Creating and Generalizing Linear Networks

Presentation and White Paper

provided by

Henry Jackson

Intergraph

1 Introduction

1.1 Data Intelligence through Spatial Relationships and Feature Information

There are several methods for adding intelligence to map data. They include distinctions in features classes, theme groupings, attribute information, hierarchies, networks, and topology. Topology, hierarchies and networks provide spatial relationships between simple digital feature data. Since maps are scaled version of the world, the accuracy and precision of individual features is devalued as scale increases. It is the spatial relationships between features that holds the map together to convey a sense of reality to the user. If spatial relationships between features are changed the user can quickly become disoriented when trying to understand the map.

It is important to understand that the spatial relationship between features can greatly affect the generalization process. It can influence both the selection of features to be generalized, and the output result of each generalization function. Features that are isolated from other map data will be generalized quite differently from features that share topology with surrounding features.

The topology between features can be used to validate and prevent certain unwanted changes that might occur during the generalization process. Feature attribution can be extremely important in determining which features can be generalized and under which circumstances. Feature hierarchies can provide structure, order and priority between related feature objects.

1.2 Why map producers avoid hierarchical data models

The insertion of feature hierarchies into digital map data can be a difficult and costly experience. As data is captured, the cartographer is usually focused on a very small portion of the earth surface. Data capture in a localized area generally produces large collections of individual features with very little focus on networks, hierarchies or feature relationships. Many map producers have opted for the quick collection of map data that is typically used to produce maps within a small range of the source scale. The value added information that feature hierarchies provides is not worth the effort to collect when the digital data is only used as a temporary vehicle to get to the production of single scale hardcopy products. Another important reason map producers exclude hierarchies in the data model is the cost both in time and money. There are limited choices of digital mapping systems that provide capabilities to capture hierarchies and other feature relationship details. These systems historically are very expensive and costly in time for training personnel to utilize such complex systems.

1.3 Why hierarchies are important

As map production moves toward cartographic databases, multi use data and rapid map products, the introduction of intelligence into the data becomes very important. When map products are produced from data that has been captured at a different scale, model generalization becomes necessary in the map production flow. Feature hierarchies provide structure to linear network features of the map. This information helps provide the order and priority of the features that constitute the linear networks. The order and importance of features within the linear network must be evaluated during model generalization to remove less important features which conflict with other features in the network. The elimination of features from linear networks is known as *Line Typification*.

1.4 Automated Digital Map Generalization and "The Blind Cartographer"

As commonly recognized, generalization is an integral process that extracts the important and relevant spatial information from reality. When producing smaller scale maps from a base map (usually large scale) or changing the map purpose to emphasize different content, decisions need to be made as to what features are to be retained and how they are to be represented within the available map space. Many less important features must be omitted and others modified correspondingly. The most important rule that guides the entire generalization process is to preserve the basic structure and characteristics of the geographic data and represent them in a legible manner.

Model generalization is necessary when the map content exceeds the capability of graphic representation. It involves a series of operations: from choosing the appropriate map content to re-grouping and arranging map features in relation to the reduced map space. The principal function of generalization is to distinguish between important and unimportant geographic features.

As software applications are developed to attempt the generalization process it becomes painfully obvious that manual generalization methods involve a cartographers, experience, training, and ability to intuitively construct detailed information from graphic representations that are not easily perceived in simply digital feature data. When a cartographer views graphic map data he can quickly detect physical relationships between features. Is the feature inside, outside, touching, to the left, close, far, or contiguous are all easily identifiable feature relationships to the trained eye of the cartographer. Color, symbology and detailed characteristics also aid the cartographer when attempting to distinguish feature information. Simple digital data is void of the characteristics that provide spatial relationships and feature information. The computer with it's make shift rules, algorithms and tools acts like a blind cartographer when it attempts to perform generalization tasks without the aid of intelligent map data.

