INTERFACE DESIGN AND KNOWLEDGE ACQUISITION FOR CARTOGRAPHIC GENERALIZATION

Harry Chang
Robert B. McMaster
Department of Geography
414 Social Sciences Building
University of Minnesota
Minneapolis, MN 55455

ABSTRACT

This paper reports on a research project that has designed a user interface for cartographic generalization that is to be used for acquiring procedural knowledge from maps. This graphical user interface consists of a series of pull-down menus, slider bars, and multiple windows to compare the results of the various generalizations. For instance, a user selecting the generalization operator, ***simplify***, is given a series of specific algorithms to select, each presented in its own window. Each window is provided with a slider bar that allows the individual to control the degree of simplification. An additional feature allows the user to apply a second operator, ***smooth***, to the data, and to overlay the original, ungeneralized, feature as well. Through the utilization of multiple windows, the user is able to compare the results of different simplification and smoothing algorithms. A last feature allows an individual to ***animate***, or generalize a feature in real time. The design provides for experimentation and enables the user to directly manipulate the tolerance values and to ascertain the quality of the generalization.

Additional operators, such as enhance, amalgamate, and aggregate, are now being added to the system. When complete, the system will allow users to generalize a database while their progress--in terms of operations and tolerance value selection--is logged. The information retrieved from these user sessions will allow the acquisition of procedural knowledge. Procedural knowledge allows control of the individual generalization operators and algorithms. The acquisition of such procedural knowledge will, ultimately, allow cartographers to build a partial knowledge base for cartographic generalization and design. The paper reports on the design of the user interface, early progress in knowledge acquisition, and the plans for development of the knowledge base.

KNOWLEDGE ACQUISITION

Several researchers, including McGraw and Harbison-Briggs (1989), have suggested several methods for acquiring knowledge, including: (1) interviewing experts, (2) learning by being told, (3) and learning by observation. **Interviewing** involves the knowledge engineer meeting with and extracting knowledge from the domain expert. In **learning by being told**, the expert "is responsible for expressing and refining the knowledge" while the knowledge engineer handles the design work (McGraw and Harbison-Briggs, 1989, p. 9). The third approach, learning by observation, is more complicated. Here the expert is allowed to interact with sample problems or case studies, or even use previous case histories. In the field of cartography and GIS, important initial decisions will involve how, exactly, do we acquire cartographic knowledge, or, specifically, which of these techniques do we use.

Types of Knowledge

In knowledge acquisition in general, and in cartographic applications specifically, there is great difficulty in the transfer of information from the expert to the knowledge base. It has been asserted, for instance, that specific knowledge acquisition techniques should be applied to different types of knowledge. This raises the important question, "What are the general forms of human knowledge?" Three types will be discussed here: procedural, declarative, and semantic.

Procedural knowledge encompasses intrinsic human abilities including motor skills, such as walking, and mental activities, such as language. Examples of procedural knowledge from cartography include generalization operations, such as smoothing and displacement, as well as other activities including classification and interpolation in symbolization. For cartographic generalization, Armstrong has defined procedural knowledge as that "which is necessary to select appropriate operators for performing generalization tasks" (Armstrong, 1991, p. 89). As an example, certain generalization operations, such as displacement and merging, require sophisticated knowledge on feature conflict and spatial interference.

In a general sense, declarative knowledge represents the facts that we carry around. For instance, we carry thousands of bits of information, such as our street address, spouses birthday, and color of our car. In cartography, declarative knowledge would include the matrix of Bertin's visual variables, rules for dot placement, and position for type placement. Semantic knowledge, as described by McGraw and Harbison-Briggs (1989, p. 22), "represents one of the two theoretical types of long-term memory. It reflects cognitive structure, organization, and representation." Furthermore, the authors state, "Because this type of knowledge includes memories for vocabulary, concepts, facts, definitions and relationships among facts, it is of primary importance to the knowledge engineer." In cartography, semantic knowledge is often associated with the generating processes of a feature. A river, for instance, may geometrically behave much differently depending on whether it is mountainous or coastal. Thus knowledge of the geomorphological metamorphosis of a river may be necessary for both generalization and symbolization. It appears that this is similar to what Armstrong calls "structural knowledge".

It should be noted that Armstrong also identifies geometrical knowledge as feature descriptions encompassing absolute and relative locations. Geometrical knowledge, in the format of spatial primitives, is contained in the Digital Cartographic Data Standard.

