
Geographic Information Services

ON THE RESULTS OF A WORKSHOP ON GENERIC DATA MODEL FOR LINEAR REFERENCING SYSTEMS

Authored by
Workshop Participants with Alan Vonderohe
(Principal Investigator Supported by The National Cooperative Highway Research Program Project)

PREFACE

The first draft of this report was prepared from notes and other materials developed during the workshop, audio tape recordings of the workshop sessions, and follow-up discussions with workshop participants. All participants were given the opportunity to review the first draft and provide responses. These responses were assimilated to identify both consensus revisions to the first draft and significant points of contention. The responders were informed of these results and asked to provide their opinions on each of the points of contention. The responders were also asked to provide measures of the relative importance of their positions on each of the points of contention (i.e., "critical", "strong preference", "weak preference"). The revised model presented in this second draft is true to the model as developed during the workshop and to the revisions on which there was consensus. Remaining significant points of contention are summarized in Appendix B.

ACKNOWLEDGMENTS

The commitment and effort of the workshop participants is gratefully acknowledged. This report is a result of their work.

Participant	Affiliation
Teresa Adams Hillary Armstrong Mark Bradford Jim Carroll	University of Wisconsin - Madison Sandia National Federal Highway Administration Michigan Department of Transportation

Richard Church	University of California - Santa Barbara
Chih-Lin Chou	University of Wisconsin - Madison
Ron Cihon	Washington Department of Transportation
Don Diget	Michigan Department of Transportation
Charles Dingman	-United States Bureau of the Census
Ken Dueker	Portland State University
Joe Ferreir	Massachusetts Institute of Technology
David Fletcher	GeoDigm
Cecil Goodwin	University of Tennessee
Stephen Gordon	Oak Ridge National Laboratory
Edward Granzow	Urban Analysis Group
Robert Harris	GIS Trans
Charlie Hickman	United States Geological
Dale Honeycutt	ESRI
Wen-Jing Huang	University of Wisconsin - Madison
Peggi Knight	Iowa Department of Transportation
Roy Larson	Minneapolis / St. Paul Metropolitan Council
Simon Lewis	GIS Trans
Frank Lockfeld	Center for Urban Analysis, Santa Clara County
David Loukes	GeoPlan
Bill McCall	Iowa State University
Daniel McHugh	New York City Transit Authority
Allisoun Moore	Maryland State Highway Administration
Tim Nyerges (Facilitator)	University of Washington
Paula Okunieff	Viggen Corporation
Ken Opiel	National Cooperative Highway Research Program
Tom Ries	Wisconsin Department of Transportation
Jay Sandhu	ESRI
Bill Schuman	Intergraph
Susan Scott	SEI Technology
Robert Smith	University of Wisconsin - Madison
Bruce Spear	United States Department of Transportation
Todd Stellhorn	ESRI
Jung-Gon Sung	University of Wisconsin - Madison
James Tucker	GDS Corporation
Allan Vonderohe (PI)	University of Wisconsin - Madison
Kirk Weaver	New Jersey Department of Transportation
Lynn Williams	Intergraph

INTRODUCTION

On August 6, 1994, forty-two transportation professionals, systems developers,and academics came together in a workshop in Milwaukee, Wisconsin with the objective of preparing a draft consensus conceptual data model, at the entity- relationship level, for linear referencing systems. Workshop participants, selected for their expertise in linear referencing systems and modeling, provided a broad representation of local, state, and federal transportation and mapping agencies;

consultants, data providers, and software providers from the private sector; and researchers from national laboratories and universities. Recognizing that it was not feasible to derive a data model that would meet all the needs of all application areas, a generic model was sought that met common needs and formed a core that could be extended as needed in specific application areas. The resulting draft data model, in object modeling form, associates transportation data with multiple cartographic representations and multiple network models through a single linear datum. The datum links the data model to real-world features and provides the referencing space that enables transformations among linear referencing methods, networks, and cartographic representations at various scales. The data model supports a set of fundamental operations that cause data to flow between the database world and the real world. The data model as presented is intended to represent the requirements for a linear referencing model - it is not intended as a specification.

BACKGROUND

Previous work on NCHRP 20-27 led to the recommendation that transportation agencies develop conceptual organizing principles founded upon the notion of location as a data integrator (Vonderohe, et.al., 1993). Another result of NCHRP 20-27 was a suggested technological framework (server net) for support of GIS-T and transportation computing in general. The continuation phase of the research is developing generic functional and data models to complement the technological model. The transportation agencies towards whom the data integration recommendation was directed manage vast stores of linearly referenced data. Any generic data model for GIS-T must include linear referencing components. Linear referencing systems are used in nearly all application areas that are based upon networks, including infrastructure management, transit, freight, intelligent transportation systems, waterway navigation, hydrological analysis, utilities management, and seismological sensing.

The significance of linear referencing methods and systems to transportation applications has been recognized for some time. An early NCHRP publication (Synthesis 21, 1974) made the distinction between methods and systems, classified a number of linear referencing methods, and made recommendations for their improvement. Transportation agencies from time to time have studied the location referencing methods they use and sought to adopt standards for them (Briggs and Chatfield, 1987). During one study, the Michigan Department of Transportation identified 38 location referencing methods in use by the agency. More recently, some state DOTs have developed formal data models for location referencing (Deighton and Blake, 1993; Ries, 1993; Scarponcini, 1994). Some have succeeded in raising the issue to the policy level (Deighton and Blake, 1993). One state DOT (Wisconsin) has recognized "location control management" as a formal business area in its information strategy plan. Early research in geographic information systems in transportation led to identification of the need for, and subsequent development of, dynamic segmentation as a critical function for managing linearly referenced data (Fletcher, 1987; Dueker, 1987; Nyerges and Dueker, 1988; Nyerges, 1990). More recently, the underlying data models that support current implementations of dynamic segmentation have been examined (Dueker, 1992).

A recent executive-level commitment to the concept of a National Spatial Data Infrastructure (Mapping Science Committee, 1993) has spurred interest in the development of standards and common models for data. The Ground Transportation Subcommittee of the Federal Geographic Data Committee has recommended that a linear referencing system be incorporated in standards efforts. The Federal Highway Administration is incorporating linear referencing systems from the states in the Highway Performance Monitoring System (FHWA, 1993). The ITS community is striving for standards and a generic data model for linear referencing. And the GIS-T Pooled Fund Study Team is incorporating a linear referencing system model in Phase B of their research.

At the same time, a number of workers are now addressing various aspects of the data conflation problem, typically associated with attempts to integrate census data tied to TIGER/line files with attribute data tied to other representations of the same street network that do not coincide with TIGER (e.g., Brace and Peterson, 1994; Clark and Bain, 1994; Peterman, 1994).

Given all these activities and interests in data sharing and integration, the need for a common, generic data model for linear referencing systems is compelling. The workshop was convened to address this need.

WORKSHOP STRUCTURE

All participants received a package of materials in advance of the workshop. The package included information on the research project, the objectives of the workshop, a preliminary program, and a set of pre-conference papers whose authors had been asked to give presentations. The final program for the workshop appears in Appendix A.

A modified form of the "Technology of Participation" method was used to structure the workshop, with Professor Tim Nyerges of the University of Washington serving as facilitator. Following introductory remarks concerning workshop methods, twelve invited technical presentations were made concerning various aspects of linear referencing systems data modeling. During the presentations, participants identified issues and wrote them on large cards which were posted during breaks. Following the presentations, participants were encouraged to identify additional issues and post them. The issues were then clustered into topic areas by the group as a whole. The clustered issues were then synthesized and gaps were identified. It was decided that the topic areas "Terms and Definitions" and "Scoping" were most critical to development of the model, and a discussion of these topics by the group as a whole ensued. Considerable progress was made and, given the time constraints of the workshop, it was decided to forego the planned breakout sessions and continue as a group of the whole to address the two critical topic areas. This led to the ultimate collective development of an object-based data model which appeared in an earlier draft of this report. The data model presented below includes consensus revisions based on responses to the first draft provided by workshop participants.

TOPIC AREAS AND ISSUES

The critical topic areas and issues identified during the workshop were:

1. **Terms and Definitions** - Example issues: Standardized, unambiguous, definitions must be developed for common terminology. Terms such as "traversal" should be used instead

of "route", "path", or "trip", all of which might be subclasses of "traversal". Terms such as "anchor point" and "anchor section" should be used because "control point" and "control section" have other meanings. The term "distance" can mean "odometer distance", "posted distance", or "cogo distance".

