Understanding Universal Exchange Formats

J. Raul Ramirez
Center for Mapping, The Ohio State University, Columbus, OH 43212

ABSTRACT: During the last decade, the use of computer-aided mapping systems (CAMS) has increased tremendously within the cartographic community. This, together with the fact that so many different private and commercial CAMS have been used, has generated the need for the design of Universal Exchange Formats (UEF). These UEF are digital formats designed to facilitate the exchange of cartographic data among the different databases and computer systems. In this paper, UEFs are studied from the theoretical and the practical point of view in order to understand their need, usefulness, and limitations. A set of formulas is introduced to help the reader in deciding when a UEF should be used.

INTRODUCTION

THE ADVENT OF digital cartographic products has revolutionized the field of cartography by introducing new and useful products. It has also generated an almost unmanageable amount of data and new problems. This stems partially from the fact that digital cartographic data are generated by digitizing old, conventional, hard-copy maps; by photogrammetric stereocompilation; by satellite means; and by all kinds of thematic applications (Ramirez, 1988). Two of the most important new problems are the evaluation of the quality of the digital data and the exchange of digital information. In this paper only the problem of digital data exchange is addressed. The interchange of digital cartographic information is a major problem due to the different structures of the databases used by the private and commercial computer-aided mapping systems (CAMS). Moellering (1983) characterized the problem:

> Each year millions of dollars are being spent to reorganize, reformat, process, verify, and check digital cartographic databases that one agency or organization obtains from another.

The problem of different structures can be described as the problem of different sets of data having exactly the same information but stored so that it is impossible to recognize or to directly use that information from any CAMS other than the one used to generate the information (Ramirez, 1988). Among the factors that could produce different database structures, the following are the most common: number of files used to store the data, format of the files, coordinate system, and precision of the data.

THE PROBLEM OF DATA EXCHANGE

In order to transfer digital cartographic data between two different CAMS, it is necessary to perform a database transformation from the original to the target system (Figure 1). A database transformation is the set of operations that reads and interprets the original database and writes it into another file with the database structure of the target system. The database transformation is performed by computer programs called TRANSLATORS.

It is necessary to have a deep understanding of the structures of the original and the target databases in order to perform a database transformation. This is a major problem because the information about database structures is considered privileged and/or confidential by most developers. An alternative action to this would be to perform the database transformation from the original to an intermediate database and then to perform a second transformation from the intermediate to the target database (Figure 2).

This is a viable alternative because the structure of the intermediate database is always well known. Therefore, there are no major problems in developing the translator from the original, because all systems have some way to extract information to the intermediate database format. Also, there usually are translators available to perform the transformation from the original to the intermediate database. Intermediate databases have been introduced by many developers of CAMS and CAD packages as the channel to transfer data in and out of their database structures. Usually the intermediate database is an ASCII file with a well-documented structure. The transformation from the intermediate to the corresponding database structure is done sometimes without user intervention by computer programs provided by the CAMS or CAD developers. The best known example of this kind of exchange formats is the .DXF file of Autodesk's Autocad.

UNIVERSAL EXCHANGE FORMATS

A major problem with the intermediate databases described above is their CAMS/CAD dependency. One translator is needed to send and one is needed to receive data per CAMS or CAD system. Because there are more than 50 systems in use in the United States, a particular user needs to have approximately 100 translators to interchange data with all of them.

For some time now people have been aware of the increasing problems of data exchange and a great deal of money and effort have been dedicated to develop an efficient solution. As a re-

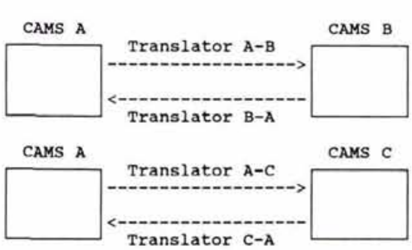

FIG. 1. Schematic representation of direct data exchange between two CAMS.