It is clear from the variety of knowledge types that different techniques will be necessary for successful knowledge acquisition in cartography. In existing rule bases for symbolization and generalization, such as those reported in McMaster (1991) for the Defense Mapping Agency and Mark (1991) for the United States Geological Survey, there exists a wealth of information on geometrical and descriptive knowledge. Transferring this factual information into a set of codified practical rules will be challenging. Buttenfield (1989, 1991) makes good progress at extracting structural or semantic knowledge using structure signatures, and knowledge from her measures will be critical in differentiating geomorphological line types and perhaps adjusting tolerance values. Other semantic information may be added at the time of encoding, such as the USGS's Enhanced Digital Line Graph data. There is a paucity of work, however, in the acquisition of procedural knowledge for cartography. As a generalized model in terms of knowledge type and extraction technique, McGraw and Harbison-Briggs (1989) present the following structure:

KNOWLEDGE	ACTIVITY	SUGGESTED TECHNIQUE
Declarative	Identifying general (conscious) heuristics	Interviews
Procedural	Identifying routine procedure/tasks	Structured interview Process Tracing Simulations
Semantic	Identifying major concepts and vocabulary	Repertory Grid Concept Sorting
Semantic	Identifying decision-making making procedures/heuristics	Task Analysis Process Tracing
Episodic	Identifying analogical problem solving heuristics	Simulations Process Tracing

[from McGraw and Harbison-Briggs (1989, p. 23)]

Those involved in building expert systems and knowledge bases in cartography and GIS will have to increasingly use innovative techniques for knowledge extraction. Simulations provide a good method for acquiring the difficult-to-extract procedural knowledge. The design of task-oriented user interfaces for procedural knowledge acquisition that allow the user to experiment and simulate specific cartographic processes appears to have great potential. One such task is cartographic generalization.

CARTOGRAPHIC GENERALIZATION

Although the process of digital cartographic generalization, a significant aspect of visualization, advanced quickly during the period 1965 - 1980, little progress has been made during the 1980s. Most of the initial progress resulted from work in the development of algorithms (such as the well known Douglas and Peucker line simplification routine, a variety of techniques for smoothing data, and algorithms for displacement, such as those by Nickerson), and attempts to analyze both the geometric and perceptual quality of those algorithms. Recent attempts at developing a more comprehensive approach to the digital generalization of map features--such as the application of simplification, smoothing, and enhancement routines either iteratively or simultaneously--have not been, for the most part, successful (McMaster, 1989). This stems in part from our lack of procedural information--or knowledge--on generalization. Such procedural knowledge includes decisions on which techniques are applied to actually generalize map information, the sequence in which these techniques are applied, and what tolerance values, or parameters, are used. Until researchers working in the spatial sciences have access to such procedural knowledge, a comprehensive approach to cartographic generalization will not be possible.

Currently, there are no commonly available logical interfaces that enable individuals to "experiment" with map generalization and ultimately to acquire such procedural knowledge. It is, however, worth mentioning a few of the previous packages with generalization capability. Perhaps the first attempt to include generalization within a mapping package was the Harvard Laboratory's SYMAP. With SYMAP, the generalization of surfaces was possible through ELECTIVE 38: Trend Surface Analysis. The user, after constructing the required database, could select both the order of the surface and the creation of a residual map. Much later, with SAS/GRAPH, the possibility for line generalization of coordinate strings was made available. SAS/GRAPH included PROCEDURE GREDUCE which applied the Douglas (Peucker) algorithm. The algorithm was controlled with two parameters: the NX value (which represented a density level of 1-5) and the EX value (which allowed the specification of a distance tolerance). The lack of any interactive user interface required a user to iterate through the NX and EX parameters until desirable output was obtained. It was a confusing, frustrating, and time-consuming process. More recently, one can point to the interface (TRANSFORMATION WINDOW) provided by the MAP II Processor as an example of a much improved raster-based design. Here, the user selects both a MAP and a FILTERING technique from a window. After this selection, a set of specific FILTERING techniques is displayed. The user may APPLY the filtering technique, view the displayed map, and immediately select an alternative technique.

Two very important aspects of this generalization user interface include (1) an on-line (hopefully hypermedia) help system and (2) the ability to, for certain operations, perform "animated" generalization. For instance, a user could request an animated simplification of a contour line where, as the line is continually reduced, the user could stop the sequence at an acceptable level.

