A LINEAR GEOGRAPHICAL CODE FOR MANAGEMENT INFORMATION SYSTEMS

Jiri Kallab
Univac Federal Systems Div.
2121 Wisconsin Ave. NW
Washington, D.C. 20007

In making the data base for a management information system, it is useful to code geographical data to correspond geographically with the real physical world.

Since joining the Univac Div. of Sperry Rand Corp. in 1953, Jiri Kallab has worked with management and foreign operations and development. Mr. Kallab majored in Business Administration, and has studied economics, market research, and international trade and development.

The success of business operations in our competitive economy often depends upon the ability to correctly and rapidly accumulate and analyze masses of data directly related to geographical distribution and concentrations. Marketing activities of both perishable and durable goods are concerned with the location of key distribution centers and the means of reaching potential users by the shortest, as well as alternative, routes. Manufacturers wish to analyze the relationship of resources to a new plant; in personnel recruitment the objective is to locate areas of concentration of certain classes of skilled labor. Economists and statisticians in every major field of human endeavor are pressing for improved results which rely heavily on geographical data.

Real-time computer systems, which represent huge networks, require the consideration of geography. In the original layout, in collecting data, in arranging data in the data-base, and in organizing the output logically, we need to consider adjacency.

This article reviews briefly the existing tools in market and economic analysis, and describes a coding system for linear geographical location. The article also recommends ways and rules for implementing the linear geographical code in management information systems.

Existing Tools in Market and Economic Analysis

In any management analysis related to economics and marketing, two criteria have to be considered: (1) the geography of the market, and (2) the industries involved.

Serious problems arise from the fact that market data are rarely available in a form which clearly defines the relationships between various levels of geographical subdivisions and adjacencies. They are most often reported only by general geographical references and broken down alphabetically in an arrangement difficult to manipulate. Such data have to be reclassified and plotted laboriously before they can be made available to management in usable form. Consequently, they are often furnished too late for practical use. Certainly, data so arranged cannot take full advantage of electronic data processing equipment for geographical analysis and plotting.

Let us briefly review the events in a corporation to illustrate the advisability of storing data in a two-dimensional geographical sequence which will be readily available for data processing and analysis.

In a complex manufacturing or distributing enterprise, market data are reflected over and over again in daily, hourly and sometimes minute operations. Function and location seem to be just two sides of the same coin.

The changing economic environment in which a company operates has to be compared frequently to forecasts which have served as a basis for production schedules and marketing quotas. No matter how reliable the indicators are, unless the trends are studied regionally, it is difficult to identify the areas requiring special attention and action. The indicators vary according to the category of industry or business. In an industry producing and marketing consumer goods, the emphasis will be on population data, income groups, and telephone households. An industry manufacturing computers and business equipment will rely mainly on survey data furnished periodically by the company's field organization. Industrial equipment, similarly, has to depend on its own marketing organization, and on physical surveys, to detect the need for capital investment goods. In each case, location and adjacency are constants which do not change in any of these studies.

Development and growth are intrinsically related to geographical locations, and any activity that affects them has to be superimposed on the area where it occurred. Resources, labor market, engineering skills, demand, production facilities, capital and finance have to be studied and considered in relation to locations and adjacencies. Channels of distribution, transportation, warehousing policies, and sales organizations, too, have to fan out from the physical location. To a considerable degree, location will influence the pricing policies, impact of competition, and the interrelation of price and profit.

A New Concept

The linear location coding system, the principles of which are outlined briefly in the following paragraphs, has been tested in volume applications, and offers considerable improvement over methods now used in the aggregation, processing and analysis of data. The concept is based on the realization that coordinates of latitude and longitude are an alignment of physical spaces of certain size, that these can be arranged into a line of two-dimensional spaces, forming a grid, and that each grid can be similarly subdivided.

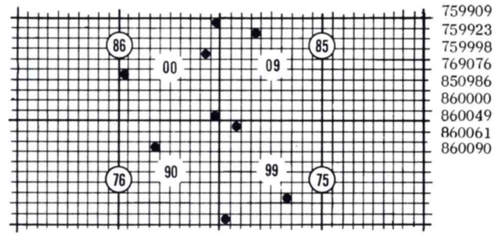
Each level of these units has to be identified by a two-digit number, the first digit signifying distance from South to North, and the second the distance from East to West, in units of that level.

