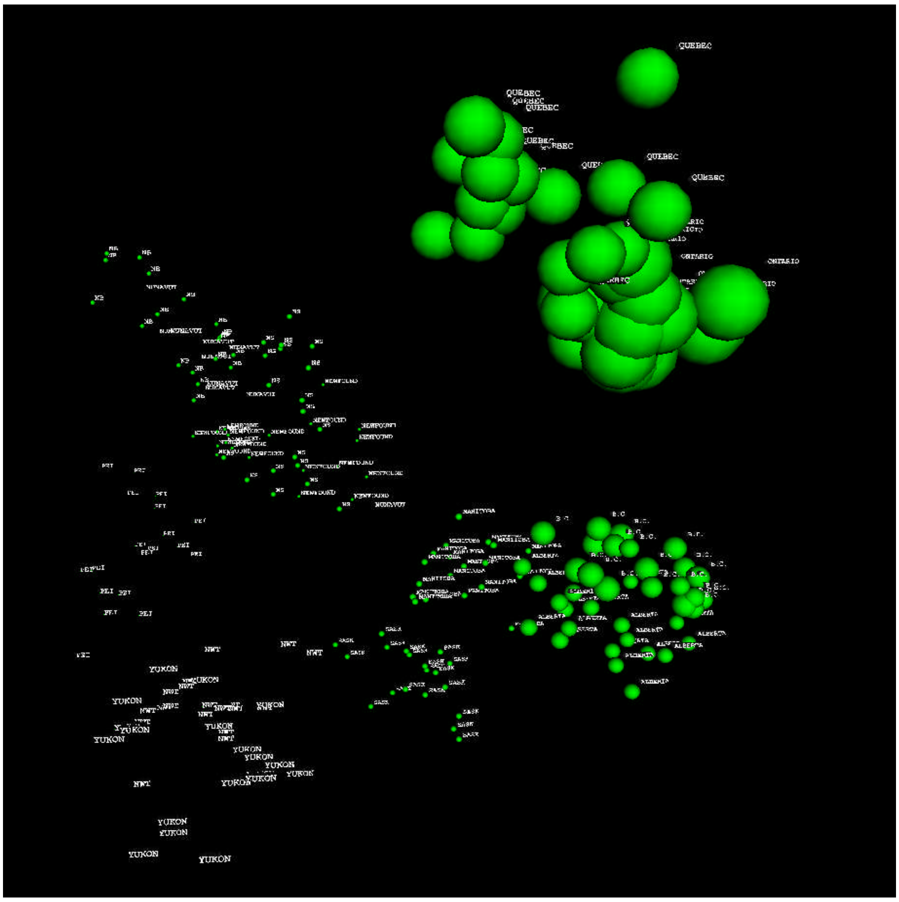


Figure 3 – Spring Season

The primary cluster patterns described in the Complete Cluster section are the same for each season. Although distance comparison is not possible here, proximity by season maintains the relationships between the different clusters *OQ*, *RR*, and *LR.* The sub-clusters of *RR* and *LR* are also maintained suggesting a very persistent seasonal and provincial relationship.

Time Animated

The time animated display was rendered using two intervals (speeds); one second intervals and 0.2 second intervals and the mean link strength for audio frequency.

At one second interval the animation showed a clear oscillation between seasons as previously described in the Seasonal and Complete sections above. At the faster speed (0.2) it became evident there was a massive shift when NUNAVUT appears. This is because the province was created after the start of the data set and the creation caused in large shift in population densities of the area.

As previously noted, there are multiple instances of seasonal disparities within the non-animated displays. The one second interval indicated a secondary oscillation in the data. Further analysis found winter and spring oscillating on an annual basis with regard to amount of migration and thus they appear in multiple clusters within the still displays. This oscillation is repetitive from year to year.

Intonation further supports the oscillation periods. Audio pitch rises for summer and fall and lowers for spring and winter indicating a greater migration during the summer and fall; lower for spring and winter. Careful attention to the winter and spring tones suggests the secondary oscillation previously mentioned.

Future Research

Jacob's Ladder 11.0© demonstrates a method for visualizing multidimensional data over time. This test demonstrates the social process of Canadian migration between provinces and provides a series of displays for viewing clustering relationships among the provinces. Animation of social change provides a new method of pattern recognition and advanced animation controls allow researchers to find patterns of change in large data sets that may be difficult to find through mathematical pattern recognition alone.

The use of VRML, Windows™ and Microsoft Excel™ allows for rapid distribution of content, large scale data sets and common application environments already in use by the social science community. Advanced controls in animation, visualization and intonation provide near limitless display variation for analysis.

Jacob's Ladder 11.0© is an ongoing project that explores data visualization of animated social change, cognitive modeling and any other related projects that can find use of such software. Theoretical applications include the modeling and social process, social cognition, animation of social network / graph theory, neural network growth and activation, neural network energy distributions and flow, information theory, chaos theory, spatial cognition, chemistry, biology, astrophysics, and particle physics.

Feature improvements will include polygon objects with face color and transparency[8], orbital dimensional representations, connection link properties and visualization, real-time data feeds, animation and non-animation combined display, basic functionality improvements, output file size reduction, character delimited input files, database query inputs, keyboard controls and shadowed animations.

[8] With regards to color this changes the number of potential combinations from $(N-3)^3$ to $(N-3)^{3F}$ where N is the number of unique nodes/objects in the space, and F is the number of faces available on the polygon used for each object. Example: F=6 (cube) |N=13 or $(13-3)^{3*6}$ or 10^{18} combinations.

References

Barnett, G. A. (1992). "The Acoustical Presentation of Technical Information." Journal of Technical Writing and Communication(22): 39-52.

Barnett, G. A., E. J. Sung, et al. (2003). An examination of Canada's Interprovincial Networks. Sunbelt Social Networks Conference. Cancun, Mexico.

Barnett, G. A., E. J. Sung, et al. (2005). A network analysis among the Canadian Provinces, SUNY Buffalo.

Bélanger, A. (1999). Report on the demographic situation in Canada 1998-1999: Current demographic analysis. Statistics Canada (Catalogue no. 91-209-XPE), Ottawa.

Borgatti, S., M. Everett, et al. (2002). UCINET 6 for Windows - Software for Social Network Analysis, Analytic Technologies, Inc.

Carey, R. and G. Bell (1997). Annotated VRML 97 Reference Manual, A-W Developers Press is a division of Addison Wesley Longman, Inc.

Seary, A. J. (2005). MultiNet: An interactive program for analysing and visualizing complex networks., Simon Fraser University.

Woelfel, J. (1998). Galileo System Manual: System Overview and Program Descriptions, The Galileo Company.

Woelfel, J. and J. Gilham (1977) "The Galileo system of measurement: Preliminary evidence for precision, stability, and equivalence to traditional measures." Human Communication Research **3** (3): 223-234.

Woelfel, J., R. A. Holmes, et al. (1988). Rotation to Congruence for General Riemann Surfaces under Theoretical Constraints. International Communication Association. Philadelphia, PA.