2. **Scoping** - Example issues: Are geographic, spatial, cartographic, and temporal objects modeled in the same domain? What is the conceptual extent of the term "linear referencing system?" Is there a set of core requirements for a host of applications? We must account for vector, non-planar models used in commercial GIS. What are the primary functions that must be supported? Is linear referencing broad enough to address workshop goals?
3. **Schema Constructs** - Example issues: What are the primary building blocks of linear referencing method? We must identify the differences between structural data model requirements and functional views. Are linear referencing system components hierarchical objects or interdependent and relational? We must relate topology, one-dimensional, two-dimensional, and three-dimensional space. Can one conceptual data model accommodate feature-based systems, planar graph-based systems, and non-planar graph-based systems?
4. **Transformation** - Example issues: How do we uniquely and unambiguously identify locations for transforming that information between dissimilar IS? What are the basic transformations between linear referencing methods? How do we link multiple linear referencing methods together for data integration in the network domain? What are the rules for aggregation to support more generalized reporting?
5. **Multi-Dimensionality** - Example issues: To address conflation, 1) should we develop unique identifiers for certain features; 2) should we standardize on topology or geometry (x,y,z)? Do we need a GPS/GIS linkage? Linear referencing systems must be linked to higher-dimensional systems, including those that model time. Spatial proximity is not a surrogate for network topology.
6. **Methods and Coding** - Example issues: How are the "lowest common denominator" sites identified? What is the appropriate datum structure? What do we mean by the "location" of a bridge - is it the center, one end, the other end? The model must be able to handle very large databases. What are the rules for establishing linear referencing method starting and ending points? What is the best method for referencing ramps? What geographic features are assigned external identifiers?
7. **Data Integrity** - Example issues: How do we ensure data integrity if we have multiple linear referencing methods? What are the referential integrity rules? Can versions be coordinated by adding version numbers to unique identifiers? What are the implications of alignment changes?
8. **Institutional Policy** - Example issues: Do users need to understand linear referencing systems? What policies and procedures are required for a linear referencing system? What cost constraints are associated with a linear referencing system?

KEY CONCEPTS

NCHRP Synthesis 21 on Highway Location Reference Methods (NCHRP, 1974) contains two fundamental premises adopted by the workshop participants:

1. There is a clear distinction between linear referencing methods and linear referencing systems: "A highway location reference system is a set of office and field procedures that includes a highway location reference method. The latter is a way to identify a specific location with respect to a known point." The workshop participants included within the concept of system a means for transformation among various methods. Thus, "milepoint", "reference post", and "engineering stationing" are methods. The policies, records, and procedures that relate these methods are the system.
2. The location of any unknown point along a linear feature can be determined by specifying the direction and distance from any known point to the unknown point. All linear referencing methods are based on this. The workshop participants concluded that the premise is true for two and three dimensions also. Each additional dimension removes a constraint on direction.

It was concluded that, given the time limitations of the workshop, we could not develop a data model that supported all the needs of all possible application areas. What was being sought instead was a core, generic data model that could be extended to meet specific needs of various applications. It was also clear that, because of the time constraint, we could not develop a robust and elegant specification. Rather, we were pursuing a model that addressed the requirements for linear referencing systems. p Multiplicity became a theme. There is a central need to integrate not only multiple scales of geography and cartographic (coordinate-based) data from multiple sources, but also multiple network models, each of them necessary for particular applications. It was decided that the data model must support certain fundamental operations:

1. **Locate** - Establishment of the location of an unknown point in the field by reference to objects in the "real world".
2. **Position** - Translation of a real-world location into a database location.
3. **Place** -Translation of a database location into a real-world location (the inverse of the "position" operation).
4. **Transform** - Conversion between various linear referencing methods, represented by database locations; between various cartographic representations; and between methods and cartographic representations.

It was expected that if these operations are supported, then the model should support higher-level operations such as those associated with GIS (e.g., overlay, connectivity, proximity) and those associated with network analysis (e.g., pathfinding, routing, location, and allocation).

THE DATA MODEL

Overview

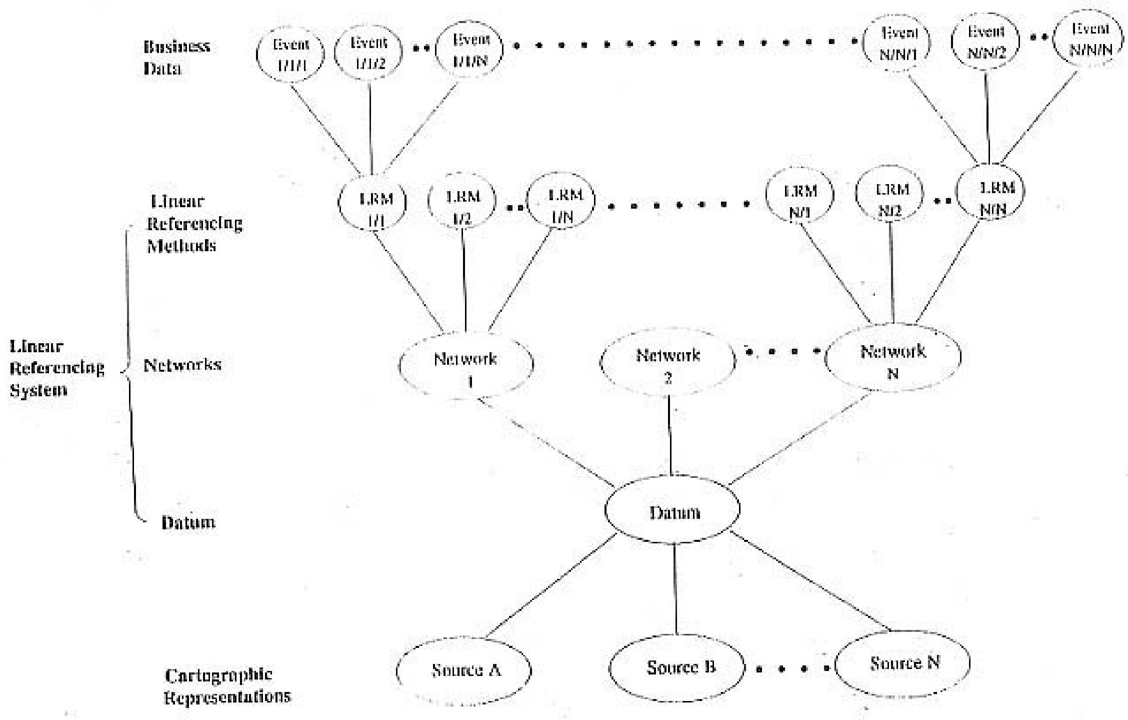

Figure 1

Figure 1 presents a conceptual overview of the data model. The central notion is that of a linear datum that supports multiple cartographic representations (at any scale) and multiple network models (for various application areas). The datum provides the fundamental referencing space for transformations among various linear referencing methods, network models, and cartographic representations. It also links the model to the "real world" through attributes that describe its location and spatial characteristics in real-world references and measures.

Cartographic representations provide coordinate references, the basis for to-scale visualization of the model, and linkages to two-dimensional and three-dimensional GIS databases. Network models provide the topological framework for pathfinding, routing, location/allocation, transshipment, and flow operations.

A number of linear referencing methods might be associated with each network model. These methods might be those associated with infrastructure management, such as reference post, milepoint, or engineering stationing. They might also be those associated with navigation (requiring recognizable landmarks or navigation aids), or with transit (timing points), or with a host of other application areas. Each linear referencing method ties a collection of business data to the model, thereby providing a means for integration of those data. The linear referencing system can be thought of as all those components of the model that provide methods for location referencing of business data, transformations among those methods, and linkage of the model to the "real world" and its cartographic representations.

