

NEW SAS/GRAPH FEATURES

Allan Russell, SAS Institute

Introduction

The SAS/GRAPH product is an evolving product. New features, as suggested by users or developed in-house, are constantly considered and added. This paper describes the new SAS/GRAPH features added in the past year. All of these new features are available in release 4.06 of the SAS/GRAPH product for the Digital Equipment Corporation VAX™ 11/7xx Series of minicomputers. They will also be available in the next major release of the SAS/GRAPH products for all supported machines.

There have been changes in most of the SAS/GRAPH procedures. The changes in GCHART, GPLOT, GMAP, G3GRID, GPROJECT, GSLIDE, GCONTOUR, and G3D will be presented as well as a list of new device drivers. The changes are grouped by procedure except for the ANNOTATE facility, which is introduced separately

GPROJECT

Three small changes have been made to the GPROJECT procedure, which is used for projecting maps. The first change moves the origin of the projected (output) data to the approximate center. This reduces the small distortions of the projection techniques (formerly the origin was in the lower left corner of the projected data). The second change is the output of the two parallels (ALBERS or LAMBERT projections) or poles (GNOMON projection) as a note in the log. This can be useful when combining and checking data.

GMAP

To allow you more control over the scaling of choropleth maps, the XSIZE= and YSIZE= options have been added. The XSIZE= and/or YSIZE= option can be independently specified in character cells, percent of viewing area, inches, or centimeters. Two uses of the XSIZE= and YSIZE= options are (1) to create a special effect by squeezing or stretching a map, and (2) to remove the distortion of scale normally introduced by varying resolution of graphics devices.

GPLOT

Three new options are available in the GPLOT procedure. The SKIPMISS option causes a plot to be discontinuous when a missing value is found in the data. Many types of data have inherent missing values, such as data from stores or other businesses that have no sales data from Sundays and holidays. The time axis for such data may include all days to emphasize the difference between sales (or other variables) on the days preceding and following the day with the missing value. The other two features, FRAME and CFRAME=, are related to each other. The FRAME option causes the area defined by the axis to be closed on all four sides. The CFRAME= option causes the background of the axis area to be filled with the color specified. The FRAME option can be used to divide a plot from titles and labels and give a graph a more professional look. The CFRAME= option can introduce more color to a graph and emphasize the plot.

GCHART

HBAR, VBAR, and PIE charts all have new features. Formerly, the HBAR, and VBAR reference line options, VREF= and HREF= respectively, accepted only a single number. Now a list of reference lines can be specified, following the normal syntax for specifying a numeric list. These reference lines can be used to represent critical points associated with a chart, such as projected company goals or points established by government regulation. The FRAME and CFRAME= options that are available in GPLOT are also available for HBAR and VBAR charts. The FRAME option causes the axis area to be closed on all four sides and the CFRAME= option causes the background of the axis area to be filled the specified color.

To allow you to customize your PIE charts, new options have been provided to control the placement of values on PIE charts. The MATCHCOLOR option causes the values associated with a slice to appear in the same color as the slice. This overrides the CTEXT= option which affects only the heading and group labels when MATCHCOLOR is specified. The MATCHCOLOR option helps to tie values to the proper slice, especially in pies with small slices.

Three new options control the location on the values associated with each slice. The values associated with a slice are (1) the name which is the value of the PIE variable, (2) the value which is the frequency count, percent sum, or mean depending on the TYPE= option, and (3) the percent which is the percent of the total pie the slice represents. The new options are NAME=, VALUE=, and PERCENT=, and each can specify one of four methods for positioning. The possible methods are OUTSIDE, INSIDE, ARROW, and NONE. The OUTSIDE method corresponds to the way values have previously been positioned on pie charts. This means there are 64 ways to place the values on a PIE chart. The default method for NAME= and VALUE= is OUTSIDE. The default method for PERCENT= is NONE. Changing the positioning method also effects the radius of the pie. The largest radius pie is obtained by specifying INSIDE or NONE for all methods. When any method is OUTSIDE, the radius of the pie must be reduced to allow room for the text. If the ARROW method is chosen, the radius is reduced even more. Remember that the length of the text for the names, values and percents can also effect the radius of the pie.

Three more options are new to the PIE statement. These all concern the addition of a group variable to pies. The GROUP= option accepts a variable and uses the values of that variable to produce one pie for each group value. This is very similar to the way groups work in VBAR, HBAR, and BLOCK charts. Each PIE comes out on a separate page (or screen) by default, but by using the ACROSS= and DOWN= options more than one pie can be placed on a page.