1.5 Software Applications

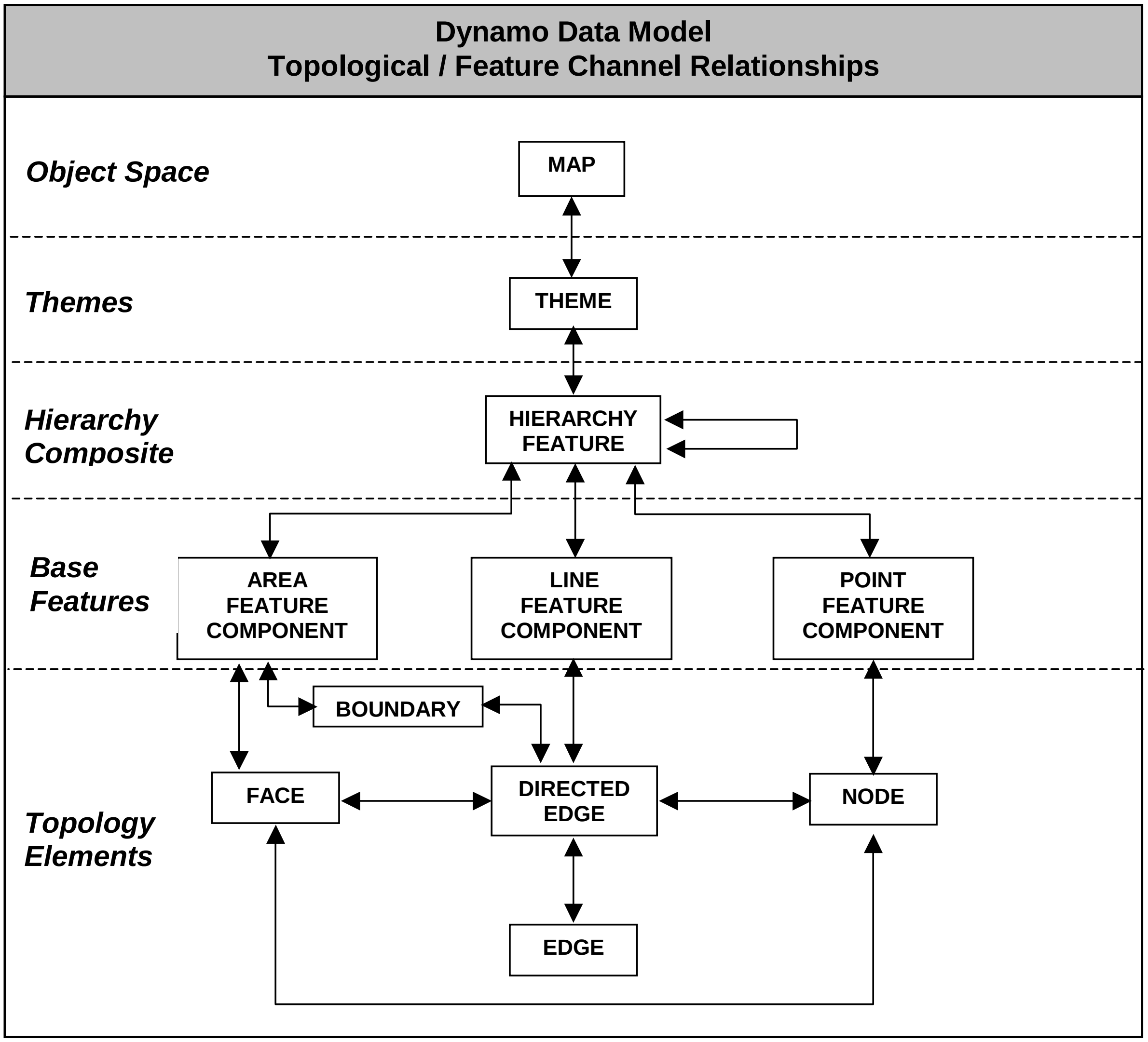
This paper includes information and examples from software applications provided by Intergraph. Dynamo is a Mapping application for collecting and integrating digital data and map products. DynaGEN is an Generalization application primarily consisting of model generalization tools that can be used in both a batch and interactive environment.

2 Hierarchical Schema for Network features

Linear Network features are some of the most prevalent and important features displayed on map products. The primary examples of linear networks are transportation road networks and drainage river systems. Even from great distances above the earth, these long linear features and network patterns are visible and distinct. The importance of the network is to link individual features that are physically connected, and to imply order, structure and priority to the seemingly unrelated pieces of spatial data. Feature Networks can be defined using many different factors. However the primary cartographic qualities used for constructing networks include attribution, length, straightness, continuity and patterning. This discussion will be limited to transportation networks specifically road type features.