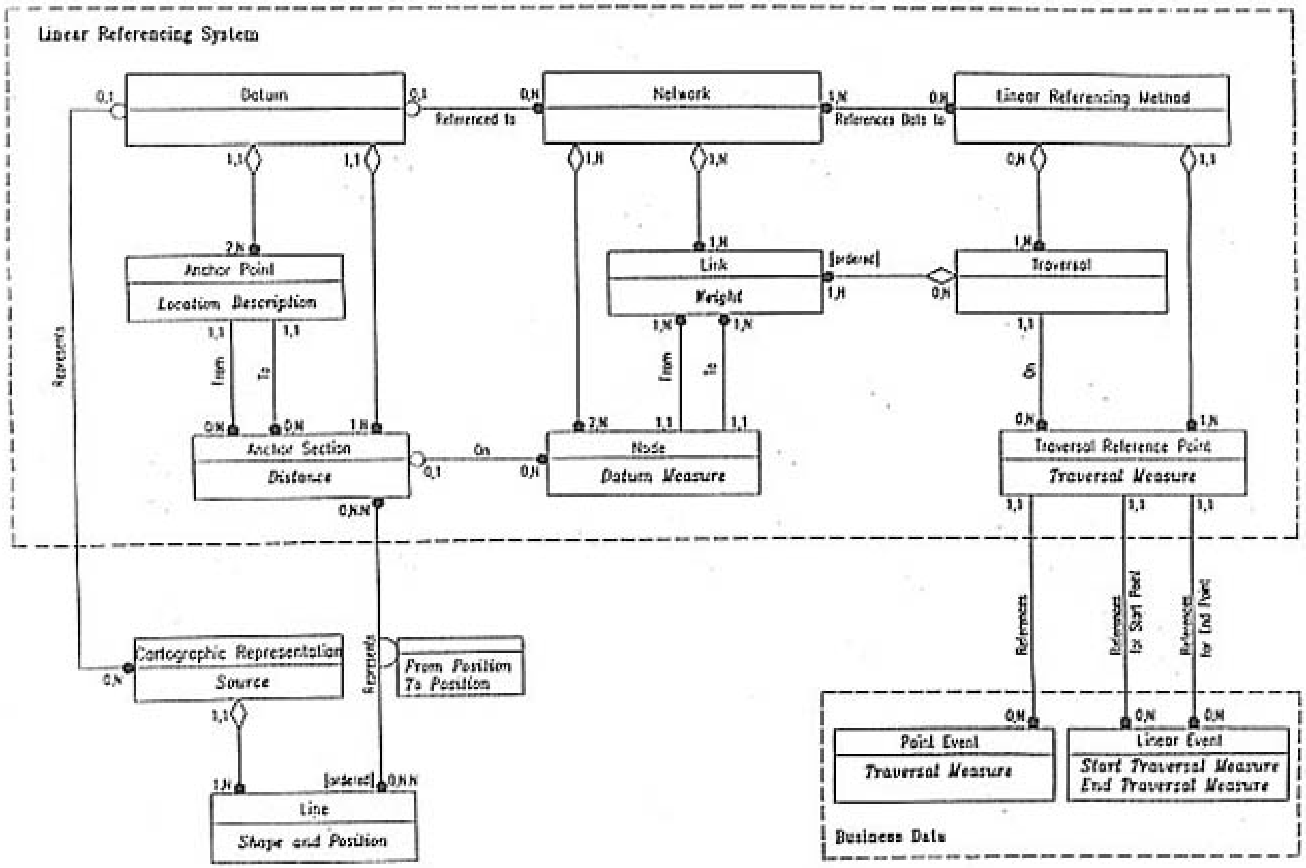

Figure 2

The object model diagram appears in Figure 2. The diagramming method is slightly modified from that of Rumbaugh, et. al., (1991). The modification being that low and high cardinalities are shown with Arabic numerals on both ends of all associations. Standard notation includes the name of an object class in the upper half of a rectangular box; attributes of the class in the lower half of the same box; an association between two object classes denoted by a line connecting the boxes; an association descriptor written on the connecting line for all associations except aggregations; attributes of an association appearing in the lower half of a rectangular box tied to the association's connecting line by a half loop; "many" cardinality indicated by a filled circle; "zero or one" cardinality indicated by an empty circle; "exactly one" cardinality indicated by lack of a circle; and aggregation indicated by a diamond symbol.

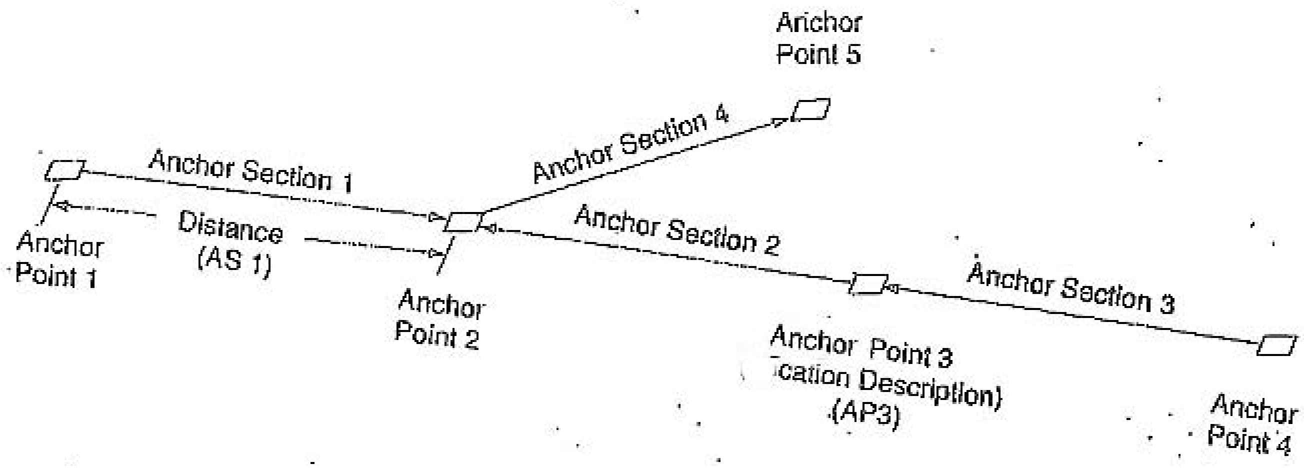

Figure 3

Object Classes and Their Attributes Linear Datum - The collection of objects which serve as the basis for locating the linear referencing system in the real world (see Figure 3). The datum relates the database representation to the real world and provides the domain for transformations among linear referencing methods and among cartographic representations. The datum consists of a connected set of anchor sections that have anchor points at their junctions and termini. No attributes are assigned to datums. Anchor Point - A zero-dimensional location that can be uniquely identified in the real-world in such a way that its position can be determined and recovered in the field. Each anchor point has a "location description" attribute which provides the information necessary for determining and recovering the anchor point's position in the field. Forms of location descriptions can vary and can be quantitative or descriptive or both. Example values include: the intersection of the centerlines of Oak Street and Maple Street; and 1.2 miles south of the Post Office on Route 9. Anchor points can be understood as one-dimensional control points, in that they serve the same purpose as geodetic control points in two and three dimensions. That is, they are the fundamental objects to which all other objects are directly or indirectly tied. Anchor Section - A continuous, directed, non-branching linear feature, connecting two anchor points, whose real-world length (in distance metrics), can be determined in the field. Anchor sections are directed by specifying a "from" anchor point and a "to" anchor point. Anchor sections have a "distance" attribute which is the length of the anchor section measured on the ground. Values are expressed in units of linear distance measure (e.g., kilometers). Anchor sections provide the fundamental referencing space. The collection of anchor sections in a given linear referencing system is analogous to the ellipsoid surface in a geodetic datum or the map projection surface in a two-dimensional Cartesian referencing system.

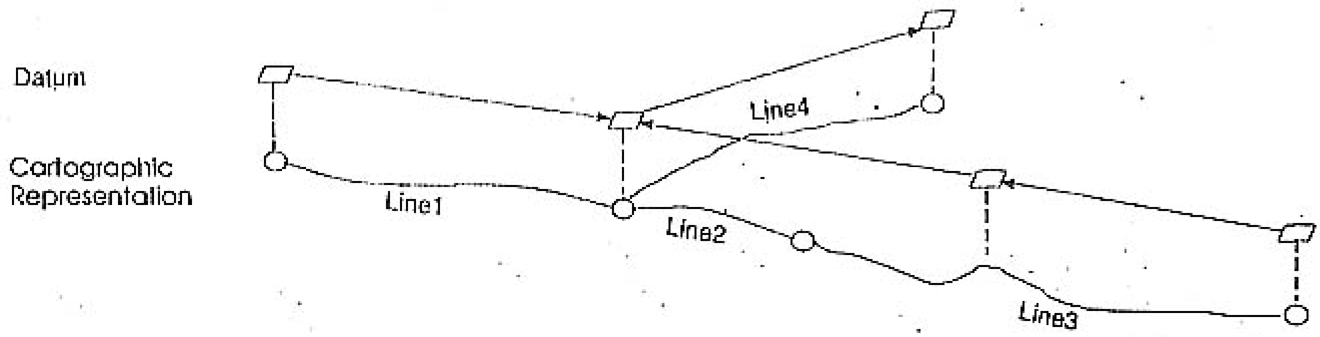

Figure 4

Cartographic Representation - A set of lines that can be mapped to a linear datum (see Figure 4). The set of lines can be either fully or partially connected. That is, the set can consist of groups that are externally unconnected but internally connected. Cartographic representations have a "source" attribute that denotes the source (scale and lineage) of the object. Scale values are expressed as ratios or as equations that relate distances measured on the source form of the cartographic representation to distances measured on the ground. Cartographic representations provide coordinate references; the basis for to-scale visualization of other components of the linear referencing system model; and linkages to extended topological, vector- based GIS data models. Line - "A generic term for a one-dimensional object." (SDTS, 1992). SDTS goes on to define five specific kinds of lines: 1) line segment, 2) string, 3) arc, 4) link, 5) chain. A line, as defined herein, can be any of these except a link. This is because lines, as defined herein, have a "shape and position" attribute. According to SDTS, a line segment is a direct line between two points, a string is a connected nonbranching sequence of line segments, an arc is a locus of points that forms a curve that is defined by a mathematical expression, and a chain is a directed nonbranching sequence of nonintersecting line segments and (or) arcs bounded by nodes, not necessarily distinct, at each end. Shape and position are provided either by the x,y,z coordinates of points associated with line segments or by the mathematical expressions associated with arcs. Possibilities for types of coordinate values include Cartesian and geographic (lat / long/ elev). Possibilities for mathematical expressions include splines and polynomials. Network - A graph without two-dimensional objects or chains. If projected onto a two-dimensional surface, a network can have either more than one node at a point and (or) intersecting links without corresponding nodes.

