U.S. DEPARTMENT OF THE INTERIOR

GEOLOGICAL SURVEY

Digitizing Procedure for the National Coal Resources Data System
of the U.S. Geological Survey

By Judy F. Hunter, Sharon S. Crowley, and Antoinette L. Medlin

Introduction

The National Coal Resources Data System (NCRDS) of the U.S. Geological Survey needs certain graphic information for a quadrangle area in order to assess coal resources in it. A set of map overlays for a quadrangle area depicting latitude and longitude locations, data-point locations, political boundaries, coal beds, strip and deep-mined areas, mined-out areas, coal-burn areas, Federal and State land ownership, and other geologic information is essential to coal-resource assessment. The map overlays are converted into a digital form that can be used in the coal-resource assessment program, GARNET, and other graphics programs. These digital data must meet particular format requirements, which in turn demands adherence to certain digitizing procedures. The purpose of this paper is to describe those requirements and procedures.

File Formats

Each quadrangle requires a point-data file and files for political boundaries, outcrops, deep- or strip-mined areas, burned areas, and any other type of boundary that may be present. Each file should be recorded in a fixed-block-size format that does not exceed 2150 decimal 8-bit bytes or characters. Also, the number of data must be a multiple of 43 decimal bytes. These data must be coded in EBCDIC IBM-standard characters and recorded on 9-track 800- or 1600-bpi (bits per inch) IBM-compatible computer tape.

These data must be represented using X and Y coordinates recorded as real numbers that indicate an absolute distance from the origin in units of 0.01 inches. The origin of the X and Y axes is the lower left hand corner of the map border to ensure that all positional data for each overlay in a quadrangle will be in reference to the same X and Y axes. Each X and Y coordinate pair has an identifier associated with it to indicate "point" data or "line" data.

Point Data

The point-data file contains reference points with associated latitude and longitude values and data points with associated identifiers. The number of reference points required varies according to the scale of the map. A 1:24,000 scale map requires eight reference points: four corner points and four interior points. A 1:100,000 scale map requires 30 reference points: four corner points, twelve points on the map border, and fourteen interior points.

A latitude and longitude value is assigned to each reference point. The format for the latitude and longitude values is shown in the POINT DATA SAMPLE (fig. 1). The values are given in degrees, minutes, and seconds

with a space between each as shown in the illustration.

The identifiers for the data points may contain as many as 16 characters but may not contain any special character other than a hyphen. The identifiers on the map for the data point must be the same as the identifiers on the stratigraphic data forms submitted to NCRDS for the same data points. The information on the stratigraphic forms is combined by matching the identifiers; the identifiers must be exactly alike.

The exact format for the final point-data file is shown in the POINT-DATA SAMPLE (fig. 1). The first and second columns of numbers, which are right justified, are the X and Y coordinates, respectively, in inches for each reference and data point in the point-data file. The third column of characters, which is left justified, contains the latitude and longitude values for each reference point or the identifier associated with each data point.

In order to obtain the proper format in the final point-data file, the reference points are digitized starting in the lower left hand corner of the map and proceeding counterclockwise around the map to digitize the four corner points. The interior reference points are digitized next in the same order. The data points follow the reference points at random. On a 1:100,000 scale map, the first reference point is the same as a 1:24,000 scale map, but all sixteen reference points are digitized in a counterclockwise order around the border rather than just the corner points. Also as depicted in figure 1, the file requires the last line to state the number of data points contained in it. This digitizing procedure will ensure the proper sequence for the computer software which interprets the digital data.

Line Data

As with point data, all line data must show X and Y coordinates in inches for each point on the line, and each line must be assigned a label to identify it. The format is shown in the LINE-DATA SAMPLE (fig. 2). The first and second columns, which are right justified, are the X and Y coordinates, respectively, in inches for each point on the line. The third column, which is left justified, is the identifier associated with the line. The spacing is the same as that for the POINT-DATA SAMPLE (fig. 1).