2.1 Dynamo Data Model

Dynamo supports themes, composite features (owning multiple base features) and base features (owning topological elements). Themes organize geographical data into general categories (i.e. Hydrography). Composite features organize or relate base features into networks and/or groups. The Composite features add a hierarchy structure to geographical data. Composite features can have attribution but not topology. Base features describe real world objects as cartographic entities. There are 3 types of Base Feature Components: Areas, Lines, and Points. Base features are composed of topology and typically have attribution that is consistent throughout the extent of the feature. Channels are the relationships between feature entities and topological elements.



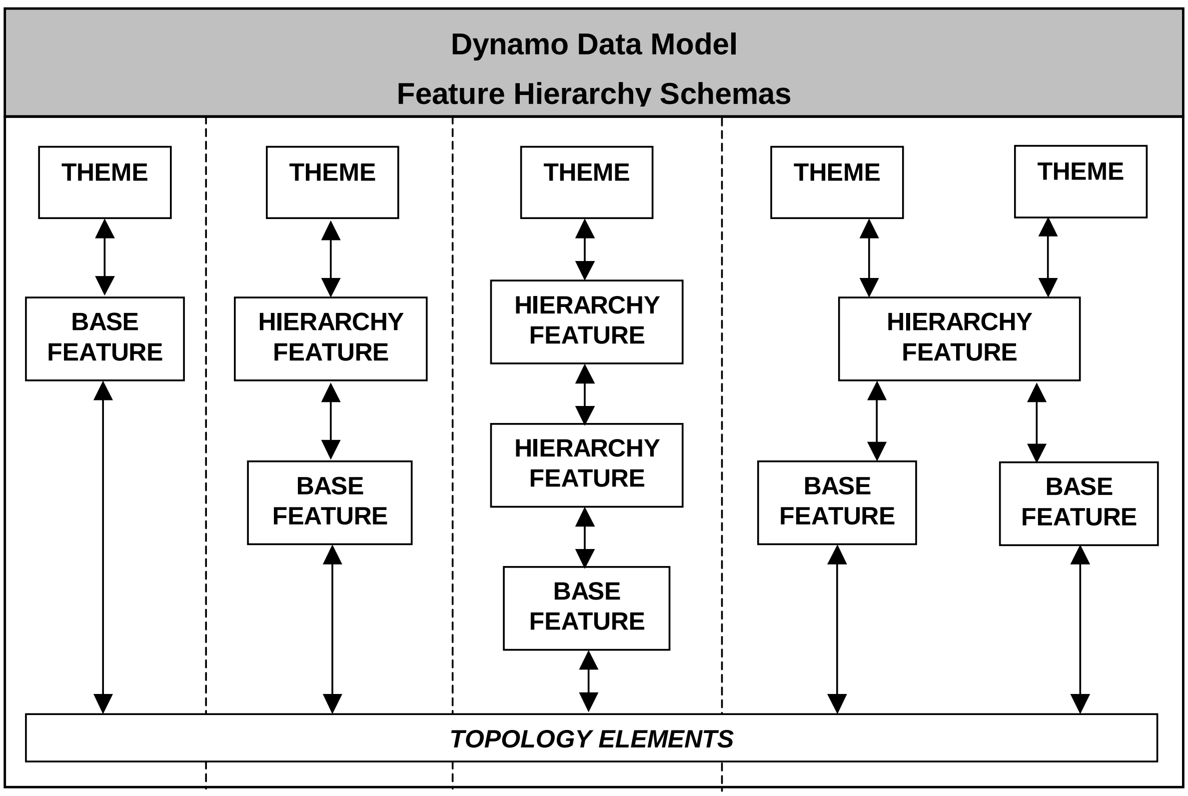
2.2 Feature Hierarchies

The features defined in a schema are organized into three classes: themes, hierarchy features, and base features. The data model can benefit in several ways from feature hierarchy features. Hierarchy features allow base features to be connected into networks or can be grouped to describe facilities. It is extremely difficult if not impossible to construct linear networks without the use of feature hierarchies. Since generalization operations rely on well constructed linear networks, the data model must include feature hierarchies to facilitate sophisticated computer assisted generalization tasks.