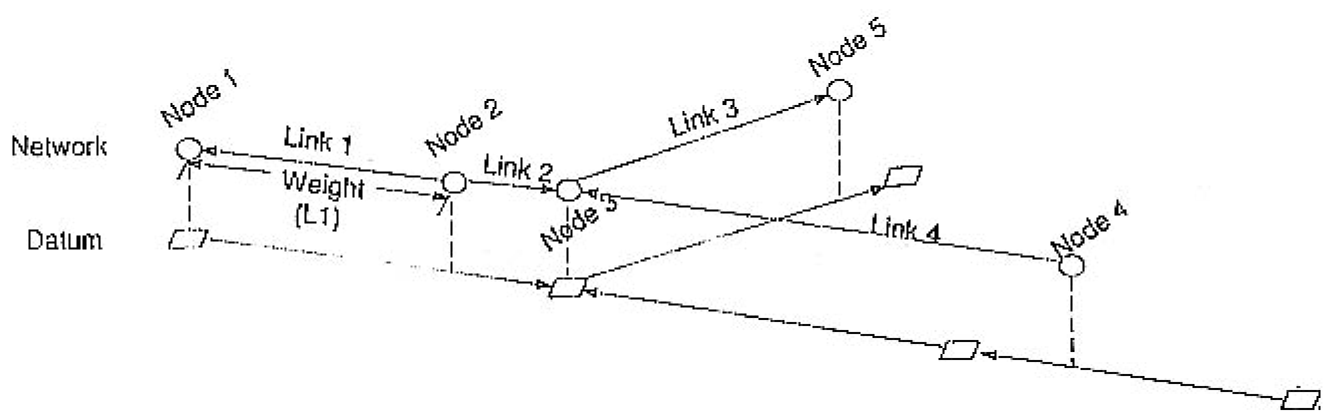

Figure 5

Note: This a modification of the definition provided by the Spatial Data Transfer Standard. Modification is necessary to exclude chains. Within the context of the linear referencing system data model, a network is an aggregate of nodes and links and is, thus, a purely topological object (see Figure 5). The network component of the model provides the basis for analytical operations such as pathfinding and flow. No attributes are assigned to networks.

Node - A zero-dimensional object that is a topological junction of two or more links, or an end point of a link.

Note: This is a modification of the definition provided by the Spatial Data Transfer Standard. Modification is necessary to remove reference to chains. In this data model, nodes do not have coordinates. They are located geometrically by reference to the datum.

Each node has a "datum measure" attribute which is used to locate it on an anchor section. "Datum measure" is an offset measured from the "from" anchor point of the anchor section. "Datum measure" is expressed as a distance measure in the same units as the "distance" attribute of the associated anchor section.

Link - A topological connection between two ordered nodes.

Note: This is a modification of the definition provided by the Spatial Data Transfer Standard. Modification is necessary to require directionality. Each link has a "weight" attribute that is a linear measure of impedance associated with travel along the link. Weights are often expressed in distance measure, but they could be in other linear metrics such as travel time or cost.

Linear Referencing Method - A mechanism for finding and stating the location of an unknown point along a network by referencing it to a known point.

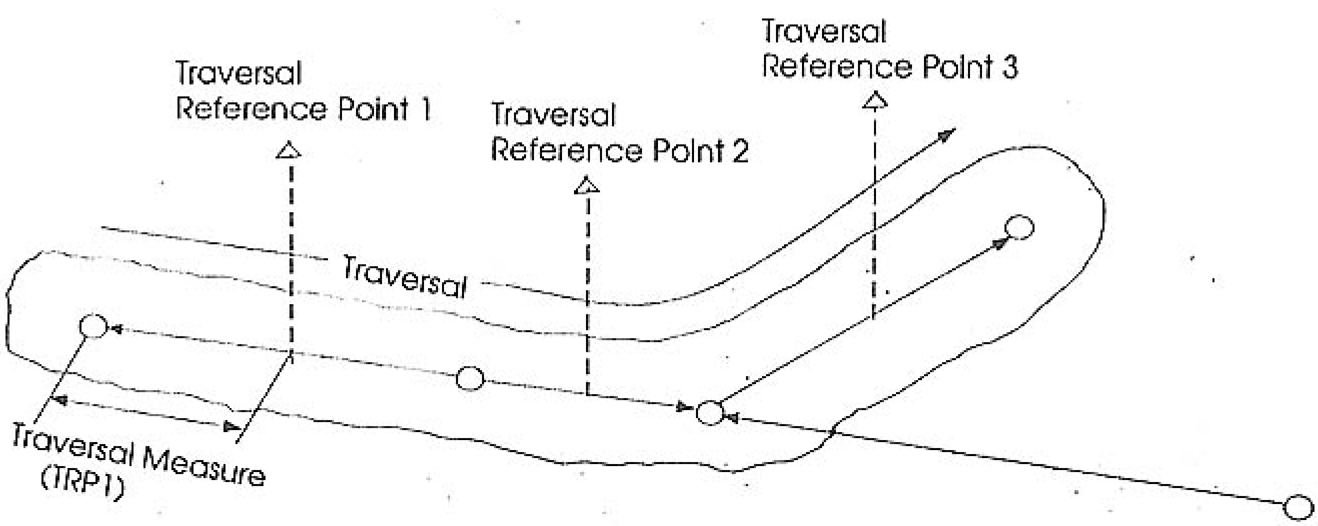

Figure 6

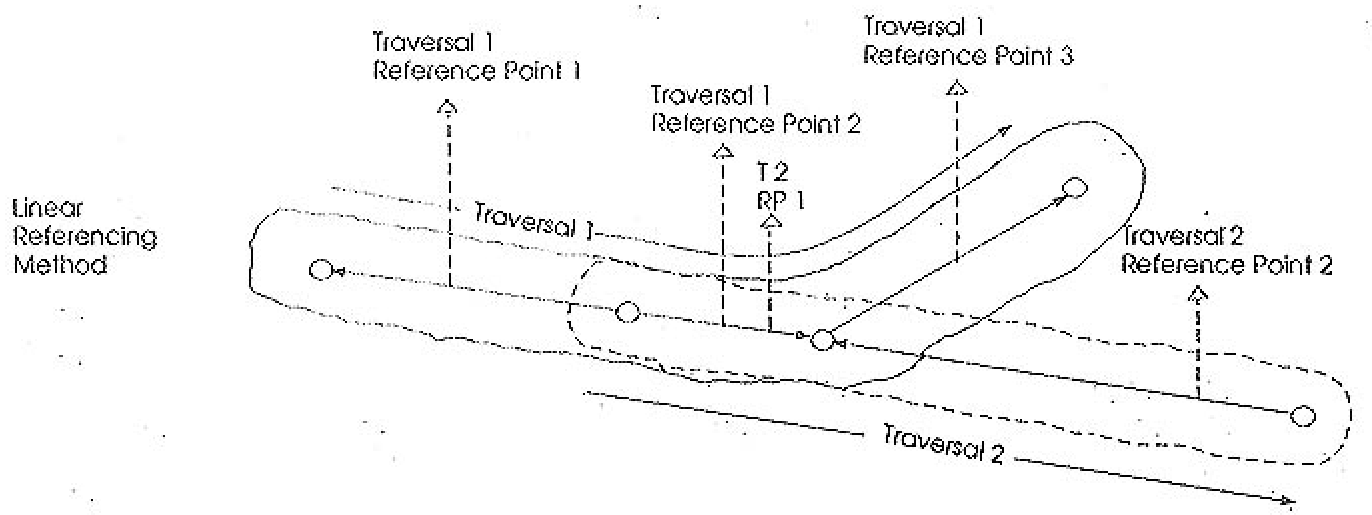

Figure 7

Note: This is a modification of the definition provided by Deighton and Blake (1993). There are many kinds of linear referencing methods (e.g., milepoint, reference post, and engineering stationing). All linear referencing methods consist of traversals and associated traversal reference points, that together provide a set of known points, a metric, and a direction for referencing the locations of unknown points (see Figures 6 and 7). No attributes are assigned to linear referencing methods.