If data are grouped in the data base in a form resembling as closely as possible the physical world, then advanced mathematical methods for the solution of transportation problems, computations of the center of gravity, various statistical methods and correlations, etc., can all be used to best advantage. The value of the Input/Output analysis can be enhanced. The economic profile of a community or region can be obtained either by geography, or by the category of industry, still showing the important geographical relationships. Also, a number of significant indicators and ratios, of interest for the management information system, can be computed.

The learning of the principles of decimal spacial interpolation is essential for the multiple uses of the linear location code in geographical coverage.

The concept of the linear location code is illustrated in Figure 1.

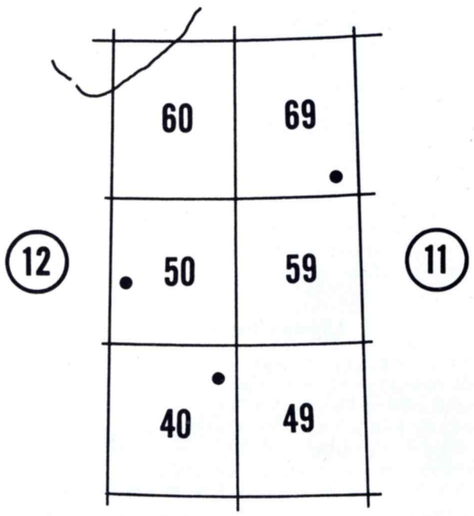
Figure 1
Linear Location Code



There is an alignment of horizontal and vertical identifications throughout units of each level. This is a very important feature and must be followed throughout the entire grid.

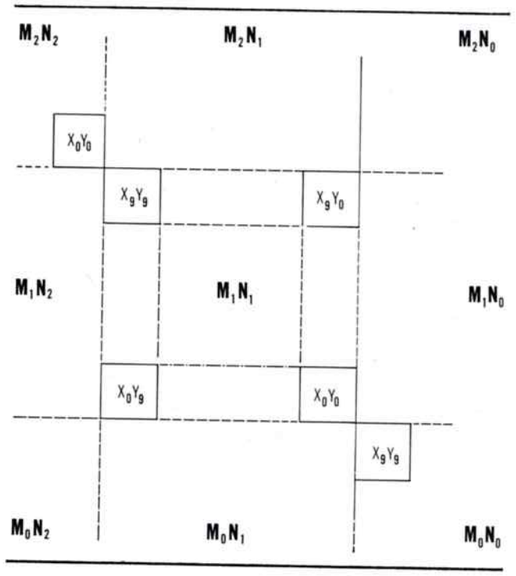
In world-wide or nation-wide applications, a space of one degree of latitude and longitude seems advisable as an intermediate level and then major emphasis would have to be placed upon cities as smallest units. The term "level", as used in this paper, is equivalent to the "power" of a numerical system. (See Diagram 1.) Thus, the codes of typical cities might appear as shown in Figure 2.

Figure 2
Linear Code in National Coverage



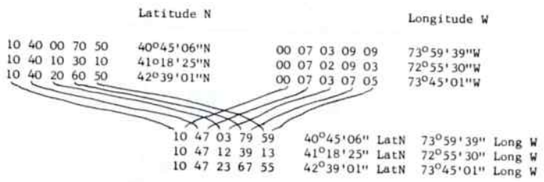

```
11 69 01  Ft. Wayne
12 40 71  Indianapolis
12 50 49  Lafayette
```

Diagram 1 — Linear Location Code Concept



If source data is in the form of latitude and longitude, the two numbers can be consolidated into the linear code number by a computer program as shown in Figure 3, and then printed for visual comparison or geographical tabulation.

Figure 3
Conversion to the Linear Code of Locations Identified by Latitude and Longitude



Examples of world-wide linear code, developed from latitude and longitude identifications, are shown in Figure 4. (See also Hemispheres Map, Diagram 2.)