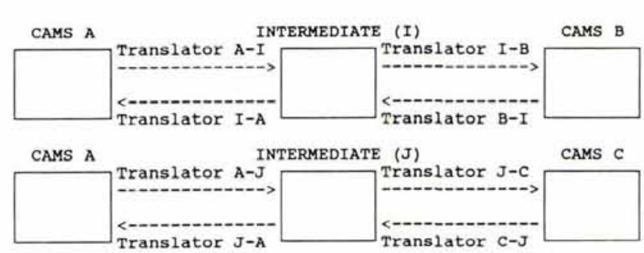

FIG. 2. Schematic representation of data exchange using CAMS/CAD intermediate database.

0099-1112/91/5701–89$03.00/0

sult, the idea of a Universal Exchange Format (UEF) has gained acceptance in the mapping community. Following Ramirez (1988) in this discussion, the term FORMAT is used to indicate a specific exchange method or some general guidelines for data exchange.

A UEF is a well-documented (generally available) structure for the storage of digital cartographic data. It is a universal intermediate database structure. Universal means that any CAMS/CAD could access that structure to store and retrieve exchange data; intermediate means that only two translators per system are needed. This includes one to transform the original database into the universal database and a second to transform the universal database into the target database (Figure 3).

Some of the best known universal exchange formats are the Initial Graphic Exchange Specification (IGES) developed by the U.S. National Bureau of Standards (1983), and the Space Data Transfer Standard (SDTS) developed by U.S. Geological Survey (1990).

There are two CAD related intermediate formats that can be considered as universal exchange formats because of the wide use of their databases: the Standard Interchange Format (SIF) from Intergraph Corporation (Intergraph Corporation, 1986) and the Autocad's Drawing Interchange File (DXF) from Autodesk, Inc. (Autodesk Inc., 1986). These database formats are the "de facto" standard for collection of geometric mapping information, and more and more CAMS are using them for drafting, editing, and database storage.

USING UNIVERSAL EXCHANGE FORMATS

Figures 1, 2, and 3 show the different alternatives to data exchange for individual files. In Figure 1, data exchange takes place between databases A and B and A and C, and this will be called Case I. This action is accomplished by applying, in each case, a single transformation to the database A (A to B in the first case and A to C in the second), for a total of two transformations.

In Figure 2, data exchange takes place between databases A and B and A and C, and this will be called Case II. This is accomplished by applying two transformations to the database A (A to I and I to B in the first case and A to J and J to C in the second) for a total of four transformations.

In Figure 3, data exchange between A and B and A and C, as shown in the upper window in Figure 3, is accomplished by applying a total of three transformations (A to U and then U to B and U to C). This will be called Case III.

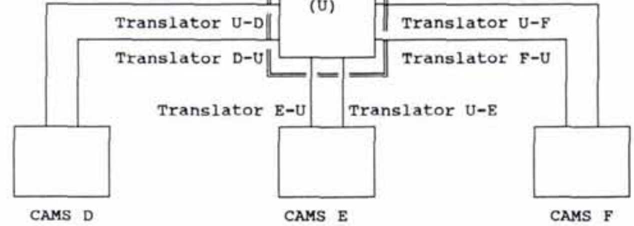

FIG. 3. Schematic representation of data exchange using the Universal Exchange Format approach.

Several obvious questions are encountered with respect to exchange methods. They include: "When should universal exchange formats be used?" (assuming any of the three alternatives can be selected) and "When should the other alternatives be used?" These questions can be answered from the point of view of the number of transformations to be applied. The following formulas can be used:

To only send or receive data from one CAMS to m other CAMS:

Case I: $S = m$ (1)
Case II: $S = 2m$ (2)
Case III: $S = m + 1$ (3)

where S is the total number of tranformations.

To send and receive data from one CAMS to m other CAMS:

Case I: $SR = 2m$ (4)
Case II: $SR = 4m$ (5)
Case III: $SR = 2\,(m + 1)$ (6)

where SR is the total number of transformations.

To send and receive data among n ($n = m + 1$) CAMS:

Case I: $T = (m + 1)m$ (7)
Case II: $T = 2(m + 1)m$ (8)
Case III: $T = 2(m + 1)$ (9)

where T is the total number of transformations. These formulas have been used to build Tables 1, 3, and 5.

Table 1 shows the number of database transformations that need to be performed to send cartographic information from one CAMS to m different CAMS, or to receive cartographic information from m different CAMS by direct exchange (Case I), by intermediate format (Case II), and by Universal Exchange Format (Case III).