Traversal - An ordered and directed, but not necessarily connected, set of whole links. Coding conventions are required for establishing traversal directionality (in contrast to link directionality) and for specifying non-connected traversals. No attributes are assigned to traversals.

Note: It is the intent of the workshop participants to allow dendritic traversals but specific implications of this for the model have not been investigated.

Traversal Reference Point - A zero-dimensional location along a traversal that is used to reference events along the traversal. Each traversal reference point has a "traversal measure" attribute which is used to locate it along the traversal. "Traversal measure" is an offset measured from the initial node in the traversal to the traversal reference point. It is in the same units as the "weight" attribute of the links in the traversal.

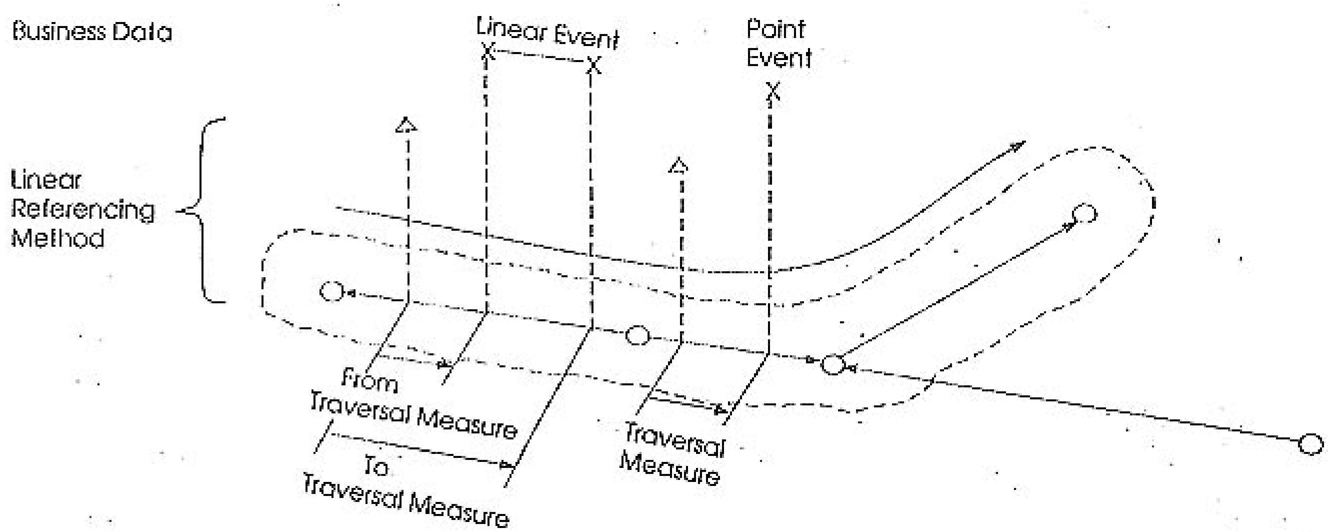

Figure 8

Point Event - A zero-dimensional phenomenon, that occurs along a traversal and is described in terms of its attributes in the extended database (see Figure 8). Examples of point events include signs and accidents . Each point event in the linear referencing system data model has an "traversal measure" attribute. "Traversal measure" is an offset measured from the referenced traversal reference point to the point event. Point event traversal measures are in the same units as the traversal measures of the traversal reference points that they reference. A positive point event traversal measure expresses measurement in the direction of the traversal. A negative point event traversal measure expresses measurement against the direction of the traversal. Point events will typically have additional attributes in the extended database.

Linear Event - A one-dimensional phenomenon that occurs along a traversal and is described in terms of its attributes in the extended database (see Figure 8). Examples of linear events include pavement types, speed zones, and construction projects. Each linear event in the linear referencing system data model has "start traversal measure" and "end traversal measure" attributes that locate the linear event along the traversal. The traversal measures are offsets measured from the traversal reference points that they individually reference. Linear event traversal measures are in the same units as the traversal measures of the traversal reference points that they reference. Rules for direction of measurement are identical to those of point event traversal measures. Linear events will typically have additional attributes in the extended database.

Associations and Their Attributes An anchor section goes from an anchor point -to-an anchor point. Each anchor section is associated with two, not necessarily distinct, anchor points in this way. Any number (0,N) of anchor sections can go from or to a given anchor point. A link goes from a node to a node. Each link is associated with two, not necessarily distinct, nodes in this way. At least one link must go from or to a given node.

A traversal reference point must be on one and only one traversal. A traversal can have any number (0,N) of traversal reference points on it. A point event must reference one and only one traversal reference point. A traversal reference point can be referenced by any number (0,N) of point events. A linear event must 1) reference its start point to one and only one traversal reference point and 2) reference its end point to one and only one traversal reference point. The traversal reference points that are referenced for the start and end of a linear event do not necessarily have to be distinct. A traversal reference point can be referenced by any number (0,N) of linear events. A number of aggregation associations appear in the model. A datum is composed of at least two, not necessarily distinct, anchor points and at least one anchor section. A cartographic representation is composed of at least one line. A network is composed of at least one link and at least two, not necessarily distinct, nodes. A linear referencing method is composed of at least one traversal and at least one traversal reference point. A traversal is composed of at least one link. The links are ordered in this association, thereby, giving direction to the traversal (see Figure 6). A link can be a component of many traversals (see Figure 7). A cartographic representation can represent zero or one datum.