USER INTERFACE DESIGN

A project in the Department of Geography at the University of Minnesota involves the development of such a graphical user interface (GUI), designed specifically to gain the "procedural" knowledge involved in map generalization (Figure 1). Using a Sun SPARCstation, a user interface (designed using SUNPhigs) has been developed. The basis for the user interface is a set of multiple windows that allows the user to experiment with different simplification and smoothing algorithms. For each of the four windows created, a separate simplification algorithm is available. The user, through direct manipulation of a series of sliding bars for each window, may quickly change the scale, position of the feature, and tolerance value. It is also possible, for each of the simplification windows, to pull down a procedure window that allows:

(1) overlay of the original feature, (2) smoothing of the feature using a series of algorithms, and (3) animation. In the animated sequence the feature is simplified from the densest set of points to a caricaturized version in real time.

For instance, using window [1] a user could (1) simplify a feature using the Lang algorithm, (2) smooth the feature with a five-point moving average, and (3) overlay the original for comparative purposes. The system is now being improved with the addition of geometrical measures, such as change in angularity, and the option of stopping the animation at any time. Future development will incorporate additional generalization operators. Eventually, the user interface will be used with the generalization test data set, developed by the NCGIA, to gain procedural knowledge on generalization from trained professional cartographers, or "domain engineers". As a cartographer works with the image via the interface, a generalization "log" will be maintained. Such a log will record, for each feature, the application and sequencing of operators, along with specific tolerance values. Such knowledge will be used to determine, in a holistic manner, how maps are generalized by evaluating the relationship among operators, parameters, and features. In effect, crucial procedural information will be recorded.

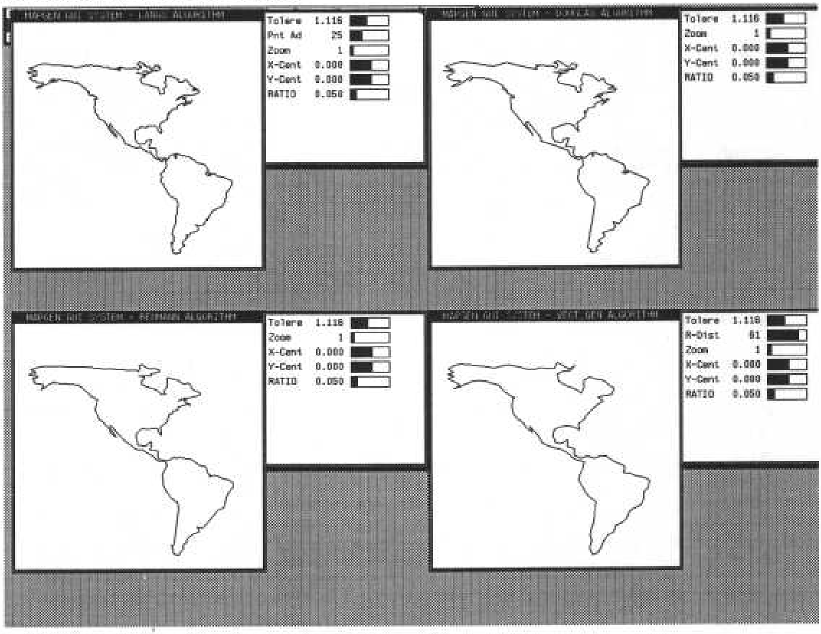

Figure 1. Generalization user interface with four simplification operators.

As GISs become more sophisticated, and the potential for quality cartographic output is finally realized, it will be necessary to pursue the application of expert systems. This will initially be in fairly narrow domains, such as typographic placement, symbolization selection, and generalization. The significant problem in the use of expert systems is in acquiring and formalizing the necessary cartographic knowledge. Capable domain and knowledge engineers must be identified, specific knowledge acquisition techniques need to be developed, and well-thought out simulations for complex cartographic processes should be designed. In developing a research agenda, then, these are the critical initial issues that need to be addressed.