The number of transformations in data exchange is an important factor because each additional tranformation increases the cost of data exchange. The cost of a transformation is related, among other things, to the amount of cartographic data to be exchanged, the structure of the original and target databases, the efficiency of the translators involved, the hardware used, and various other factors. Table 2 shows the cost ratios Case II/Case I and Case III/Case I assuming the same cost for all tranformations.

Table 2 indicates that data exchange, using an intermediate format, is always twice as expensive as direct data exchange. On the other hand, data exchange using a universal format is twice as expensive when data are sent or received only from another CAMS. Yet its cost is closer (although always higher) than the cost of the direct data exchange when the number of CAMS increased.

Table 3 shows the number of transformations for Cases I, II, and III when data are sent and received from one CAMS to

TABLE 1. TRANSFORMATIONS (S) TO SEND OR RECEIVE CARTOGRAPHIC INFORMATION FROM ONE SYSTEM TO M SYSTEMS

m	Case I	Case II	Case III
1	1	2	2
2	2	4	3
3	3	6	4
4	4	8	5
5	5	10	6
10	10	20	11
20	20	40	21
30	30	60	31
40	40	80	41
50	50	100	51

TABLE 2. COST RATIOS TO SEND OR RECEIVE CARTOGRAPHIC INFORMATION FROM ONE SYSTEM TO M SYSTEMS. CASE II/CASE I IS EQUAL TO 2.0 FOR ALL VALUES OF M.

m	Case III/Case I
1	2.000
2	1.500
3	1.333
4	1.250
5	1.200
10	1.100
20	1.050
30	1.033
40	1.025
50	1.020

TABLE 3. TRANSFORMATIONS (*SR*) TO SEND AND RECEIVE CARTOGRAPHIC INFORMATION FROM ONE SYSTEM TO M SYSTEMS.

m	Case I	Case II	Case III
1	2	4	4
2	4	8	6
3	6	12	8
4	8	16	10
5	10	20	12
10	20	40	22
20	40	80	42
30	60	120	62
40	80	160	82
50	100	200	102

TABLE 4. COST RATIOS TO SEND AND RECEIVE CARTOGRAPHIC INFORMATION FROM ONE SYSTEM TO M SYSTEMS. CASE II/CASE I IS EQUAL TO 2.0 FOR ALL VALUES OF M.

m	Case III/Case I
1	2.000
2	1.500
3	1.333
4	1.250
5	1.200
10	1.100
20	1.050
30	1.033
40	1.025
50	1.020

several other CAMS, and Table 4 shows the cost ratios for Table 3.

Table 4 shows the same ratios as Table 2, and, therefore, the same considerations applied in the discussion of that Table are valid here.

Table 5 shows the number of transformations for Cases I, II, and III when data are exchanged among all CAMS. Each system must be able to exchange (send and receive) cartographic data with all other systems. Table 6 shows the cost ratios for Table 5.

Table 6 shows the cases in which universal exchange formats are more useful: i.e., when cartographic data are sent and received (Case III) among a minimum of three different CAMS. As shown by this table, the usefulness of Universal Exchange Formats increases with the number of CAMS involved. It is important to point out that, when three $(m+1)$ systems are involved, the direct data exchange (Case I) and the Universal Exchange Format (Case III) require exactly the same number of transformations.

TABLE 5. TRANSFORMATIONS (T) TO EXCHANGE CARTOGRAPHIC INFORMATION AMONG (M + 1) SYSTEMS

m	Case I	Case II	Case III
1	2	4	4
2	6	12	6
3	12	24	8
4	20	40	10
5	30	60	12
10	110	220	22
20	420	840	42
30	930	1860	62
40	1640	3280	82
50	2550	5100	102

TABLE 6. COST RATIOS TO EXCHANGE CARTOGRAPHIC INFORMATION AMONG (M + 1) SYSTEMS. CASE II/CASE I IS EQUAL TO 2.0 FOR ALL VALUES OF M

m	Case III/Case I
1	2.000
2	1.000
3	0.667
4	0.500
5	0.400
10	0.200
20	0.100
30	0.067
40	0.050
50	0.040

The analysis of the above tables indicates that the exchange method should be selected depending on the scope of the data exchange. If Cases I, II, and III are available, direct exchanges must be selected to send cartographic data from a single specific CAMS to several CAMS or to receive data from several CAMS to a single CAMS. This selection is based on the minimum number of transformations needed. To exchange (send and receive) cartographic data among several CAMS (more than three), the UEF approach must be used. Again, the selection is based on the number of transformations.