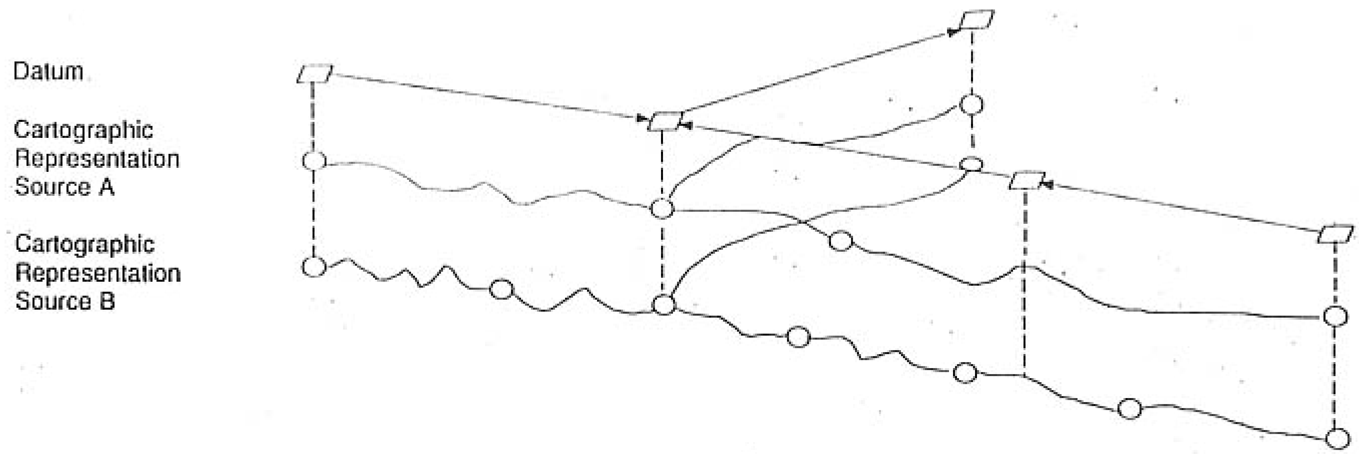

Figure 9

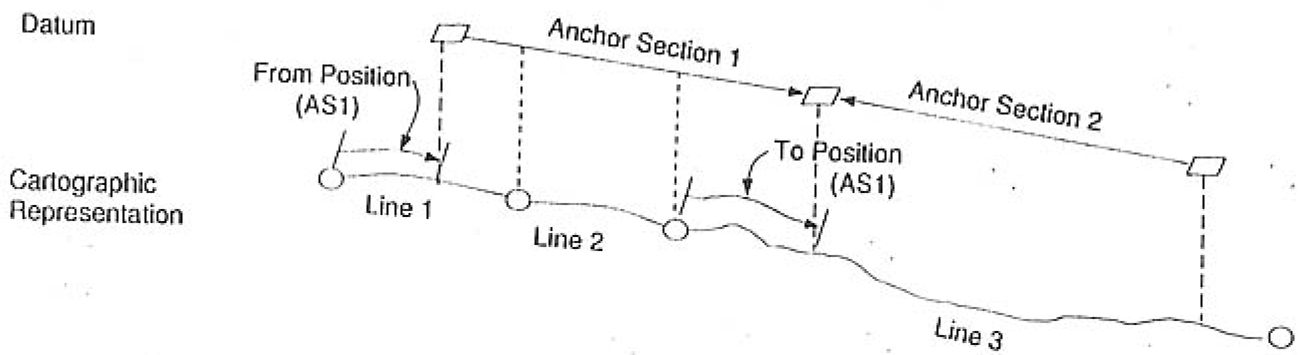

Figure 10

A datum can be represented by any number (0,N) of cartographic representations (see Figure 9). A line can represent any number (0,N.N) of anchor sections, including as many as two partial anchor sections (see Figure 10). An anchor section can be represented by any number (0,N.N) of lines, including as many as two partial lines. The lines are ordered and the association is assigned attributes to resolve the many-to-many and partial- object mappings. The association "represents," between line and anchor section, has from position" and "to position" attributes. "From position" specifies the percentage of the first line in the list to be used as an offset from that line's start point to map the beginning of the anchor section onto that line. "To position" specifies the percentage of the last line in the list to be used as an offset from that line's start point to map the end of the anchor section onto that line.

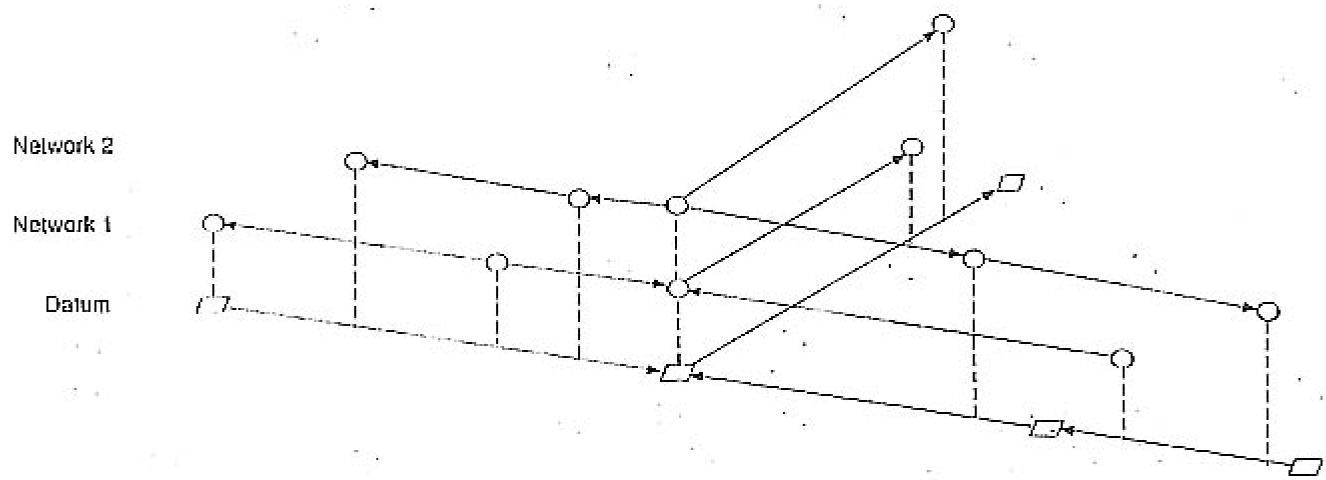

Figure 11

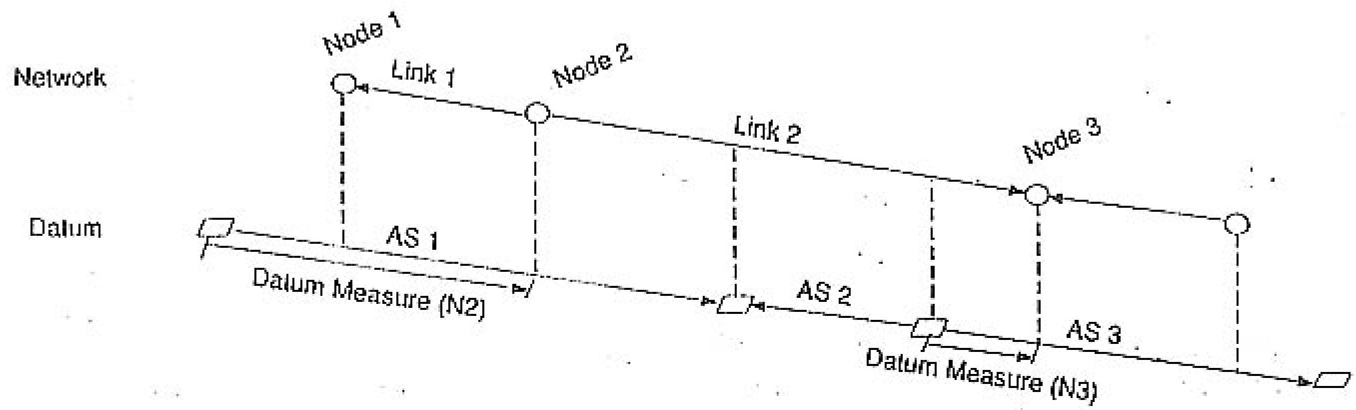

Figure 12

A network is referenced to zero or one datum. A datum can have any number (0,N) of networks referenced to it (see Figure 11). A node can be on zero or one anchor section. An anchor section can have any number (0,N) of nodes on it (see Figure 12). A linear referencing method must reference data to at least one network.

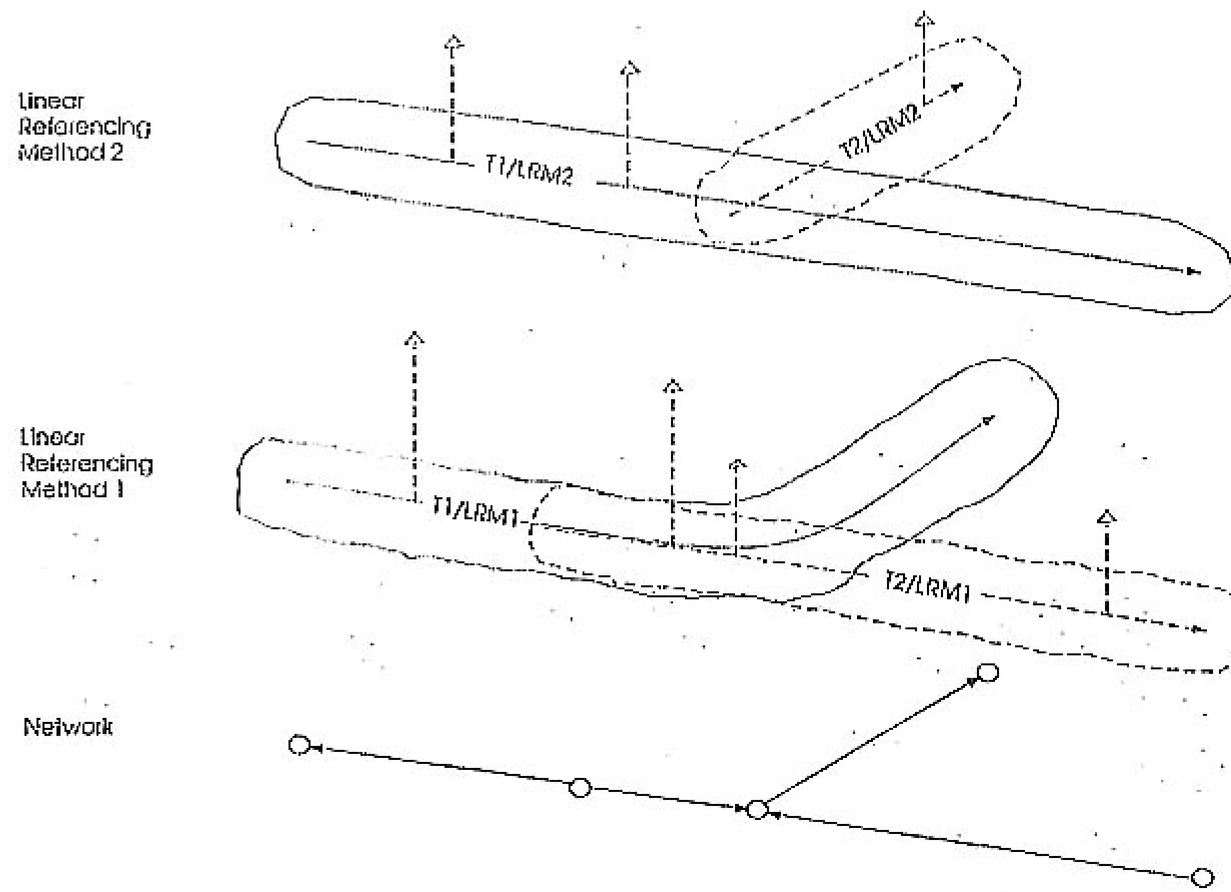

Figure 13

A network can have any number (0,N) of linear referencing methods that reference data to it (see Figure 13).