The hierarchy feature is considered a ***composite feature*** because it is composed of other feature elements. It can own one or more base features. Hierarchy features can also own or be owned by other hierarchy features. Hierarchy features may own attribution which is relative to each of the base features it owns. However the base feature need not have consistent attribution. Hierarchy features exist only to link base features and other hierarchy features and may not own topology. Base features are considered ***component features*** because they can be linked together as components of hierarchy features. Base features are homogeneous in that no attribute may vary within the extent of the feature. Base features are defined according to the kind of geometry they contain: point, line or area. Base features are considered the lowest feature level. They do not need to be spatially subdivided because the attribution is consistent throughout the feature. Base features are ***composed of*** topological elements.

The Dynamo Data Model supports several type of configuration between Themes and Topology. The minimum hierarchy is Theme to Base Feature to Topology element. A Theme must always be the top layer and Topology must always be the lowest layer. The table and diagram below illustrate the valid Dynamo hierarchy relationships and required elements of the data model.

Relationship	Dynamo Supported
Theme to Theme	not supported
Theme to Hierarchy Feature	typically one-to-many but will support many-to-many
Theme to Base Feature	typically one-to-many but will support many-to-many
Theme to Topology Element	not supported
Hierarchy Feature to Hierarchy Feature	typically one-to-many but will support many-to-many
Hierarchy Feature to Base Feature	one-to-many but will support many-to-many
Hierarchy Feature to Topology Element	not supported
Base Feature to Topology	many-to-many



3 Building Linear Networks

3.1 Usefulness of Feature Networks

The level of feature networking required for map production is dependant on the requirements of the map product and the amount and type of generalization to be performed. The network information can also be used to determine when and where generalization is required and to what set of features it would be applicable. The cartographic rules for structuring the network is also dependant on product and map production requirements. Feature networks add a level of intelligence to the map product that may not always be appreciated by the end user especially when hardcopy maps are the final product. Digital map products on the other hand can supply users with an unlimited wealth of information about the spatial data and increase the usability of the product. It's the map producer who can benefit most from feature networks because sophisticated mapping applications can derive feature relations without human intervention. This increases the level of automation in the production process.

3.2 Feature Merging

Many of the map data collection systems capture linear network features as short individual un-related features. Even when there is no attribute change or obvious reason for the abrupt break in the feature, some data is fragmented to a level where each feature has only a single topological element associated with it. In some data schemas insignificant attribute changes can also cause the network features to be fragmented. Even when hierarchy features and feature networks are not produced these fragmented features need to be merged together to form larger homogeneous features. Fragmented feature data can cause numerous problems with map production software including feature selection, generalization and symbolization.

3.3 Transportation Systems and Grids

Transportation system networks are sets of systematic features which generally require connectivity to other parts of the network. These types of features seem to have no beginning and no end. Transportation systems can contain main arteries and smaller appendages which serve to connect parts of the network or service other map features. In some respects a transportation system is similar to the circulation system of the human body. Roads seem to go where roads are needed. In some situations they are prefect grids while others form bridges between main lines. In any case there is a systematic approach to their existence and an ordering to their importance. Usually transportation systems are composed of line feature components but in some instances can include point feature components like bridges, traffic circles and interchanges.

There can be many different types of transportation systems including Major Roads (Highway/Interstates/Routes) (HIR's), Roads, Divided Roads, Streets, Grids, Casements. Railroads and Trails. Casements are a special network intended to delineate the edges of both sides of transportation type features. Transportation systems are not limited to road type features. Railroad systems with multiple tracks and railroad yards can also exist. In areas where numerous Trails, Cart Tracks or Paths are portrayed, there may be a need to network or produce hierarchies for these feature type.

3.4 Constructing Feature Hierarchies and Networks

There are several steps involved in generating useful linear feature networks. First a schema must be developed which supports the necessary attribution and feature relationships. The design of the schema must consider the types of features in the data model and the product requirements of the map product. The next step is the create a data dictionary with the necessary hierarchy features and attribute assignments. An automated process can then assign the necessary parent hierarchy features to low level feature components collected during data capture. This in no way constitutes a complete linear network but lays the ground work for constructing an intelligent hierarchy of linear features. Additional attribute information and reassignment of some portions of the network are expected and must be carried out by the cartographer.

3.5 Complete Hierarchies Operator in DynaGEN

The "Complete Hierarchy" operator of DynaGEN adds new hierarchy features to the object space and connects base feature class to them as defined in the Hierarchy Network table. The operator generates a one-to-one relationship between base feature and hierarchy feature. Feature hierarchies are used to connect multiple base features to a common parent feature. This aids in construction of feature networks when base feature attribution differs such to create a fragmented set of features. The data dictionary must defined the feature class hierarchies and define all feature to feature and feature to base feature connections.