There are several additional factors to be considered in selecting the exchange method. One factor is the cost of developing or acquiring the database translators. As previously indicated in the direct exchange method, two translators per target CAMS are needed to send and receive information. On the other hand, only a total of two translators are required in the Universal Exchange Format method, regardless of the number of target CAMS involved. Therefore, it is more expensive to develop all the translators needed for the direct exchange than for the Universal Exchange Format approach, assuming that the cost of development per translator is the same.

Another factor to be considered is the number of files to be translated to a particular database format. The greater the number of files, the greater the number of total database transformations to be performed. For example, if there are f files to be translated, then the total number of transformations can be obtained from Tables 1, 2, and 3 by multiplying each entry (e) under Case I, Case II, and Case III by f. If the cost of each transformation is c (assuming the same cost for all transformations), then the total cost for each case is obtained by multiplying efc. The cost differences for Tables 1, 3, and 5 are given by

Table 1:
$$\text{Case II} - \text{Case I} = mfc \quad (10)$$
$$\text{Case III} - \text{Case I} = fc \quad (11)$$
Table 3:
$$\text{Case II} - \text{Case I} = 2mfc \quad (12)$$
$$\text{Case III} - \text{Case I} = 2fc \quad (13)$$

Table 5: Case II − Case I = $m(m+1)fc$ (14)
Case III − Case I = $[2 - m(m-1)]fc$ (15)

Whenever the cost difference for translations is greater than the program development cost, it is better to select the direct exchange method. This is because program development is a one-time cost, but the translation is a cumulative cost. Hence, cartographic data exchange among more than three systems is always less expensive if the Universal Exchange Format approach is used.

Finally, a note of caution must be given with respect to the assumption of similar costs and execution times in the development of Equations 10 to 15, and Tables 2, 4, and 6. These assumptions are just approximations. If needed, more accurate Equations can be derived from Tables 1, 3, and 5.

CONCLUSIONS

Selection of a data exchange methodology is a major component of any map production system. The formulas and tables provided here can be used to determine the best method for data exchange in a given case. Universal Exchange Formats are the best choice for data exchange under the proper conditions. However, the other methods of data exchange are better under circumstances where the general capabilities of UEF may not meet the requirements of a given project.

REFERENCES

Autodesk Inc., 1986. *AutoCAD Reference Manual*, Oakland: Command Technology Corp.

Federal Interagency Coordinating Committee on Digital Cartography, 1986. *Federal Geographic Exchange Format*, FICCDC.

Moellering, H., 1983. Issues Facing the Development of Digital Cartographic Data Standards: Background and Introduction, *Issues in Digital Cartographic Data Standards*, No. 3.

Intergraph Corporation, 1986. *Standard Interchange Format (SIF) Command Language Implementation Guide*, Intergraph Corporation Document No. DIXD7330, Huntsville, Alabama.

Ramirez, J. R., 1988. *A Map Representation Theory for the Evaluation of Digital Exchange Formats*, Report No. 389, Department of Geodetic Science and Surveying, The Ohio State University.

U.S. Geological Survey, 1990. *The Spatial Data Transfer Standard*, Draft: January 24, 1990.

U.S. National Bureau of Standards, 1983. *Initial Graphic Exchange Specification (IGES) Version 2.0*, U.S. Department of Commerce, National Technical Information Center Publication no, PB83-127448, Washington D.C.

(Received 22 May 1989; accepted 22 August 1989; revised 16 April 1990)

Forthcoming Articles

Patrick A. Agbu and *Egide Nizeyimana*, Comparisons between Spectral Mapping Units Derived from SPOT Image Texture and Field Soil Map Units.