The identifiers should label the lines as accurately as possible. One quadrangle may contain one or several coal beds with associated stripped or mined-out areas. The Fire Creek Coal Bed should be labelled as completely as possible within the 16 characters allowed to identify it. The first label may be FIRE CRK COAL 1. Numbering the outcrop lines within a bed will enable the operater to accurately label all the lines in one coal bed, such as FIRE CRK COAL 1, FIRE CRK COAL 2, FIRE CRK COAL 3, or FIRE STRIP 1, FIRE STRIP 2, FIRE STRIP 3, etc. As shown in the LINE-DATA SAMPLE (fig. 2), the different types of lines are set apart by the identifier.

2

FIGURE 1—POINT-DATA SAMPLE

	(Coordinates)				(Identifier)			
	(X)		(Y)					
Columns	1	9	14	22	28			43
	0.000			.029	32	15	00	
	0.000			.029	88	45	00	
	19.332			.023	32	15	00	
	19.332			.023	88	37	00	
	19.336		22.731		32	22	00	
	19.336		22.731		88	37	00	
	.023		22.742		32	22	00	
	.023		22.742		88	45	00	
	6.458		7.605		32	17	00	
	6.458		7.605		88	42	00	
	12.889		7.596		32	17	00	
	12.889		7.596		88	40	00	
	12.887		15.167		32	20	00	
	12.887		15.167		88	40	00	
	6.459		15.172		32	20	00	
	6.459		15.172		88	42	00	
	5.553		18.459		THR174			
	5.509		18.344		THR173			
	13.821		21.649		THR210			
	14.518		20.925		THR211			
	14.783		20.983		THR212			
	5		POINTS DUMPED					

(alpha characters may be in upper or lower case)

Columns	Description
1-9	X coordinate in inches (f9.3)
10-13	blank
14-22	Y coordinate in inches (f9.3)
23-27	blank
28-43	latitude, longitude or data point identifier

FIGURE 2--LINE DATA SAMPLE

(Coordinates)			(Identifiers)	
(X)	(Y)			
Columns 1 9	14	22	28	43
.845		−.115	DOUGLAS COAL 1	
.835		−.120	DOUGLAS COAL 1	
.826		−.126	DOUGLAS COAL 1	
.819		−.134	DOUGLAS COAL 1	
.812		−.143	DOUGLAS COAL 1	
.807		−.153	DOUGLAS COAL 1	
.804		−.163	DOUGLAS COAL 1	
.802		−.174	DOUGLAS COAL 1	
.801		−.185	DOUGLAS COAL 1	
.801		−.197	DOUGLAS COAL 1	
.803		−.207	DOUGLAS COAL 1	
.806		−.216	DOUGLAS COAL 1	
.811		−.223	DOUGLAS COAL 1	
.817		−.229	DOUGLAS COAL 1	
.825		−.234	DOUGLAS COAL 1	
.833		−.236	DOUGLAS COAL 1	
.842		−.238	DOUGLAS COAL 1	
.850		−.238	DOUGLAS COAL 1	
.856		−.237	DOUGLAS COAL 1	
.861		−.236	DOUGLAS COAL 1	
.865		−.236	DOUGLAS COAL 1	
.868		−.235	DOUGLAS COAL 1	
.969		−.234	DOUGLAS COAL 1	
.913		−.177	DOUGLAS COAL 1	
.913		−.178	DOUGLAS COAL 2	
.914		−.178	DOUGLAS COAL 2	
.917		−.178	DOUGLAS COAL 2	
.921		−.177	DOUGLAS COAL 2	
.927		−.176	DOUGLAS COAL 2	
.935		−.175	DOUGLAS COAL 2	
.944		−.174	DOUGLAS COAL 2	
.954		−.173	DOUGLAS COAL 2	
.966		−.171	DOUGLAS COAL 2	
.977		−.170	DOUGLAS COAL 2	
.989		−.169	DOUGLAS COAL 2	
1.000		−.167	DOUGLAS COAL 2	
1.011		−.166	DOUGLAS COAL 2	
1.021		−.163	DOUGLAS COAL 2	
1.030		−.159	DOUGLAS COAL 2	
1.038		−.154	DOUGLAS COAL 2	
1.044		−.147	DOUGLAS COAL 2	
1.049		−.139	DOUGLAS COAL 2	
1.052		−.130	DOUGLAS COAL 2	
1.054		−.120	DOUGLAS COAL 2	
1.055		−.108	DOUGLAS COAL 2	
1.055		−.097	DOUGLAS COAL 2	
1.055		−.085	DOUGLAS COAL 2	
1.055		−.074	DOUGLAS COAL 2	
1.055		−.063	DOUGLAS COAL 2	