Sample 1

```
TITLE1 .C=B3 'GMAP XSIZE= and YSIZE= Options';
TITLE2 .C=B3 'XSIZE = 40 (Cells)';
TITLE3 .C=B3 'YSIZE = 3 IN';
FOOTNOTE .C=B3 'Figure 1';
PROC GMAP DATA=FL MAP=FL;
  ID STATE;
  CHORO STATE / NOLEGEND XSIZE=40 YSIZE=3 IN;
RUN;
```

Sample 2

```
TITLE1 'GMAP XSIZE= and YSIZE= Options';
TITLE2 'XSIZE = 50 PCT';
TITLE3 'YSIZE = 50 CM';
FOOTNOTE 'Figure 2';
PROC GMAP DATA=FL MAP=FL;
  ID STATE;
  CHORO STATE / NOLEGEND XSIZE=75 PCT YSIZE=50 CM;
RUN;
```

Sample 3

```
DATA TANGENT;
  IF _N_ = 1 THEN D2R = ATAN(1) / 45;
  RETAIN D2R;
  DO DEGREES = 0 TO 359;
    RADIANS = DEGREES * D2R;
    TANGENT = TAN (RADIANS);
  OUTPUT;
  END;
RUN;

SYMBOL I=J;
TITLE 'Plot with SKIPMISS & FRAME Options';
FOOTNOTE 'Figure 3';
PROC GPLOT DATA=TANGENT;
  PLOT TANGENT * DEGREES / SKIPMISS
  FRAME
  VAXIS=-30 TO 30 BY 10 VMINOR=9
  HREF=90 270
  HAXIS=0 TO 360 BY 90 HMINOR=8;
RUN;
```

Sample 4

```
PATTERN1 V=X1 C=B7;
PATTERN2 V=X2 C=B7;

TITLE 'Chart with FRAME Option';
FOOTNOTE 'Figure 4';
PROC GCHART DATA=EMPLOYEE;
  VBAR DEPT / GROUP=YEAR
  SUBGROUP=SEX
  TYPE=PCT
  G100
  FRAME;
RUN;
```

Sample 5

```
/* Figure 4 was produced by combining the
four pies from the following statements. */
PROC GCHART DATA=EMPLOYEE;
  PIE DEPT / NAME=OUTSIDE
    VALUE=NONE;
  PIE DEPT / NAME=INSIDE
    VALUE=NONE;
  PIE DEPT / NAME=ARROW
    VALUE=NONE;
  PIE DEPT / NAME=NONE
    VALUE=NONE;
RUN;
```

Sample 6

```
TITLE 'PIE Chart with Groups';
FOOTNOTE 'Figure 6';
PROC GCHART DATA=EMPLOYEE;
  PIE DEPT / GROUP=YEAR
    ACROSS=2;
RUN;
```

GMAP XSIZE= and YSIZE= Options

XSIZE = 40 (cells)