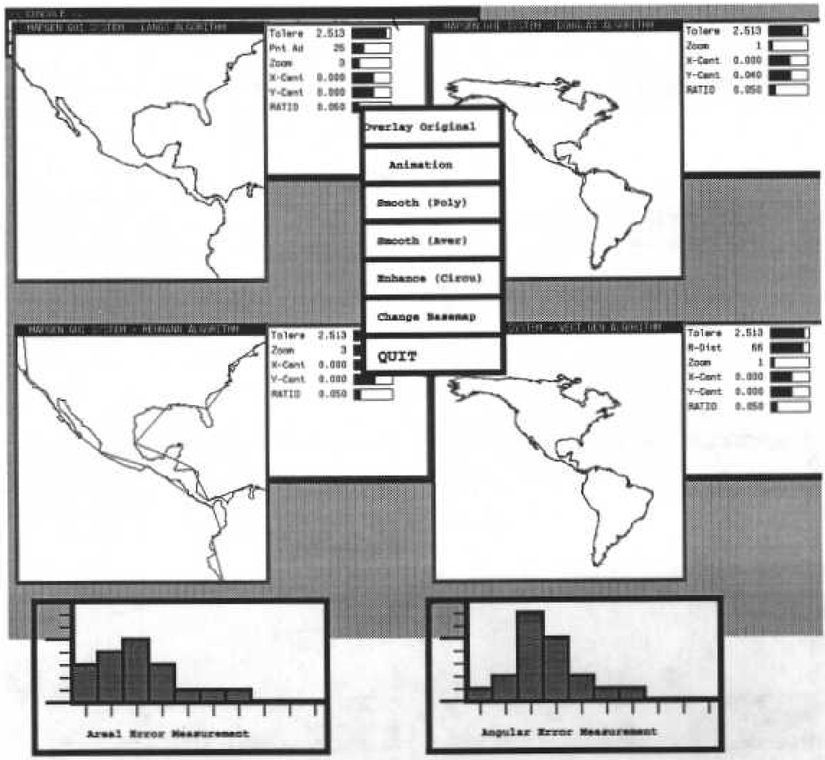

Figure 2. A variety of scales for the simplified image with simplification and original line depicted. Note the pull-down menu available for each of the original windows with selections for: (1) original overlay, (2) animation, (3) polynomial smoothing, (4) weighted-average smoothing, enhancement, and reselection of basemap.

Figure 1 depicts the interface, as copied from the Sun Workstation screen into a Postscript file. The Map Generalization System (MGS) provides the user with four windows, each with a simplification routine. It was decided, given the current level of knowledge in cartographic generalization, that most users would initiate the process through simplification. The four algorithms available include: Lang algorithm (top left), Douglas algorithm (top right), Reumann algorithm (bottom left), and Vectgen (bottom right). Descriptions of these algorithms may be found in McMaster (1987a) and are evaluated geometrically in McMaster (1987b). Originally, the interface provided six algorithms, including Jenks' and nth point, but these were later dropped for space considerations. Associated with each image window is a control window with sliding bars, including those for (1) manipulating the tolerance value, (2) zooming the image, (3) relocating the x-center and (4) y-center, and (5) a ratio value for enhancement. Some algorithms require an additional bar for control of the simplification operator. In Figure 1, each of the four map images of the North American continent has been generalized using the same linear tolerance value: 1.116 units. Note the significantly different results of the generalizations produced by each of the algorithms. At this particular level of simplification, the Lang and Douglas routine retain more of the original detail.

Note the set of bars for the Lang algorithm. The top bar can be slid to adjust the tolerance and, in real time, the image is simplified in the image window. With Lang, the look-ahead value of points, set to 25 in this example, can also be adjusted.