(alpha characters may be in upper or lower case)

4

The identifier associated with the line may contain as many as 16 alphabetic and/or numeric characters, but the hyphen is the only acceptable special character. The identifier should completely label the line according to the type of line it is, and it must be different for each line segment or polygon in a line file. A line file contains all the line segments or polygons for a quadrangle; therefore, for a complete outcrop pattern of a particular coal bed, an overlay may contain many identifiers, i.e., COAL OUTCROP 1, COAL OUTCROP 2, COAL OUTCROP 3 (see fig. 3).

Although line data can become very complicated on an overlay that contains dense geologic data, strict adherence to five general rules for digitizing line data simplify the task. The general rules are:

1. Digitize lines in a counterclockwise direction if possible.

2. All digitized lines must form closed polygons as simple island polygons, or by both endpoints intersecting the map border, or by intersecting another line (see Number 4 below). Lines cannot be left "hanging" in the quadrangle interior.

3. Lines that intersect the map border must cross to the exterior side of the border by at least 0.25 of an inch.

4. Lines that intersect other lines, such as stripped areas intersecting outcrops, must be continued 0.25 of an inch beyond the point of intersection, unless such an extension would intersect any other similar type of line for the same coal bed, such as one stripped line intersecting another stripped line. These extended lines are closed back to the beginning point of the line avoiding any other line. If closure of a line is impossible back to the beginning point, the line may be closed back to the map border (see fig. 3).

5. Detailed deep-mined areas should be generalized.

Figure 3 depicts an actual digitized coal bed including outcrop lines and strip- and deep-mined lines. The map border outline has been added to the line data for display purposes only. The dashed lines show the digitizing method used to close lines into polygons in different circumstances as explained further in the following examples.

The line labelled COAL OUTCROP 1 is an example of line data that begins and ends intersecting the map border. At "A", the line is begun 0.25 of an inch on the exterior side of the map border. The identifier of the line is added (COAL OUTCROP 1), and it is digitized in a counterclockwise direction to point "B", which is 0.25 of an inch to the exterior side (west) of the map border.

The line labelled DEEP-MINED 1 is a typical example of a line that intersects an outcrop line. Digitizing begins at point "C" where the deep-mined line meets the outcrop line. The identifier is added, and the line is digitized in a counterclockwise direction to point "D". The

5

line is continued 0.25 of an inch over the outcrop line and is closed
back to point "C" (the closure is represented by dashed lines). This
digitizing method is used often when an overlay contains strip- and
deep-mined lines.

The line labelled DEEP-MINED 2 is an example of a line that intersects
another line where it is impossible to close it back to the beginning
point of the line. In the example, the line is begun at point "E" at
least 0.25 of an inch on the exterior side (south) of the map border. The
identifier is added and the line is digitized in a counterclockwise
direction to point "F". As is it impossible to close the line back to
the beginning point of the line, the line is closed to the border
(continuing at least 0.25 of an inch outside the map border). The dashed
lines represent the closure method.

Figure 4 is an enlargement of a small area of the overlay shown in figure
3 and is used to show the method of digitizing a closely strip- and deep-
mined area. The outline of the strip-mined area is begun at point "A"
and is digitized according to the preceding example of DEEP-MINED 1 to
point "D". Once again, the dashed lines show the closure procedure.
The outline of the deep-mined area is begun at point "B" where it inter-
sects the strip-mined line. The line is not continued 0.25 of an inch
over the strip-mined line. The line is continued over the strip-mined
line at point "C", but it is contained within the area between the strip-
mined line and the outcrop line. It is closed back to point "B" (note
dashed lines).

Detailed sections of a quadrangle that contain excessively complicated
line data may be generalized to minimize digitizing time and reduce
computer costs of the resource-assessment program, GARNET. Figure 5 is
an example of a detailed deep-mined area. The heavy solid line represents
the generalized line that is digitized for the area. Only the generalized
line as shown in the example is digitized. None of the blocks within
the deep-mined area is digitized.