3.6 Hierarchy Network Merge Operator in DynaGEN

The final step is to run the automatic "Hierarchy Network" operator with the "Feature Network Blending" algorithm. This application uses input parameters and input attribute conditions to determine candidate base feature components which could be merged. Then the application merges base features based on inherent algorithm rules and reconstructs the hierarchy features to complete the hierarchy network. The constructing of feature networks is automated to a large degree but some interactive efforts will still have to be performed by the user. These include reviewing any queues generated by the process and validating the overall network for cartographic acceptability. Depending on the complexity and structure of the data model, DynaGEN is limited to the amount of hierarchy feature reconstruction that can be performed. Some high level complex networks may have to be performed by the user.

Most transportation systems are defined by connecting base feature components to form the straightest path. DynaGEN can generate parts of a transportation network by comparing feature classes and attributes to form the straightest homogeneous path along the route until the path terminates or dead ends. DynaGEN can also make cartographically intelligent decisions for merging features across intersections, forks and junctions. DynaGEN requires some external source of information to determine the existence of transportation networks. This means that through feature hierarchies all the base feature components of a network must be linked or that the user must identify the components through fences or result sets.

DynaGEN Network Blending for roads utilizes three parameters. A minimum length value is used to eliminate very short individual features that act as bridges between longer feature components. Sometimes these short components exist because of attribute differences. The minimum length parameter defines conditions for when to override the defined attribute checks defined for blending features of similar attribution. The angle 2 lines parameter defines the variance from a straight line when only 2 lines touch and the angle >2 lines parameter is used when multiple line converge. The algorithm is design to blend the straightest lines within the defined parameters.

3.7 Rules for Constructing Road Hierarchy Networks

In an effort to design a multi product cartographic database for NIMA (National Image Mapping Agency) a hierarchy schema was developed for linear road networks. Attached are the rules and a diagram illustrating examples of the data structure. DynaGEN supports this set of rules in defining and constructing feature hierarchies and networks. **See diagrams in the back of this white paper.**

4 Generalization of Linear Networks

Manual generalization of linear networks can be some of the most difficult and tedious work that the cartographer would ever have to perform. Typically the cartographer must select important features from the existing network rather than eliminating features from it. When attempting the elimination method it becomes more difficult to identify the un-important parts of the network because some of the features have been removed. Automated and/or computer assisted methods are needed to aid the cartographer in this monumental task.

4.1 Typification

Typification generalization methods makes maps more readable by reducing the complexity of a group of line features in a dense region by deleting a subset of the features while maintaining the representative pattern in their distribution. This is useful when the presence of a feature type in an area is important, but the actual number of occurrences of the feature is not.

Typification methods of generalization remove detail in the map by eliminating line features that do not represent a significant part of the network of line features. The intent is to show a representative pattern of the more significant features in the network.

4.2 DynaGEN Line Typification

The Line Typification (conflict resolution) algorithm removes line features such as roads based on their proximity to other lines. One of the primary rules of the algorithm is to maintain connectivity and continuity throughout the linear network. This operator also assumes total length is a primary condition for determining importance to the network. DynaGEN Typification permits the user to designate critical line features within the feature selection that must not be removed. This assigns a priority to the feature and re-evaluates conflicts based on this special priority

The algorithm requires two parameters. The proximity to length parameter defines what percentage of a features total length must conflict with other lines within the network. The minimum space tolerance defines the distance between lines that would constitute conflicts. Since DynaGEN operators on vector data and final symbology is not present during this task, the minimum space tolerance also can be used to simulate conditions when final symbology is available. Normally this parameter is set much higher than standard symbol widths. **See typification results as demonstrated in the oral presentation.**

4.3 Future Requirements for DynaGEN Line Typification

Intergraph recognizes the need to enhance the DynaGEN typification algorithms. One of the most important aspects of the generalization process is the ability to assign priority to features based on class and attribute information.