Aligning Irregular Areas

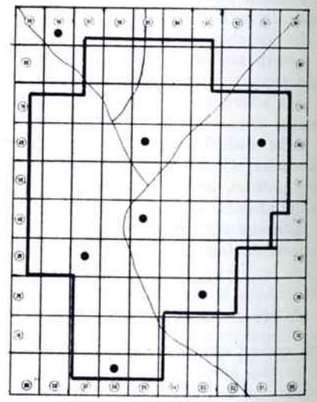
In some business applications it may be necessary to align and code irregular areas, such as states, counties, and even postal zones and census tracts. Such applications of the linear code are feasible, and offer advantages over an arbitrary numerical or alpha-numerical system, as long as the coding has the appearance of a straight-line grid.

For the purposes of a more detailed identification and expansion of the code to include unique account numbers, locations within cities, unless representing a multitude of points, can be identified by spacial interpolation. While in world-wide or nation-wide systems, the city was identified by a single code pinpointing the major section of the town, in this code, the outline of the city would be projected into a two-digit or four-digit grid, to indicate at least the relative position of data, coded within this grid. A variety of meaningless numerical customer codes can thus be replaced by a linear code in which every two-digit number has a geographical significance. (See Figure 5.)

Figure 4
World-Wide Linear Codes Converted from Latitude and Longitude Identification

34	54	30S	58	30 41W	0055580591 BUENOS AIRES
14	54	12S	172	42 59W	0178521702 PAGO PAGO
34	23	37S	151	42 12E	0252586310 SYDNEY
27	48	18S	28	12 56E	0366212708 JOHANNESBURG
33	45	10N	84	23 37W	1038347359 ATLANTA
38	14	47N	85	45 49W	1038852756 LOUISVILLE
40	45	6N	73	59 39W	1047037959 NEW YORK
41	18	25N	72	55 30W	1047123913 NEW HAVEN
42	39	1N	73	45 1W	1047236755 ALBANY
29	56	53N	90	4 10W	1120909057 NEW ORLEANS
33	27	12N	112	4 28W	1132324057 PHOENIX
34	3	15N	118	14 28W	1132480254 LOS ANGELES
34	44	42N	92	16 37W	1130427258 LITTLE ROCK
35	28	26N	97	31 4W	1130574572 OKLAHOMA CITY
37	41	30N	97	20 16W	1130776394 WICHITA
47	14	59N	122	26 15W	1143722454 TACOMA
36	24	18N	139	54 28E	1234604019 TOKYO
41	42	42N	28	18 15E	1346117710 ISTAMBUL

Figure 5
Decimal Spacial Interpolation of Urban Area



In the metropolitan and urban application of the linear code, the metropolitan code represents an extension of the world-wide code. This extension is used when a multitude of points have to be identified and converted from master files listing street addresses. Semi-automatic methods are available for this conversion. The coordinates are replaced by natural arteries, such as expressways, and the size of the spaces and districts is dictated by the physical pattern of minor units, i.e., city blocks. It is evident that in some of the districts or spaces, the capacity of the code (00 to 99) is not exhausted. This same feature acts as a balancing factor and makes possible the mandatory overall alignment of the grid with the coordinates. (See Figure 6.)

Techniques have been developed which make conversions from arbitrary systems to linear code in urban applications simple and feasible wherever there are repetitive sequential or random references to the same locations.

