The CEDRA Corporation's

COMMAND OF THE MONTH

A monthly information bulletin — March 2010

FEATURED COMMAND

Determining Duplicate Features

Application Description

We dare say that most everyone involved with GIS has come across a situation where it was needed to find duplicate features. That is, features that are the same and which reside on top of each other.

How this occurs no one "seems to know", but from time to time a copy or two of the same feature seems to pop up. The task then becomes how do we find these duplicate features.

The CEDRA Solution

To address this issue, the [Duplicate Features] command within the CEDRA-Polygon-Tools toolbar, shown in Figure 1, can be used. This command enables the user to determine which features are identical in a single layer or by comparing two layers of the same feature type.

Duplicate Features Overview

During the work flow of a project, it is possible for a user to intentionally, or unintentionally create duplicate features, which may reside in the same or in different layers. The [Duplicate Features] command assists the user in identifying such features, and if specified:

a. highlight the duplicate features,
b. remove the first or all duplicate features,
c. add attributes that are not in common, or
d. update attributes in common. This command is operable on selected features of a layer, or on all features.

Duplicate Features Operation

➤ **1** **Scroll** down in the **table of contents area**, and **select** the **layer** that contains the features to be queried for duplicity. As will be seen below, this is an optional step. For the purpose of this command description, let us refer to this layer as the *TOClayer*.

> ***Command Of The Month bulletin***
>
> *This month's issue discusses how to find duplicate features in a single layer or by comparing two layers of the same feature type.*

➤ **2** **Select** with the *Select Features from Active Layers* tool the **features** to be queried for duplicity. As will be seen below, this is an optional step, but in order to use this step, the preceding step must have been carried out.

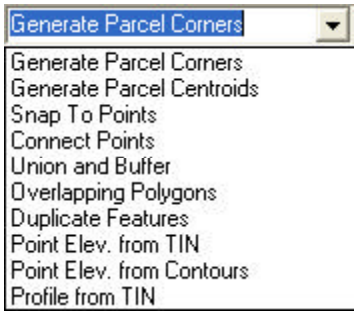

Figure 1
CEDRA-Polygon-Tools Toolbar

➤ **3** **Scroll** down in the **{CEDRA-Polygon-Tools}** menu combo box, and **select** the **[Duplicate Features]** menu command (see Figure 1) to display the multi-input dialog box of Figure 2.

➤ **4** **Scroll** down in the *Layer to be Checked:* data field, and **select**

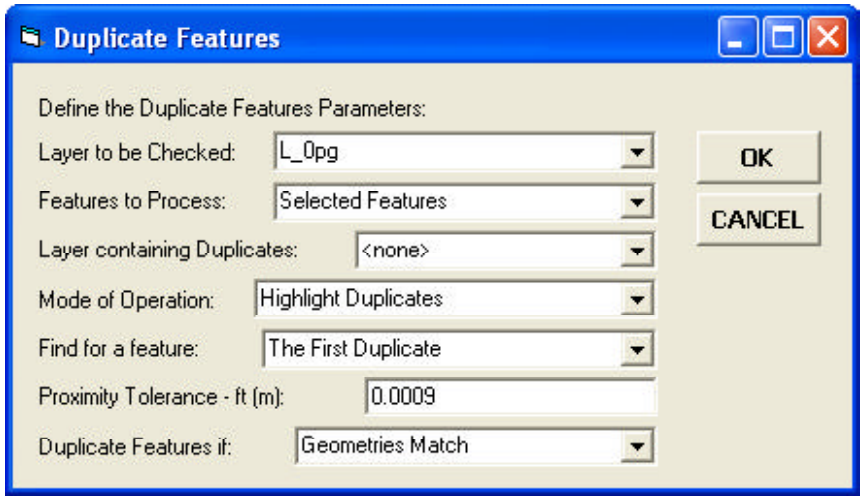

Figure 2 - Duplicate Features Command Multi-Input Dialog Box

The CEDRA Corporation™
Total CADD for Engineers™
Bridging Engineering and GIS™
151 Sully's Trail - Suite 6
Pittsford, New York 14534
Phone: 585-232-6998
Fax: 585-262-2042
E-mail: cedra@cedra.com
URL: http://www.cedra.com

the **layer to be checked** for duplicate features. For the purpose of this command description, let us refer to this layer as the *LayerA*. Note that the layer specified in this step is the layer to be processed. If features have been selected in the above steps, then *LayerA* should be the same as the *TOClayer*. Otherwise, the selection of features above will be disregarded.

➤ **5** **Scroll** down in the *Features to Process:* data field, and **select** the:

- **Selected Features** option to process only the features that have been selected in the first two steps, provided *TOClayer* and *LayerA* are the same.

- **All Features** option to process all features in *LayerA*, regardless whether any features have been selected in the first two steps, or not.

➤ **6** **Scroll** down in the *Layer containing Duplicates:* data field, and **select** the:

- **<none>** option to check for duplicate features only in the *LayerA*.