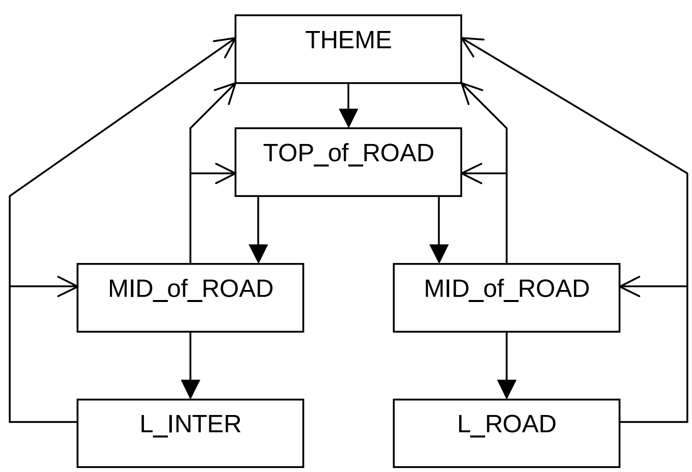
Rules for Constructing Road Hierarchies

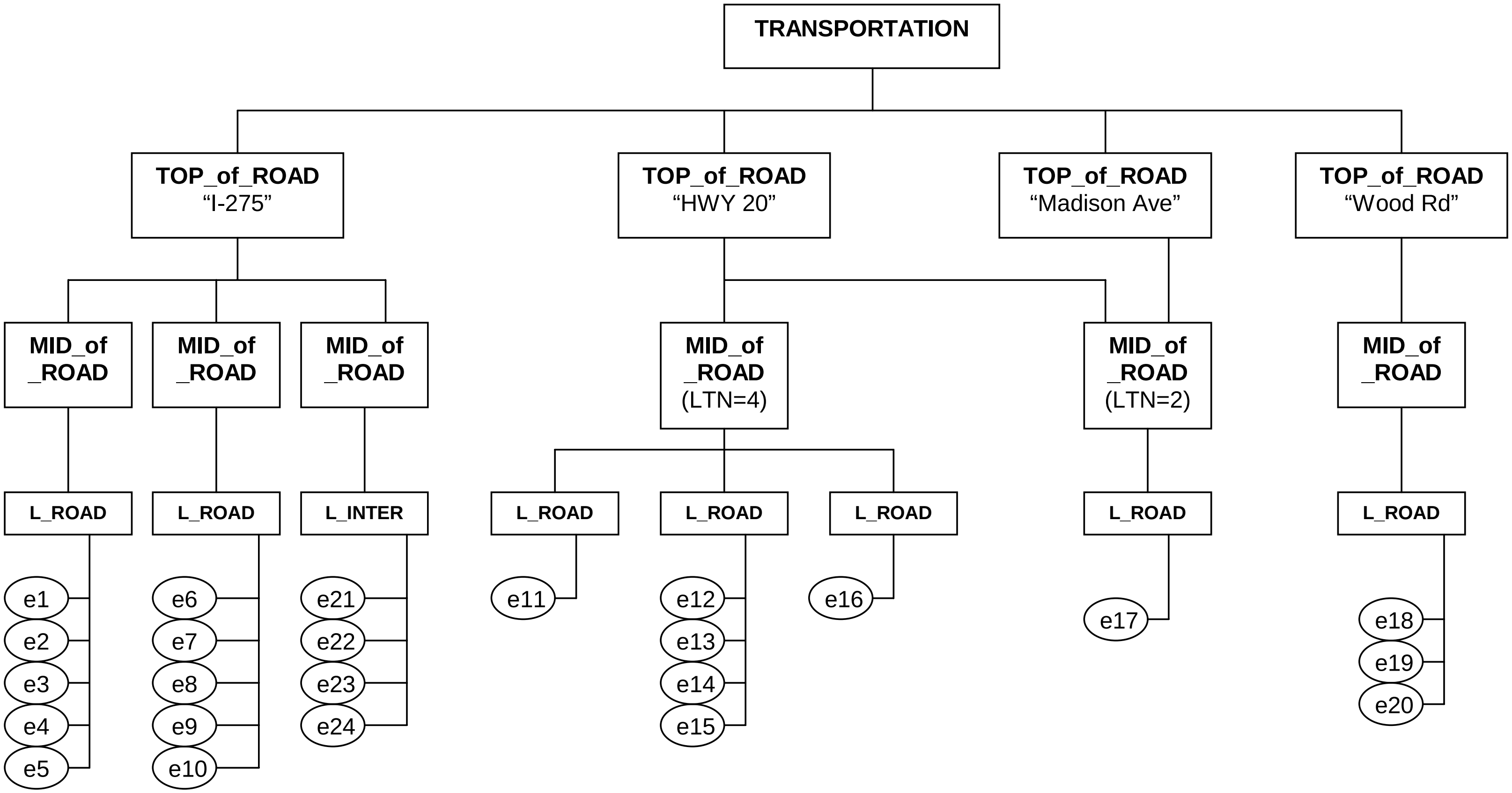
1. L_Road must be a component of Mid_of_Road or direct connection to a Theme.
2. L_Road may not be a component of more than one Mid_of_Road.
3. L_Road must be composed of one or more edges; the edges must be contiguous.