Conclusion

The format requirements outlined above may change as the coal-resource
assessment program, GARNET, and other graphics programs are modified to
meet user demands. Subsequently, the digitizing procedures may change to
meet the modified format requirements. We recommend that you contact
the staff of the National Coal Resources Data System (U.S. Geological
Survey, 956 National Center, Reston, VA 22092) before beginning a digitizing
program of your own for use in the NCRDS coal-resource assessment program.

FIGURE 3—Outcrop lines and strip- and deep-mined lines for an actual coal bed. Letters are explained in text.
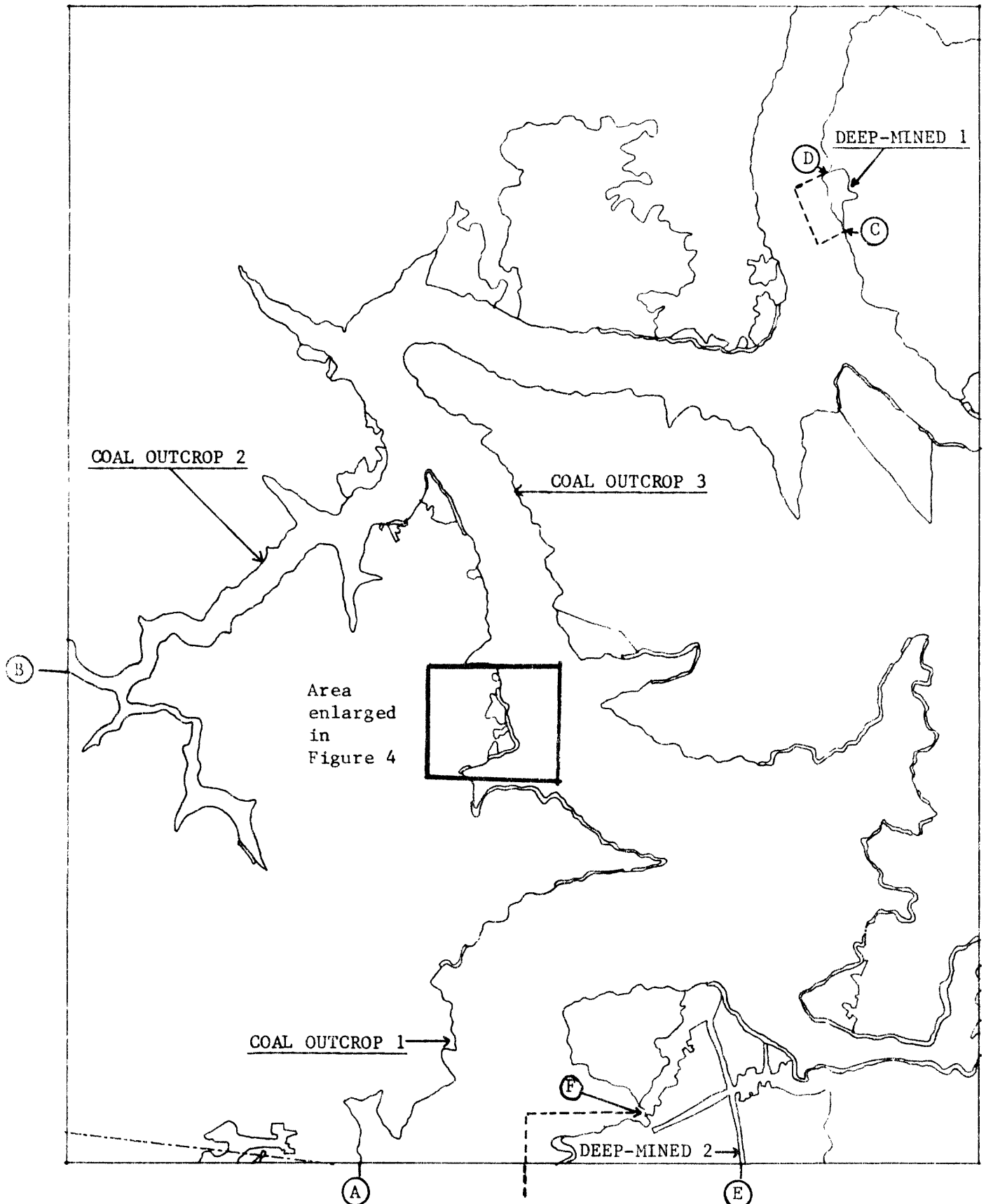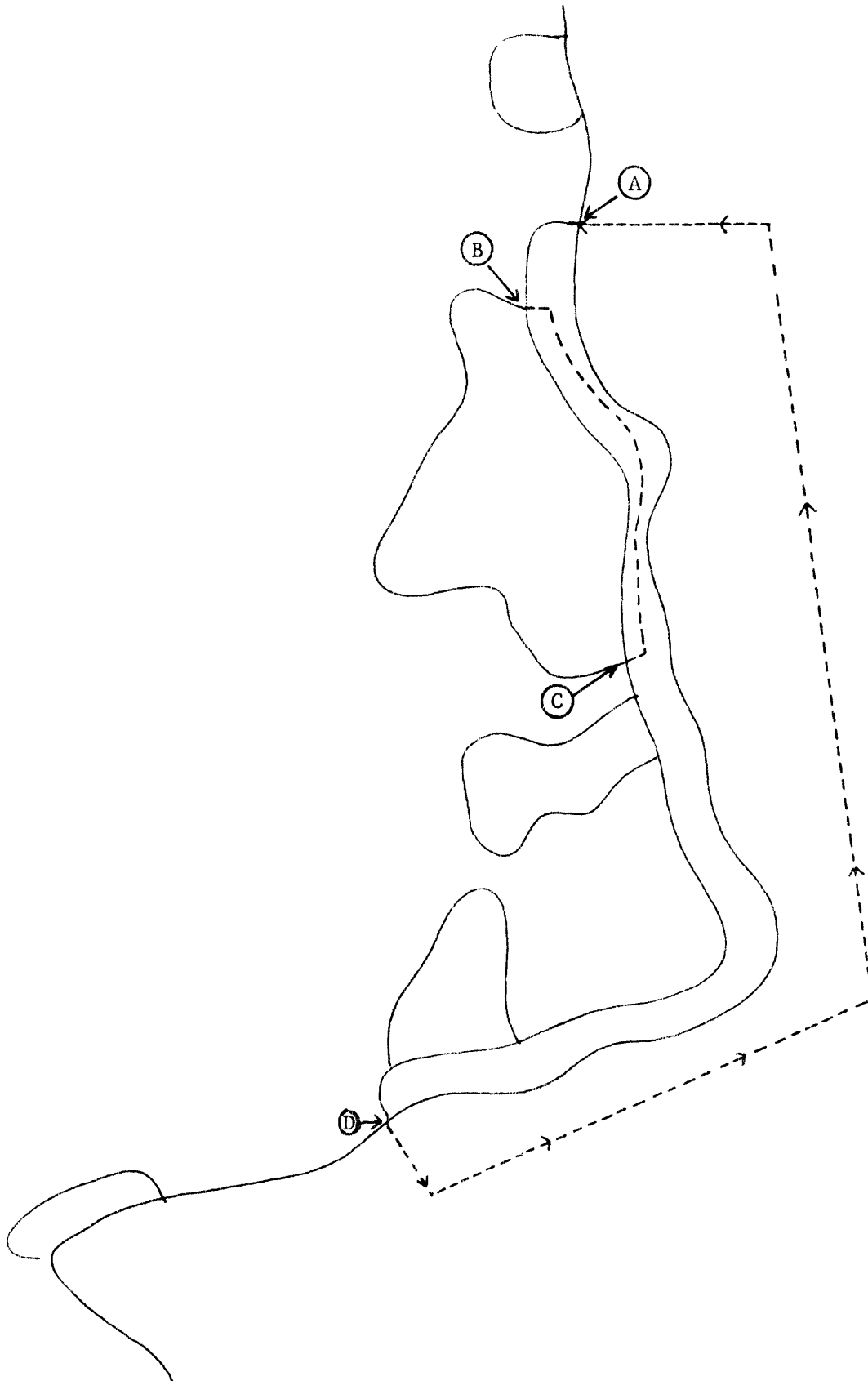
FIGURE 4--Enlargement of part of figure 3.
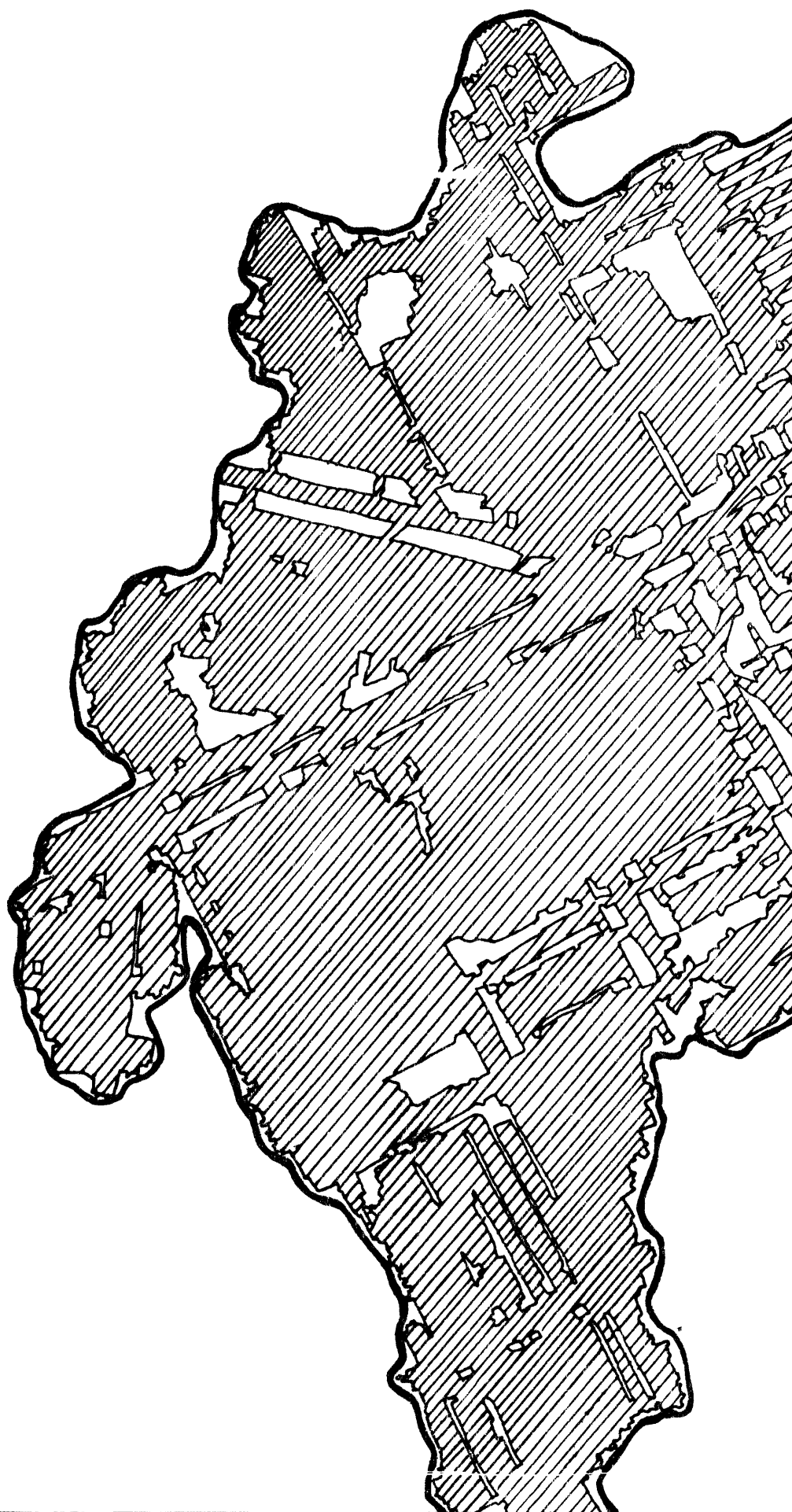Letters are explained in text.

FIGURE 5.--Detailed map of deep-mined area.
Only the heavy solid line is digitized for the area.

9