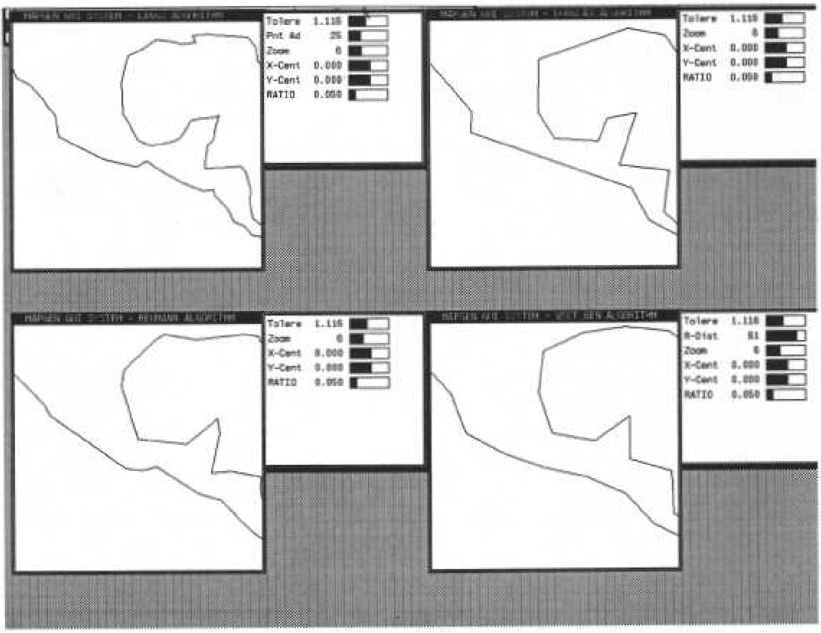

Figure 3. Use of the zoom and recenter functions for Central America.

In Figure 2, the zoom, x-center, and y-center functions have been modified to focus the image--with the same level of simplification produced by the four algorithms--on Central America. For each of the four windows, an additional menu bar may be pulled down. This menu bar allows for (1) overlay of the original line, (2) smoothing using a B-spline, (3) smoothing using a five-point weighted-moving average, (4) a simple enhancement routine, and (5) the selection of a new base map.

Figure 3 depicts a greatly enlarged (6x) view of the Central America coastline, as simplified using the same tolerance. Using the zoom and recentering functions, the user may focus in on various "problem areas". In this instance, note that Lang algorithm retains more of the original detail along the coastline at the same tolerance-level, but requires more coordinate information.

Figure 4 illustrates the overlay of the original digitized feature. This menu bar allows for the original feature to be toggled on and off for each of the windows. Figure 5 depicts a B-spline smoothing of the lines in Figure 3. Again, the B-spline can be toggled on and off. As depicted in Figure 2, the menu bar allows for an animation. Here, the user may continuously simplify a line from most to least complex in real time. An option to stop the animation at any point is now being implemented.

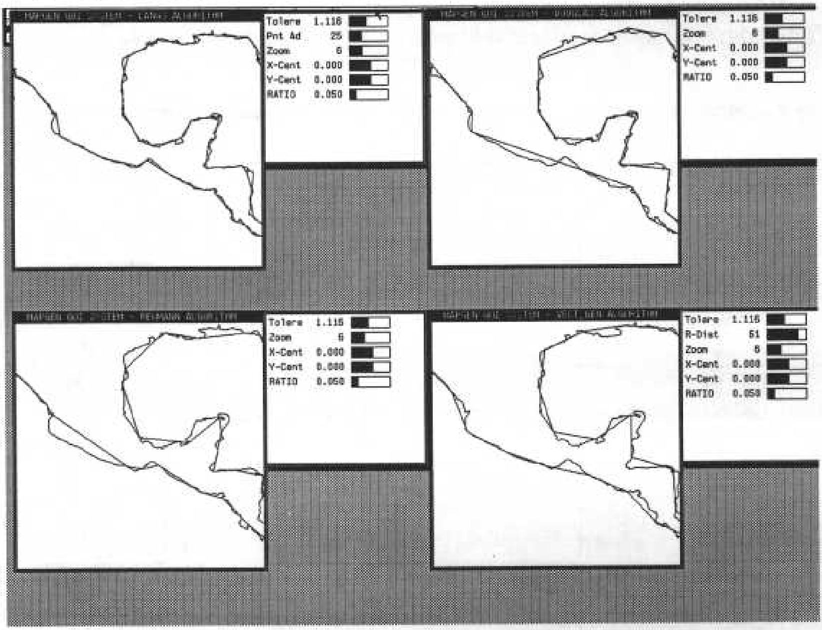

Figure 4. Overlay of the original line.

SUMMARY

This paper has reported on the development of a graphical user interface for map generalization. Such a conceptual structure was suggested in McMaster and Mark (1991). The initial work has focused on (1) the overall layout and functionality of the interface and (2) the application of simplification and smoothing operators. The user is also given the capability to overlay the original digitized feature, rescale and recenter the feature, and animate the simplification. Current work involves adding geometric measures, such as angular and areal modification, and enhancement and displacement operators.