Diagram 2 — Map of the Hemispheres Showing Linear Masses

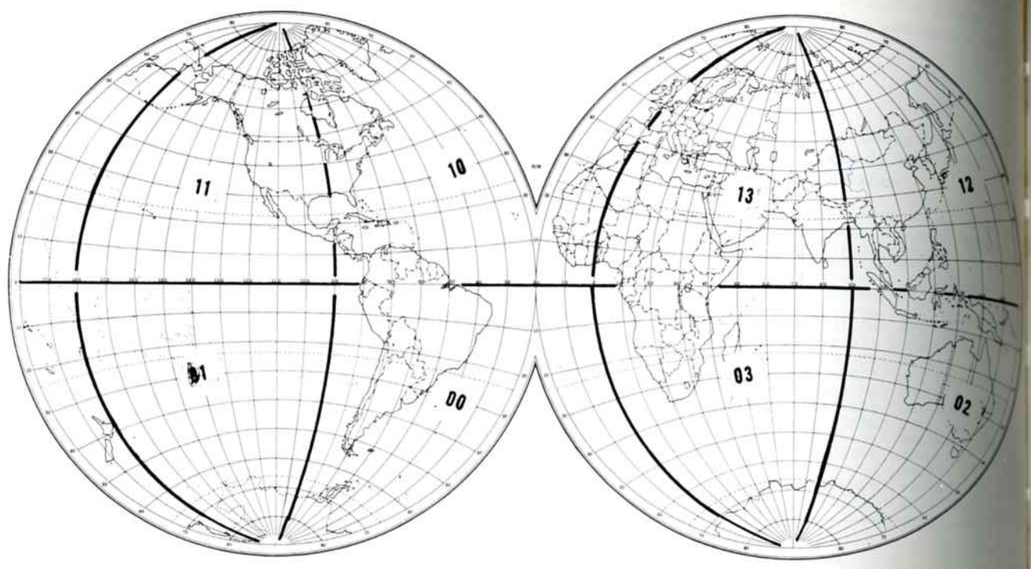
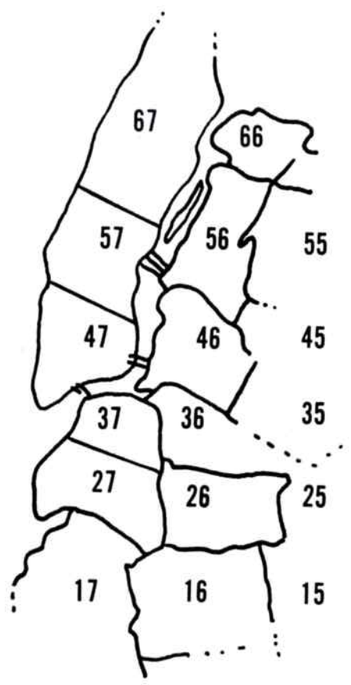


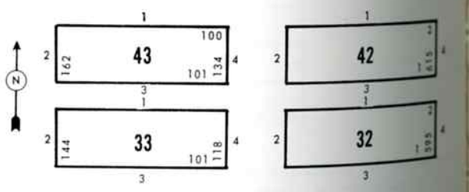
Figure 6
Linear Layout of a Metropolitan Area



Selecting Proper Maps

Great care must be given to the selection of proper maps. It is mandatory that, in metropolitan coverage, the coding maps show the number of the first building on each blockside. (See Figure 7.)

Figure 7
Blocks and Blocksides in Metropolitan Areas



The procedure to follow in developing coding maps is to start with the major divisions, using as boundary lines the major arteries, keeping in mind the coordinates of latitude and longitude, and considering the Northern blockside as "1". Starting at the North-East corner, the blocksides are recognized counterclockwise as 1 (N), 2 (W), 3 (S), 4 (E). Each blockside is considered to have 100 units, irrespective of length, and the street addresses (buildings) are converted to sequenced code according to a simple formula, which permits interpolation of any additional addresses. Within the buildings, customers are coded by alpha-numeric scales to create unique identifications and room for expansion. The term "alphanumeric" code is used to represent a code which has been initially arranged into alphabetical sequence and then assigned numerical codes in sequence, leaving room for the interpolation of new accounts. Two digits may be sufficient for the alpha-numeric code, except in the case of large buildings, where three digits are allowed for the alpha-numeric code and only one digit for the building (door-code). The rules, applicable to physical variations such as split blocks, combined blocks, parallel or opposite street numbering, are simple and easy to follow.

The linear arrangement of records by buildings, blocksides, blocks, and districts, against alphabetical dispersion by names is evident in Figure 8. The meaningful linear code, as opposed to an arbitrary customer identification, consists of two-digit codes for space (57), district (35), block (62), one-digit code for the blockside (1, 2, 3, 4), two-digit code for the building, and two-digit unique alpha-numeric customer number.

Applications of the Linear System

How does a linear geographical system affect the economics of individual companies, or even the national economy? Many people are active rearranging data, dispersed by men and machines, to the logical geographical relationships on which management decisions have to be made or actions taken. This process takes much time, and often timeliness is lost.

The increased effectiveness of company operations, and the new uses to which market data can be put, should easily absorb the existing skills. The emphasis in the processing of market data would be shifted from historical files to geographical locations and adjacencies; from inability to analyze and correct variations to constructive solutions of problems and expansion of business.

