Spatial Data Editing

- The removal of error from, and updating of digital maps.
- Cover two types of errors:
  - Location errors
  - Topological errors

**Location errors**
- Human errors in manual digitizing
- Errors in scanning and tracing including collapse lines, misshapen lines, and extra lines. (Figure 5.1)

*Figure 5.1*
Common types of digitizing errors from tracing. The thin lines are lines on the source map, and the thick lines are lines from tracing.
Errors in converting the digitized map to real-world coordinates caused by erroneous control points.

Duplicate lines from manual digitizing or tracing and form a series of tiny polygons. (Figure 5.2)

**Topological Errors**
- Errors that violate the topological relationships used in a GIS package.

Figure 5.2
Digitizing errors from duplicate lines include sliver polygons and missing labels for the sliver polygons.
- **Undershoot**: gap between arcs. (Figure 5.3)
- **Overshoot**: one arc is over extended.
- **Dangling arc**: results from undershoot or overshoot, it has the same polygon on the left and right sides.
- **Dangling node**: end of dangling arc (Figure 5.4).
- **Pseudonode**: appears on continuous arc and divides the arc unnecessarily into separate arcs (Figure 5.5). Some pseudonodes are acceptable and necessary.
  - to show different segments of the road,
  - the starting point of digitizing
– The error in the direction of arc for stream and one-way road. (Figure 5.6)
– Multiple labels error caused by open polygon. (Figure 5.7)

**Topological Editing**

- Topology-based GIS package can detect topological errors and mark errors with special symbols.
- Special action is needed to remove the specific types of digitizing errors.
- After removing errors, map topology must be rebuilt.

---

**Figure 5.6**
The from-node and to-node of an arc determine the arc’s direction.

**Figure 5.7**
Digitizing error of multiple labels due to unclosed polygons.
Correcting Digitizing Error

Global Method

• Apply specific tolerances to an entire map to remove digitizing errors.
  – CLEAN uses dangle length and fuzzy tolerance
    » Dangle length: minimum length for dangling arcs on the output coverage. The dangling arc is removed if it is shorter than the dangle length. (Figure 5.8)

Figure 5.8
The dangle length specified by the CLEAN command can remove an overshoot if the overextension is smaller than the specified length. In this diagram, the overshoot (a) is removed and the overshoot (b) remains.
Fuzzy tolerance: distance between vertices along an arc and vertices along two nearby arc. It can be used to remove the duplicate lines within the specified fuzzy tolerance. (Figure 5.9)

- Global method must be applied carefully, if dangle length is too large, it will remove undershoots as well as overshoots. (Figure 5.10)
Local Method
- Deal with selected digitizing errors from a digitized map.
- Use some tolerances to help editing.
  - Nodesnap: node is snapped to the new location within nodesnap tolerance.
  - Editdistance: search radius for selecting features for editing.

Edgematching
- Use for editing multiple maps
- Operation before joining maps because line from two maps rarely meet perfectly along the border. (Figure 5.16)

Non-topological Editing
- Work with location errors
- Use for updating and revision of spatial data
  - Delete, Move, Cut and Paste (Figure 5.17)
  - Reshape (Figure 5.18, 5.19, 5.20)
  - Split and Merge (Figure 5.21, 5.22)
Figure 5.16
The diagram shows lines from two adjacent quads do not meet perfectly. The mismatch is only visible after zooming in.

Figure 5.17
After a polygon is moved in ArcView, a void area appears in its location. GIS users, who are used to topology-based operations, may find this ArcView operation strange.

Figure 5.18
Reshape a line in ArcView by moving a selected vertex.

Figure 5.19
Reshape a line in ArcView by deleting a selected vertex.

Figure 5.20
Add a vertex to a line in ArcView.
**Other Types of Map Feature manipulation**

**Line Simplification**
- Simplifying or generalizing a line by removing some of its points.
- Use in reducing map scale for proper display.
- Use Douglas-Peucker algorithm (Figure 5.23).
  - GENERALIZE command in Arc/Info (Figure 5.24).
Figure 5.23
The Douglas-Peucker line simplification algorithm is an iterative process, which requires use of a tolerance, trend lines, and calculation of deviations of vertices to the trend line. See text for explanation.

Figure 5.24
Two views of the Puget Sound area in Washington. Created from ARC/INFO's GENERALIZE command with a weed tolerance of 2 km, Map (b) is a generalized version of Map (a).
**Line Densification**

- The process of adding new points to select lines in a map at specified interval. *(Figure 5.25)*

**Line Smoothing**

- Add new points to lines, but the location of new points is generated by mathematical functions.
- The command SPLINE in Arc/Info uses cubic polynomials function for line smoothing.

---

**Figure 5.25**

Line densification adds new vertices to an arc at a specified interval.
\[ y = \beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3 \]

**Transfering Features**

- Transfering map features from one map to another.
- Useful when different map coverages share some common boundaries. **Ex:** a soil map may share some common boundaries with a land use map.
- Save time in digitizing and ensure the maps match spatially for data analysis.