The ultimate goal for the interface is for the acquisition of procedural knowledge on generalization. Given a map image, for instance, in what sequence would an operator apply the simplification, smoothing, and displacement operators? In order to extract such knowledge, it will be necessary to create user logs within the interface that allow for the careful recording of operator application, sequencing, and the selection of tolerance values. Such a graphical user interface, when complete, will enable cartographers to begin the process of procedural knowledge acquisition for the purposes of building knowledge bases.

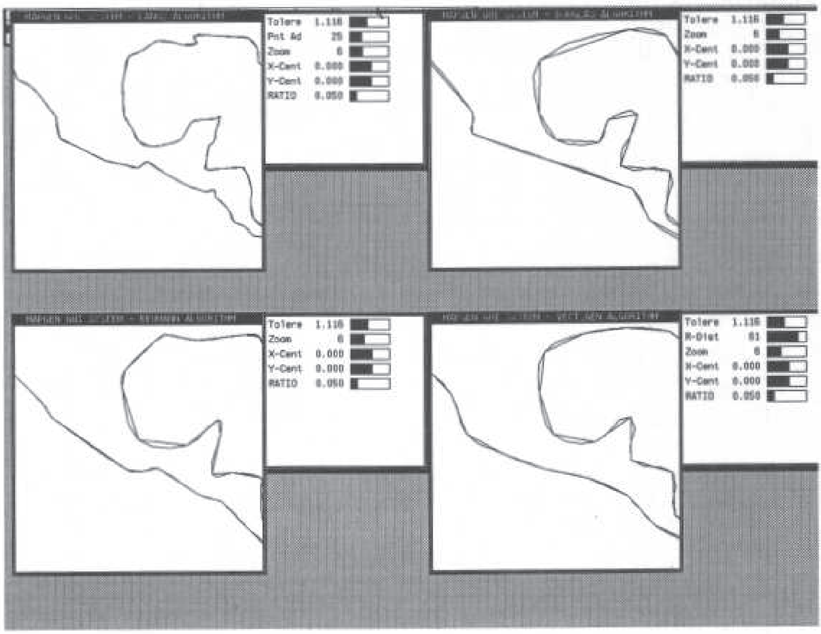

Figure 5. B-spline smoothing of Figure 3, with original simplification.

Bibliography

Armstrong, Marc P. 1991. "Knowledge Classification and Organization," in Map Generalization: Making Rules for Knowledge Representation. Barbara Buttenfield and Robert B. McMaster, (eds). (United Kingdom: Longman), pp. 86-102.

Buttenfield, Barbara P. and Robert B. McMaster. 1991. Map Generalization: Making Rules for Knowledge Representation. (United Kingdom: Longman)

Buttenfield, Barbara P. 1989. "Scale-Dependence and Self-Similarity in Cartographic Lines." Cartographica, Vol. 26, No. 1, pp. 79-100.

Mark, David S. 1991. "Object Modelling and Phenomenon-Based Generalization," in Map Generalization: Making Rules for Knowledge Representation. Barbara Buttenfield and Robert B. McMaster, (eds). (United Kingdom: Longman), pp. 103-118.

McGraw, Karen L.and Karen Harbison-Briggs 1989. Knowledge Acquisition: Principles and Guidelines. (Englewood Cliffs, N.J.: Prentice Hall).

McMaster, Robert B. 1991. "Conceptual Frameworks for Geographical Knowledge," in Map Generalization: Making Rules for Knowledge Representation. Barbara Buttenfield and Robert B. McMaster (eds). (United Kingdom: Longman), pp. 21-39.

McMaster, Robert B. and David M. Mark 1991. "The Design of a Graphical User Interface for Knowledge Acquisition in Cartographic Generalization." Proceedings GIS/LIS'91. Atlanta, Georgia, pp. 311-320.

McMaster, Robert B. 1989. "The Integration of Simplification and Smoothing Algorithms in Line Generalization." Cartographica, Vol. 26, No. 1, pp. 101-121.

McMaster, Robert B. 1987a. "Automated Line Generalization." Cartographica, Vol. 24, No. 2, pp. 74-111.

McMaster, Robert B. 1987b. "The Geometric Properties of Numerical Generalization." Geographical Analysis, Vol. 19, No. 4, pp. 330-346.

Tulving, E. 1972. "Episodic and Semantic Memory." In E. Tulving and W. Donaldson, eds. Organization of Memory. (New York: Academic Press).