F. J. Ahern, J. Sirois, W. D. McColl, R.P. Gauthier, T. T. Alfoldi, W. H. Patterson, and *T. A Erdle*, Progress Toward Improving Aerial Defoliation Survey Methods by Using Electronic Imagers.

Eric W. Augenstein, Douglas A. Stow, and *Allen S. Hope*, Evaluation of SPOT HRV-XS Data for Kelp Resource Inventories.

Emmanuel P. Baltsavias and *Dirk Stallmann*, Trinocular Vision for Automatic and Robust Three-Dimensional Determination of the Trajectories of Moving Objects.

David G. Barber and *Ellsworth F. LeDrew*, SAR Sea Ice Discrimination Using Texture Statistics: A Multivariate Approach.

D. S. Bhargava and *Dejene W. Mariam*, Effects of Suspended Particle Size and Concentration on Reflectance Measurements.

John M. Briggs and *M. Duane Nellis*, Seasonal Variation of Heterogeneity in the Tallgrass Prairie: A Quantitative Measure Using Remote Sensing.

Doug C. Brockelbank and *Ashley P. Tam*, Stereo Elevation Determination Techniques for SPOT Imagery.

Joseph P. Burns, Survival through Ethical Conduct.

Pat S. Chavez, Jr., Stuart C. Sides, and *Jeffrey A. Anderson*, Comparison of Three Different Methods to Merge Multiresolution and Multispectral Data: Landsat TM and SPOT Panchromatic.

Jean Chorowicz, Jean-Yves Preard, Richard Guillande, Claude-Roger Morasse, Daniel Prudon, and *Jean-Paul Rudant*, Dip and Strike Measured Systematically on Digitized Three-Dimensional Geological Maps.

P. K. Chua, A. H. W. Kersley, J. K. Ridley, W. Cudlip, and *C. G. Rapley*, The Determination and Use of Orthometric Heights Derived from the Seasat Radar Altimeter over Land.

Warren B. Cohen, Response of Vegetation Indices to Changes in Three Measures of Leaf Water Stress.

Warren B. Cohen, Chaparral Vegetation Reflectance and its Potential Utility for Assessment of Fire Hazard.

Liping Di and *Donald C. Rundquist*, Instantaneous Three-Channel Image Processing to Facilitate Instruction.

D. E. Flittner and *P. N. Slater*, Stability of Narrow-Band Filter Radiometers in the Solar-Reflective Range.

Sanjib K. Ghosh and *Zhengdong Shi*, Evaluating Dynamic Performance of an Analytical Plotter.

Joshua S. Greenfeld, An Operator-Based Matching System.

Nils N. Haag, Michael H. Brill, and *Eamon B. Barrett*, Invariant Relationships in Side-Looking Synthetic Aperture Radar Imagery.

T. K. Koo and *Y. B. Aw*, A Three-Dimensional Visualization Approach to Traffic Accident Mapping.

James LaGro, Jr., Assessing Patch Shape in Landscape Mosaics.

Andre G. Lareau, Flight Performance of an Airborne Minefield Detection and Reconnaissance System.

A. J. Naftel and *J. C. Boot*, An Iterative Linear Transformation Algorithm for Solution of the Collinearity Equations.

James P. Ormsby and *Dorothy K. Hall*, Spectral Properties of Fog Over the Malaspina Glacier, Alaska in Comparison to Snow, Ice, and Clouds.

Derek R. Peddle and *Steven E. Franklin*, Image Texture Processing and Data Integration for Surface Pattern Discrimination.

Mohammed Yousef H. T. Qari, Application of Landsat TM Data to Geological Studies, Al-Khabt Area, Southern Arabian Shield.

Eric J. M. Rignot, Ronald Kwok, John C. Curlander, and *Shirley S. Pang*, Automated Multisensor Registration: Requirements and Techniques.

Steven A. Sader and *J. Chris Winne*, Digital Image Analysis Hardware/Software Use at U. S. Forestry Schools.

Morris M. Thompson, Publication Aspects of Ethics in Photogrammetry.

S. D. Wall, T. G. Farr, J -P Muller, P. Lewis, and *F. W. Leberl*, Measurement of Surface Microtopography.