- **The layer** to be checked for having features that are duplicate of features in the *LayerA*. For the purpose of this command description, let us refer to the layer selected in this data field as the *LayerB*.

➤ **7** **Scroll** down in the *Mode of Operation:* data field, and **select** the:

- **Highlight Duplicates** option to highlight any of the features in the *LayerA* that have been found to have duplicates.

- **Remove Duplicates** to remove any of the features in the *LayerA* that have been found to have duplicates.

- **Add Attributes not in Common** option to add to any of the features in the *LayerA* that have been found to have duplicates in LayerB any attributes of *LayerB* that do not exist in *LayerA*.

- **Update Attributes in Common** option to update the attributes of any of the features in the *LayerA* that have been found to have duplicates in *LayerB* with the contents of the attribute fields of *LayerB* that exist in *LayerA*.

➤ **8** **Scroll** down in the *Find for a feature:* data field, and **select** the:

- **The First Duplicate** option to identify and operate only on the first duplicate feature, if there are more than one duplicate features.

- **All Duplicates** option to identify and operate on all duplicate features that may be found.

➤ **9** **Enter** in the *Proximity Tolerance - ft (m):* data field the tolerance distance in display units within which features may deviate from each other but still be considered as duplicates. For polylines, curves and polygons, all vertices must be within the tolerance for two features to be considered duplicate.

➤ **10** **Scroll** down in the *Duplicate Features if:* data field, and **select** the:

- **Geometries Match** option to select duplicates based only on a geometric match.

- **Geometries & Attributes Match** option to select duplicates not only due to a geometric match, but also on a match of attributes.

➤ **11** **Click** at either the **OK** button to continue,
or
click at the **Cancel** button to abort the command.

Having clicked at the *OK* button, and if the response in the *Mode of Operation:* data field is the:

- **Add Attributes not in Common** option, or the **Update Attributes in Common** option, the choice list dialog box of Figure 3 is displayed, in which case the user should continue with the next step.

- **Highlight Duplicates** option or the **Remove Duplicates** option, the command terminates.

When the choice list dialog box of Figure 3 is displayed, the user should:

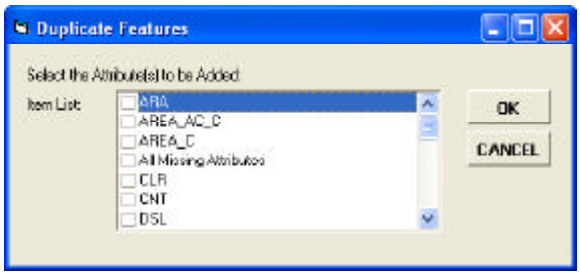

Figure 3
Attribute Selection Choice List Box

➤ **12** **Scroll** down in the *Item List:* data field, and **select** the:

- **All Missing Attributes** option to add all attributes not in common, or to update all common attributes in *LayerA* for the duplicate features.

- **the attribute(s)** to be added which is not in common, or to be updated in *LayerA* for the duplicate features.

➤ **13** **Click** at either the **OK** button to execute the command as speci-

fied in the above steps, and then terminate.
or
click at the **Cancel** button to abort the command.

Having clicked at the OK button, note the following:

1. If the **Selected Features** option is selected in the *Features to Process:* data field, and if no features have been selected, a warning message to that fact is displayed. In this case click at the *OK* button to acknowledge the message, and abort the command.

2. If two layers are to be processed and they are of different feature type, an error message similar to that shown in Figure 4 will be displayed. In this case click at the *OK* button to acknowledge the message, and abort the command.

3. When the command terminates successfully, a message is displayed in the status bar informing the user of how many duplicate features have been found and processed.

Note

In determining duplicate polyline and polygon features, the [Duplicate Features] command will reverse the direction of the source feature. That is to say, if a match can not be made in the command's initial checking, the command will reverse the direction of the feature to see if a match can be made. This process does not alter the source geometry of the feature.

So that, a duplicate polyline or polygon feature will be found even if the direction of the features are opposite to each other, provided the coordinates of the vertices are identical when the original feature is reversed.

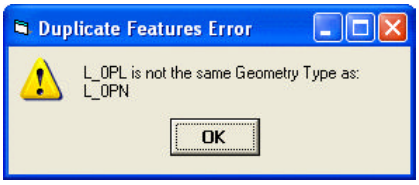

Figure 4
Different Feature Type Error Message

Summary

To our knowledge there is no "out-of-the-box" ArcGIS solution to determine duplicate features. So that, those users who come across situations where duplicate features exist will find the [Duplicate Features] command very useful in determining these features.

As always, users who have a need for functionality that is not presently available in CEDRA software should feel free to forward these requests to CEDRA, as well as, any other comments or suggestion you may have.

If you have a request for Command Of The Month, feel free to phone, fax or e-mail your request to The CEDRA Corporation.