4. L_ Interchange must be a component of Mid_of_Road or direct connection to a Theme.
5. L_ Interchange may not be a component of more than one Mid_of_Road.
6. L_ Interchange must be composed of one or more edges; the edges must be contiguous

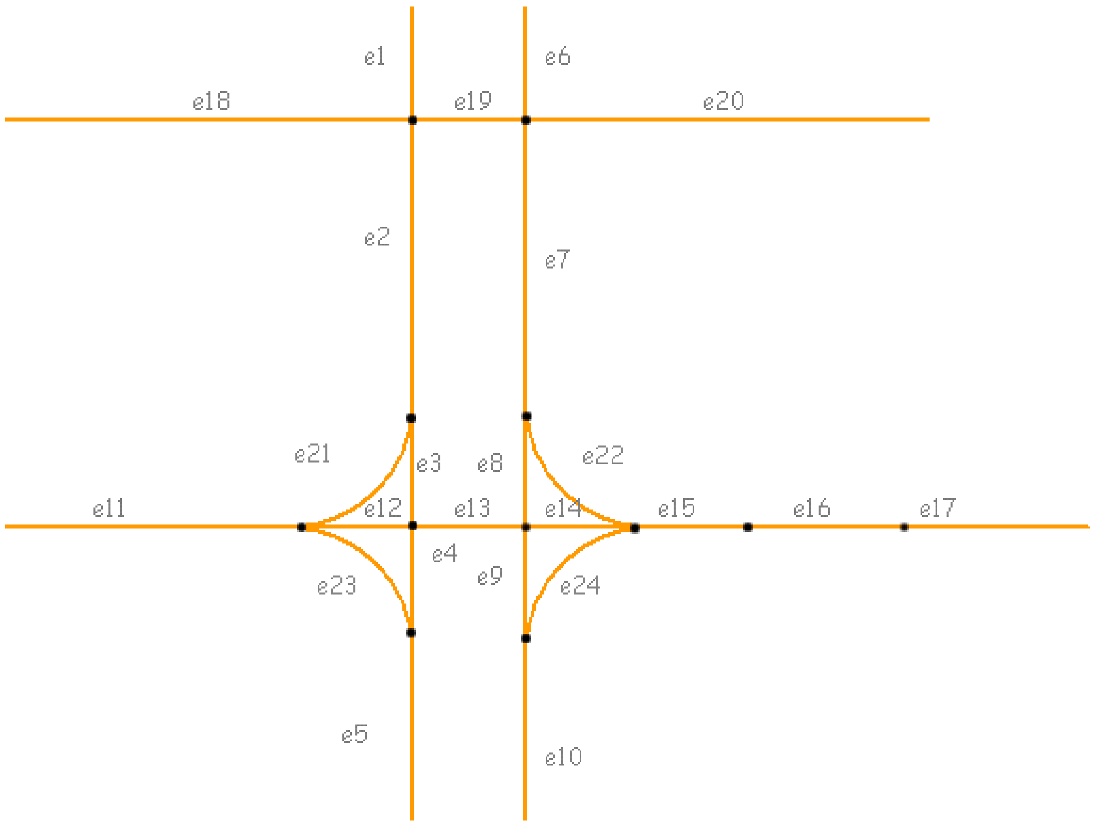
7. Mid_of_Road must be a component of Top_of_Road or direct connection to a Theme.
8. Mid_of_Road must be composed of one or more L_Road feature components and these L_Road feature components must be contiguous.
9. Mid_of_Road may be a component of multiple Top_of_Road features.
10. Mid_of_Road may additionally be composed of one or more L_Interchange features and these L_Interchange features need not be contiguous.

11. Top_of_Road must be composed of one or more Mid_of_Road features and these Mid_of_Road features need not be contiguous.
12. Top_of_Road must be connect to one or more Themes.





See attached diagram



Road component Diagram