Testing the Reliability of an Indicator

A typical example of the application of the linear system to the optimization of marketing is the case of a corporation which was anxious to test the reliability of an indicator, used to determine potentials and quotas in durable goods sold to industries, against other possible indicators. The conventional indicator was available only by industry classification; the tentative indicators, to be used in the correlation, were collected for each prospect by the field organization through a continuous census.

The enumeration cards, used in the continuous census, were created originally from questionnaires, and included a number of significant characteristics, such as standard industry code, Dun and Bradstreet rating, attitude, competition equipment, and other data.

Through a computer program it was possible to include additional indicators, and test the correlation geographically by industry, as well as overall. Once the user was satisfied as to the reliability of the selected indicators, the prospect records could be weighed, and accumulated in linear sequence by buildings, blocks, districts and spaces. Since the linear code assured shortest distance from prospect to prospect — overall as well as by industry — the lack of balance between sales territories and reservations was easily spotted. The sales territories were then equalized easily. The field organization demonstrated confidence in the new approach, and in many areas, the sequential linear territories replaced previous arbitrary allocations of prospects and reservations.

The Linear Master File

The linear master file, used in the establishment of potentials, can serve as a vehicle for recording performance, such as frequency of calls, distance factors, and sales results. Potentials, performance and booked sales are then accumulated by small and large linear geographical units, overall or industry totals are established, and variances later analyzed down to their original causes.

Modifying Territories

In another application, union opposition prevented modification of territories, assigned on the basis of seniority, and used in the computation of commissions. Actually, three territorial systems were used simultaneously in the same large metropolitan area. By identifying the points at which each system varied from the linear sequence used in computer processing, it was possible to replace the cumbersome manual computation by an electronic procedure, without modification of the union contract. (See Figure 9.)

In another union situation, where the effectiveness of deliveries and service was low, the linear system showed the areas of weakness and permitted reduction of operating expenses through increased speed in dispatching and better scheduling of deliveries of goods and services.

Figure 8
Linear Arrangement of Names By Building, Blockside, Block

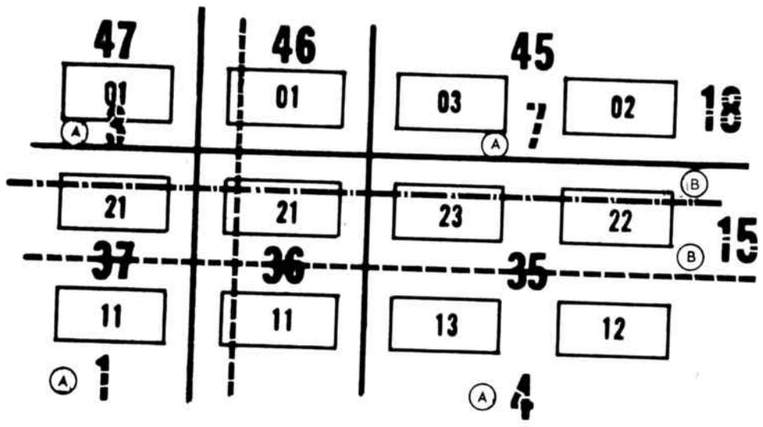
```
5735621|3800      18 E 41 ST NEW YORK
5735621|3805 AUDOGRAPH CO NY                    18 E 41 ST NEW YORK
5735621|3810 HEALTH SERVICE INC                 18 E 41 ST NEW YORK
5735621|3818 BECKER HARRY                       18 E 41 ST NEW YORK
5735621|3818 CONN GENERAL LIFE INS              18 E 41ST  NEW YORK

5735622|3000      475 5TH AVENEW YORK
5735622|3010 BARBIZON CORP                      475 5TH AVENEW YORK
5735622|3011 BENENSON REALTY CO                 475 5TH AVENEW YORK
5735622|3012 BUTTERRIESER J L ESTATE            475 5TH AVENEW YORK
5735622|3016 CESCO CONTAINER CO                 475 5TH AVENEW YORK
5735622|3031 FENESTRA INC                       475 5TH AV NEW YORK

5735622|5000      9 E 40 ST NEW YORK
5735622|5017 CAYTON INC                         9 E 40 ST  NEW YORK
5735622|5025 DODGE X SALTZMAN                   9 E 40 ST  NEW YORK

5735622|6500      292 MADISONNEW YORK
5735622|6506 AMERICAN TRANSIT ASSN              292 MADISONNEW YORK
5735622|6513 BUCHMAN X BUCHMAN                  292 MADISONNEW YORK
5735622|6542 ALFRED HOPKINS X ASSOC             292 MADISONNEW YORK
```

Figure 9
Method of Identifying Different Territorial Systems
Within a Metropolitan Linear Plan



Division Cards
Blockside Territ.