Clarification of Directionality

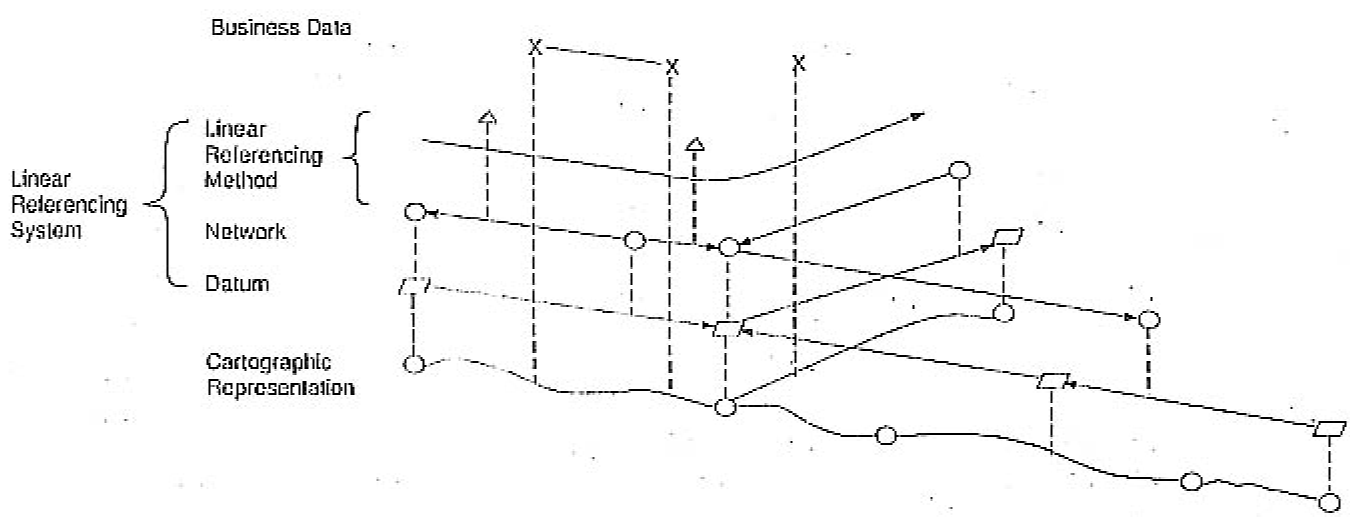

Figure 14

Figure 14 depicts the components of the model and a mapping that produces displayable coordinates for business data. Three object classes in the model have direction associated with them.

Each anchor section has direction, by definition. The direction of an anchor section could be initially established in any way, perhaps as a matter of convenience in the field. Anchor section direction is used to map lines and nodes onto anchor sections and to map anchor sections onto links.

Each link has direction because the order of the nodes it connects is specified. Link direction might be established according to the application the links will support. Link direction is used to map links onto anchor sections and to support network analysis.

Each traversal has direction, by definition. Traversal direction might by established by institutional factors (e.g., STH 10 South) or by analysis (e.g., pathfinding). Traversal direction is used to order traversal reference points and to map point events and linear events onto traversals.

Within the model, these three kinds of direction are reconciled by ordering and specifying "from / to" associations. Anchor section direction is established by specifying "from" and "to" anchor points. Link direction is established by specifying "from" and "to" nodes, then reconciled with anchor section direction by mapping the ordered nodes onto anchor sections. Link direction is reconciled with traversal direction by ordering the links that compose the traversal.

Supported Operations Using information from an implementation of this model, one can locate one's self in the field by first identifying which anchor section they are on (from a hardcopy cartographic representation or a listing of anchor sections, their anchor points, and the location descriptions of the anchor points). A measurement will then be made, along the linear facility, from the "from" anchor point, towards the "to" anchor point, to the unknown point, thus establishing its location.

A phenomenon (say, an accident) in the field can be positioned as a point event in the database merely by creating a record that identifies the traversal reference point to which the accident is referenced and specifies the traversal measure (+ or -) from the traversal reference point to the accident. The record will usually also contain values for other attributes of the point event (accident).

A point event in the database can be placed in the field by first using the traversal measure of the point event and the traversal measure of the associated traversal reference point to compute a cumulative offset from the initial node in the traversal to the point event. Then compute cumulative offsets from the initial node in the traversal to successive nodes in the traversal (using link weights) until a node is reached whose cumulative offset is greater than the cumulative offset of the point event. This node and the immediately previous node determine the link on which the point event lies. From the cumulative offsets, compute the offset of the point event from the "from" node of the link as a percentage of the weight of the link. Using the node/anchor section associations and the topology of the anchor sections, determine the anchor sections and/or portions thereof that map to the link. Determine the length of the link from the distance attributes of the anchor sections. Compute the distance from the mapped location of the link's "from" node to the point event using the percentage offset and the length of the link in distance units. Determine the anchor section that contains the point event and compute the distance from that anchor section's "from" anchor point to the point event. Produce a hardcopy cartographic representation and print out the location description of the "from" anchor point. Discover the "from" anchor point in the field and, using the cartographic representation for direction reference, lay off the distance to the phenomenon represented by the point event. A milepoint reference can be transformed into a project / engineering stationing reference on the same traversal by first comparing each project's "beginning of job" or "0+00" traversal measure to the traversal measure of the milepoint. If the milepoint is on any project, it will be on the project whose 0+00 has the largest traversal measure that is less than the traversal measure of the milepoint (here we assume that all project directions are the same as that of the traversal). For the selected project, determine if the "end-of-job" traversal measure is greater than the traversal measure of the milepoint. If so, the milepoint is on the project. Compute the offset from 0+00 to the milepoint. Express the result as engineering stationing.

Possible Derived Attributes

In any implementation of a data model, the trade-off between storing information and computing information when needed must be considered. Factors in this consideration concern efficiency and are application-specific. They depend upon the nature of the queries that are expected to be most frequently posed.

Attributes in the current data model most efficiently support queries that require mappings of networks or their components onto the datum and mappings of the datum or its components onto cartographic representations, as shown in Figure 14. Inverse mappings require computation. For example, if we are interested in mapping an anchor section onto a link or a connected set of links, we must first identify all candidate nodes through the "on" association. For each of the candidate nodes, the set of nodes to which it is directly connected must be determined through the node/link association. For each candidate node, the set of connected nodes must be compared to the candidate set. Any node not in both sets helps form a link that is partially on the anchor section. The remainder of the partial link at each end of the anchor section must then be mapped onto the adjacent anchor section. The offset for each of the two anchor points can then be determined as a percentage of the weight of the appropriate link. If such inverse mappings are expected frequently, it might be appropriate to compute anchor point offsets on links only once and store them. For some applications it might even be appropriate to compute all possible mappings through the model once and store them.

Another attribute that is used frequently in network analysis applications is the cumulative weight of a traversal - that is, the sum of the weights of its links. This is also an attribute that could be computed once and stored.