			Ⓐ	Ⓑ
35	12	1	4	15
35	22	1	7	18
35	22	2	7	15
35	23	1	7	18
35	23	2	7	15
36	11	1	4	15
36	11	2	1	15
36	11	3	4	15
36	21	1	7	18
36	21	2	3	15
36	21	3	7	15
37	11	1	1	15
37	21	1	3	18
37	21	2	3	15
45	02	1	7	18
46	01	2	3	18
46	01	3	7	18
47	01	1	3	18

Centralized Dispatching

In the case of a nationwide transportation industry, centralization of the dispatching function seemed to be the only solution to the retention of trucks and waste of time as well as facilities by decentralized dispatching offices. The linear system permitted the selection of points of origin, points of destination and reporting points in geographical sequence along the principal routes of movement and became the basis for the modification of the dispatching function.

Computation of the optimal location of manufacturing plants, distribution points, warehouses and terminals is another example where the linear system permits the accumulation of data and computer processing. Inventories and establishment of reorder points would be an application related to the optimization of the distribution network.

Impact of advertising or television commercials can be measured accurately if opinion and consumer research, sampling techniques, and comparison of expenditures to returns follow the same linear arrangement, provided by the linear master file.

Direct Mail Promotion Analysis

Direct mail promotion is another medium in dynamic marketing which is increasingly interested in being able to relate compilation data to physical locations of small centers of distribution or areas of followup activity. Comparison of data accumulated by postal zones is a cumbersome process in itself, but can be considerably simplified if linear code is used to relate geographical locations according to adjacencies. In large metropolitan and adjacent residential areas, the linear code permits identification of both irregular census tracts, as well as arbitrarily numbered postal zones, by computer methods. When linear code is used, any kind of redistricting can be accomplished in a fraction of the time otherwise required.

Advantages of Linear Coding

The benefits resulting from the application of the principle of linear coding and processing are numerous and affect many economic activities.

The dispersion and disorientation resulting from the meaningless alphabetical sequencing of geographical data is removed. With the linear code, minute identifications are not lost once data are accumulated on higher level units. Comparisons are made on higher levels, and variations can be traced to their original sources with great ease.

Tapes with pre-recorded factors can be collated with tapes recording the original data, and regional adjustments can be made through a stored program.

Cumbersome classification of data is unnecessary, as data can be extracted selectively and still remain in optimum linear relation to adjacent data. After selective processing, data can be associated in one computer run with the bulk of data through sequenced collating rather than sorting.

Reapportioning of territories, and balancing of workloads, can be accomplished in the shortest possible time, since any irregular territorial system can easily be identified against the linear system. Thus, while data are accumulated and processed on the computer by the sequenced linear system, assignment can be easily made according to other overruling nonlinear, nonaligned or nonsequential considerations.

The linear code can be subjected to arithmetic and logical operations, and relative distances of selected locations to other locations can be computed as well as accumulated.

Cumulative relative distances of each point to all other points can be arranged in order of magnitude, and the absolute, or weighted, center of data within space can be computed. Starting from the center or from any selected point, locations can be paired for optimum solution of transportation problems.

A third dimension can also be added to the two-dimensional linear code, in order to identify altitude or magnitude. This might prove useful in flight control, weather forecasting, or study of radioactivity.

The linear code, combined with the capabilities of the computers, will improve our access to information about the physical world in which we live and keep us abreast of its constant changes. Data, coded by the linear code, can be programmed to be printed in the form of statistical maps, without the need of cumbersome manual plotting.

The numerous applications of the linear code have been useful and beneficial, and seem to justify wide use of the linear geographical concept.