ADDRESSING THE ISSUES

A brief synopsis of how the model addresses issues identified by workshop participants follows:

1. **Terms and Definitions** - A data dictionary is included with the model. Definitions from the literature, particularly the Spatial Data Transfer Standard were used whenever possible. Terms having alternative meanings or interpretations in the transportation and GIS communities were avoided.
2. **Scoping** - The scope of the model is linear referencing. Primary spatial aspects of the model are topology in the datum and networks, distance measures in the datum, weights in the networks, and offsets to locate zero-dimensional objects. The model includes cartographic representations in higher dimensions to provide coordinate references, to-scale visualization, and linkages to extended GIS data models. Issues such as polygonal representation of facilities at large scales and left/right offsets to off-facility features are not directly addressed by the model, but could be treated in extensions. Four fundamental operations are supported. Many other higher-level operations are also certainly supported. Temporal dimensionality remains unaddressed, except for a few ideas on versioning (see item 7, below). The development of robust space/time abstractions is an open research area.
3. **Schema Constructs** - The model was developed to be generic and to include the core requirements of as many application areas as possible. The model is presented as an object model. Considering the sparseness of behavior included in the model, a relational form should be readily derivable. No hierarchies of classes are included in the model as presented. The model has components that are non-planar (networks) and components that are planar in many implementations (cartographic representations). Development of a feature-based model from the object model at hand should be investigated.
4. **Transformation** - Transformation is one of the four fundamental operations demonstrated to be supported by the model. Transformations are possible between linear referencing methods, between cartographic representations, and between linear referencing methods and cartographic representations. Aggregation of events to support more generalized reporting should be supported by the model, although no demonstration is provided.
5. **Multi-Dimensionality** - The datum is the fundamental reference space for transformation, for linking to higher dimensions, and for solving the conflation problem. It is a topological object that includes distance measures. Proximity operations can take place at the datum level if distance is the desired metric, at the network level if link weight is the desired metric, or at the cartographic representation level if two- or three-dimensional analysis is desired. Use of the Global Positioning System during data collection essentially provides a new cartographic representation of a linear facility for each pass in the field. Such data can be linked to the model by associating receiver positions with anchor points as they are encountered in the field.
6. **Methods and Coding** - The "lowest common denominator" is the datum. It consists of anchor points and anchor sections. Anchor points are zero-dimensional objects that must be unambiguously identifiable in the field. Therefore, a bridge cannot serve as an anchor point. The center point of a specified end of the deck of a specified bridge could serve as an anchor point. Starting and ending points of traversals are the appropriately ordered nodes of the first and last links in a traversal. The representation and referencing of ramps connecting roadways is not directly addressed by the model, but could be with extensions. Which geographic features are assigned external identifiers depends to some extent upon applications. The only features that require external identification are anchor

points. Most applications would require identifiers for traversals and traversal reference points, at a minimum. The efficiency of the model for supporting operations on very large databases remains unknown. It should be remembered that the model is not intended as a specification. Refinements are possible.

7. **Data Integrity** - The model solves many data integrity problems arising from multiple copies of data. For example, linear referencing methods do not have to be explicitly imbedded in multiple cartographic representations. There is a single datum. Changes in the datum, caused by changes in alignment, generate a cascade of changes in the database (for mappings between the datum and cartographic representations, for mappings between the datum and networks, for specification of traversals, and for offsets of traversal reference points and events). Rules that associate these necessary changes must be developed. Temporal objects could be created through the use of version identifiers. With appropriate rules for assembly of temporal objects, they could be used to track changes over time.
8. **Institutional Policy** - The greatest incentive for policy concerning linear referencing systems is the cost savings realized from data integration, data sharing, and reduction of chaos. Some agencies have already adopted policy in this regard (e.g., Utah). Use of a linear referencing system must be simple and straightforward. Even so, not all users must understand the use of all methods

Easily understood procedures must be developed for use in both the field and the office.

SUMMARY

A generic data model for linear referencing systems has been developed. The model includes multiple cartographic representations, multiple networks, and multiple linear referencing methods to which any amount of business data can be tied. The model supports integration of attributes attached to various spatial databases without requiring registration of cartographic representations in coordinate space. Instead, they are all linked to a single common datum. The model supports a set of fundamental operations that link the database world and the real world and allow transformations within the database world. The model will also support network analysis and basic GIS operations, although examples have not been developed.

The model is intended as a description of the core common requirements of as many application areas as possible. A need for extension of the data model to include particulars for specific application areas should be expected. Potential application areas include infrastructure management, transit, freight, intelligent transportation systems, urban planning, waterway navigation, and seismological testing.

Key concepts from the workshop and core elements of the data model have been incorporated into two on-going national-level efforts: 1) the GIS-T Pooled Fund Study and 2) the latest draft of the Dynamic Assignment of Network Attributes Specification from the ITS community. Ideas from the workshop are also being expressed within the context of the National Spatial Data Infrastructure (Dueker, 1995).

REFERENCES

Brace, K. and B. Peterson, (1994), "Cook County, Illinois TIGER/Line Conflation and Update Project", Proceedings, SHTO Symposium on Geographic Information Systems in Transportation, Albuquerque, N.M.

Briggs, D.W. and B.V. Chatfield, (1987), Integrated Highway Information Systems Report 133, National Cooperative Research Program, National Academy Press, Washington, D.C.

Clark, J. and L. Bain, (1994), "Integrating a Transportation Model Network into Enhanced TIGER/Line", Proceedings, SHTO Symposium on Geographic Information Systems in Transportation, Albuquerque, N.M.

Deighton, R.A. and D.G. Blake, (1993), "Improvements to Utah's Location Referencing System to Allow Data Integration", presented at the Transportation Research Board 3rd International Conference on Pavement Management, San Antonio.

Dueker, K.J., (1987), "Geographic Information Systems and Computer-Aided Mapping", Journal of the American Planning Association, Vol. 53, No. 3.

Dueker, K.J., (1995), "Access to Data: National Spatial Data Infrastructure", presented at the 74th Annual Meeting of the Transportation Research Board, Washington, D.C., January.

Dueker, K.J. and R. Vrana, (1992), "Dynamic Segmentation Revisited: A Milepoint Linear Data Model", Journal of the Urban and Regional Information Systems Association, Vol. 4, No. 2, Fall.

Federal Highway Administration, (1993), Highway Performance Monitoring System Field Manual, Office of Highway Information Management, FHWA Order M5600.1B.

Fletcher, D.R., (1987) "Modelling GIS Transportation Networks", Proceedings Vol. II, Urban and Regional Information Systems Association, 25th Annual Conference, Fort Lauderdale, FL.

Mapping Science Committee, (1993), Toward a Coordinated Spatial Data Infrastructure for the Nation, National Academy Press, Washington, D.C.

Michigan Department of Transportation, (1984), 1984 Reference System Report, Database Task Force, Reference System Review, Lansing, MI.

National Cooperative Highway Research Program, (1974), Highway Location Reference Methods, Synthesis of Highway Practice 21, National Academy of Sciences, Washington, D.C.

Nyerges, T.L., (1990), "Locational Referencing and Highway Segmentation in a Geographic Information System", ITE Journal, March.

Nyerges, T.L. and K.J. Dueker, (1988), Geographic Information Systems in Transportation, Office of Planning HPN-22, Federal Highway Administration, U.S. Department of Transportation, Washington, D.C.

Peterman, V., (1994), "Conflation: Merging Spatial Data Sets", Proceedings, SHTO Symposium on Geographic Information Systems in Transportation, Albuquerque, N.M. Ries, T., (1993), "Design Requirements for Location as a Foundation for Transportation Information Systems", Proceedings, SHTO GIS-T Symposium, Portland.

Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F, and W. Lorensen, (1991), Object- Oriented Modeling and Design, Prentice Hall, Englewood Cliffs, New Jersey.

Scarponcini, P., (1994), Location Data Modeling Effort Final Report, prepared for Minnesota Department of Transportation, Graphic Data Systems Corporation. United States Geological Survey (USGS), (1992), Spatial Data Transfer Standard: Part 1 Logical Specifications, Reston, VA.

Vonderohe, A.P., Travis, L., Smith, R.L., and V. Tsai, (1993), Adaptation of Geographic Information Systems for Transportation, Report 359, National Cooperative Highway Research Program, National Academy Press, Washington, D.C.

APPENDIX A. WORKSHOP PROGRAM

APPENDIX B. REMAINING SIGNIFICANT POINTS OF CONTENTION

The most significant unresolved point of contention has to do with the nature of traversals and the associations among linear referencing methods, networks, and the datum. Some participants believe that traversals (at least the ones containing traversal reference points) should be associated with the datum and not with networks. The assertion is that linear referencing methods should be based on datum measures and be tied together more directly by having traversals consist of ordered and directed collections of anchor sections.

The contrary position is that linear referencing are inherently network-based. A traversal has to do with getting from "here" to "there", implying nodes and links. Various linear referencing methods require various network models. For example, we will probably use different networks for highway systems, transit routes, and street addressing. Each of these applications has, or is, its own linear referencing method(s). Moreover, we would have to overpopulate the datum with anchor points and anchor sections in order to support all linear referencing methods.

Some have suggested that this problem be addressed by making a distinction between 1) traversals that have persistence and institutional significance and are used for linear referencing and 2) traversals that are not persistent, might result from analysis (e.g., pathfinding), and might not be used for linear referencing. It has been suggested that these two kinds of traversals be called "routes" and "paths". However, these two words are used extensively by the transportation community, usually with meanings other than those that are needed for this purpose. SHTO has glossary definitions for both "route" and "path". These definitions do not make the kind of distinction that is necessary to resolve the issue at hand. An additional consideration (not necessarily a point of contention) is that of dendritic traversals. It has been suggested that the Environmental Protection Agency's river reach model (RF3) be examined for its treatment of traversals with branches. This task was not undertaken by the Research Team.