YSIZE = 3 IN

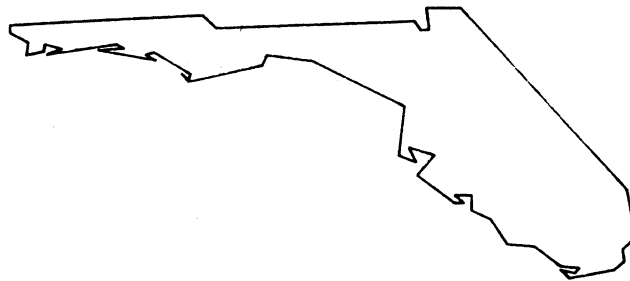


Figure 1

GMAP XSIZE= and YSIZE= Options

XSIZE = 50 PCT

YSIZE = 50 CM

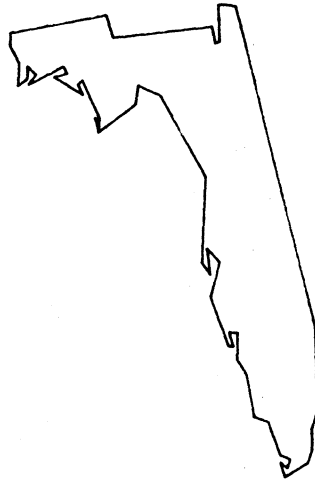


Figure 2

Plot with SKIPMISS & FRAME Options

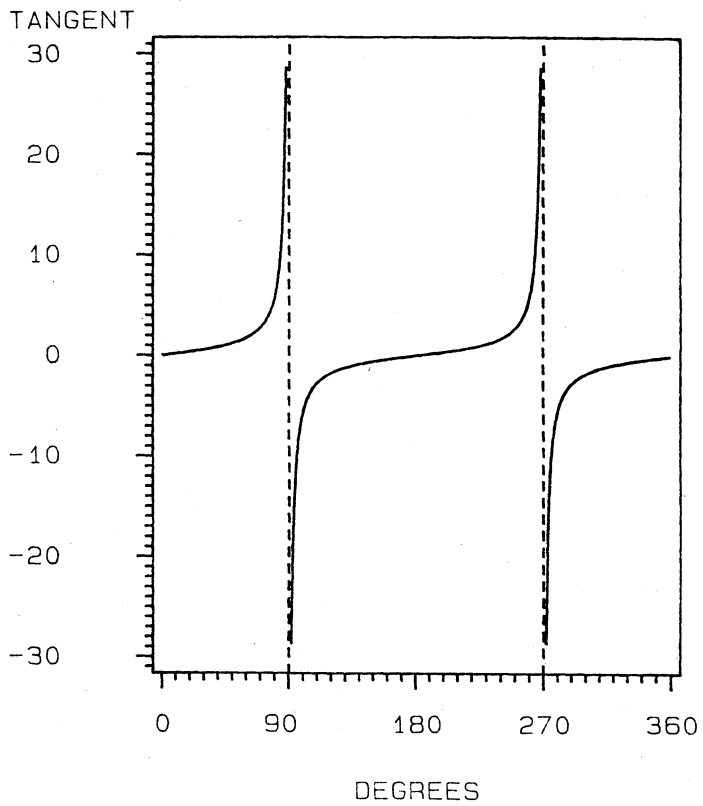


Figure 3

Chart with FRAME Option

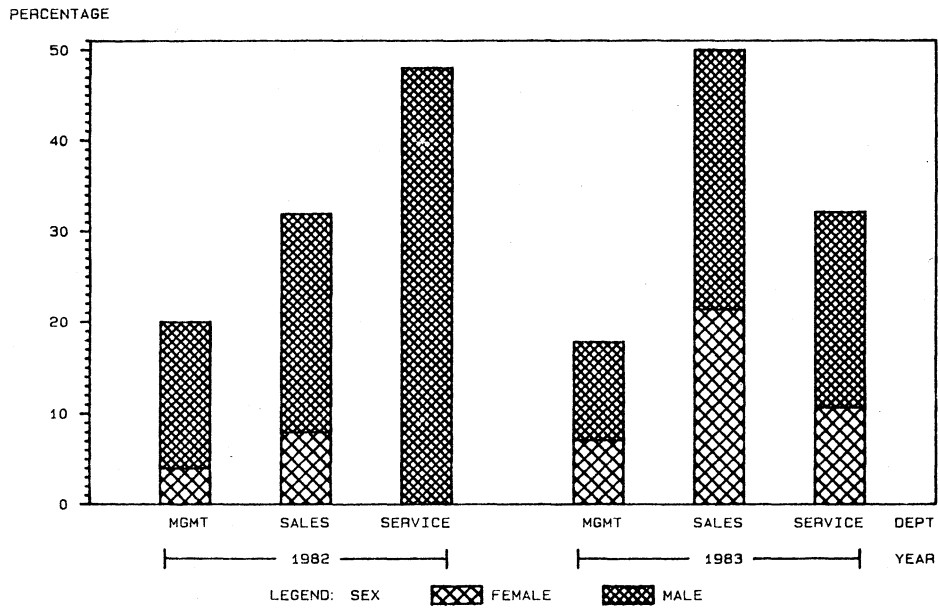


Figure 4

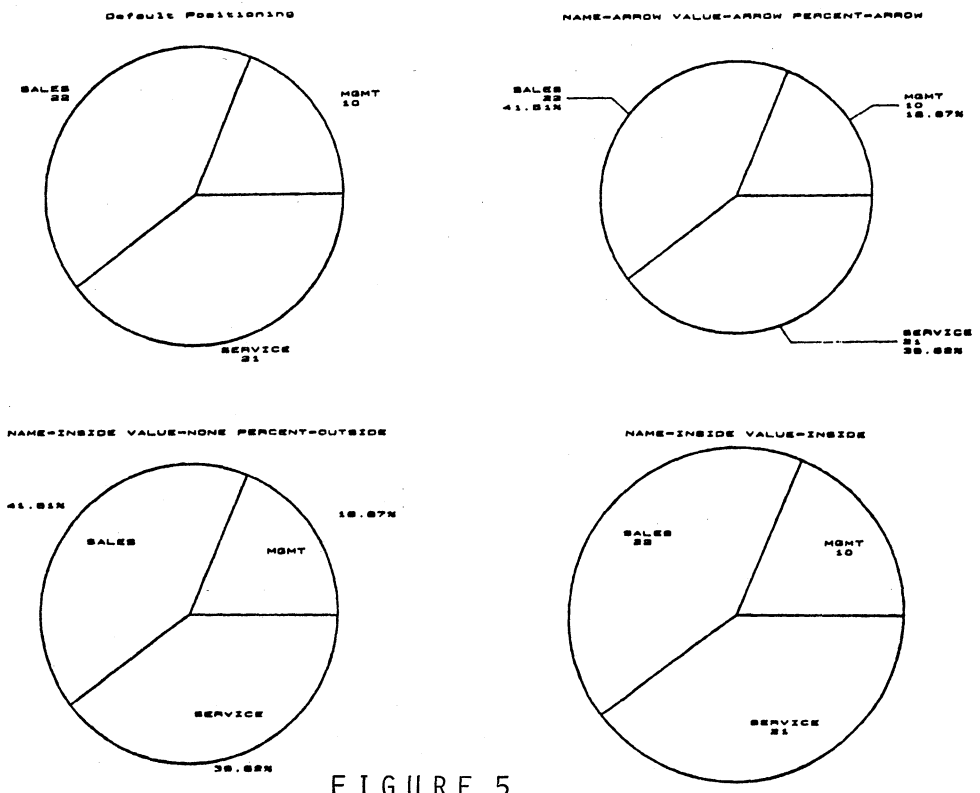


FIGURE 5

Pie Chart with Groups

FREQUENCY OF DEPT

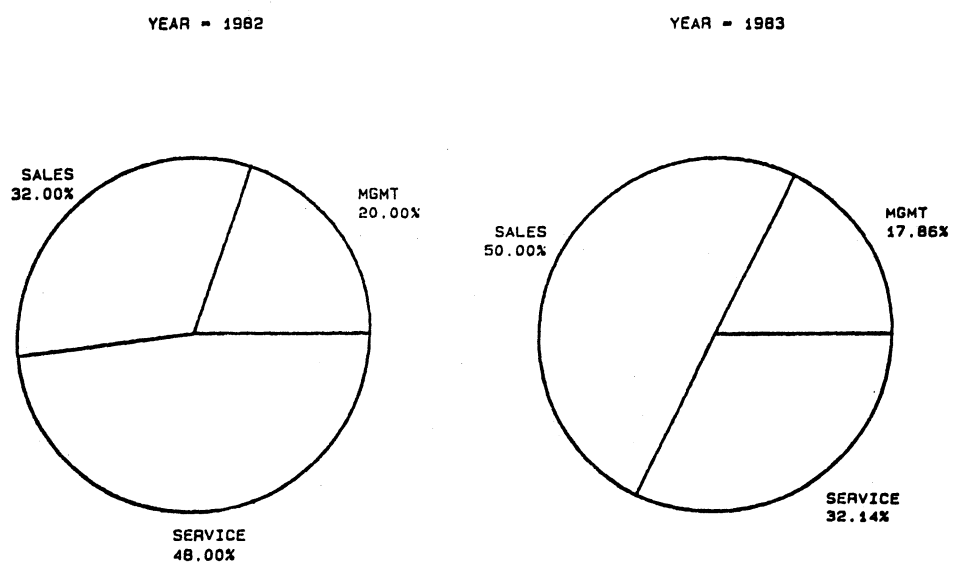


Figure 6

THE SAS/GRAPH ANNOTATE FACILITY

Introduction

The ANNOTATE facility permits you to customize SAS/GRAPH procedure output from variables and values contained in non-graphic SAS data sets. The facility is capable of defining polygons, controlling pen movement, placing text on the page, and so on. It adds to existing procedures the capability of producing custom graphic output from a DATA step.

This paper is intended only as an overview of the facility, some of its capabilities and its limitations. Complete documentation is contained in Technical Report P-128; Changes and Enhancements in the Base SAS and SAS/GRAPH Products under VMS, Version 4.

The Problem

Until now, placement of additional graphics on a given plot, whether simple lines, special symbols or a more complex application, was tedious at best. Clever DATA steps were created, or many attempts made until the proper '.M=' and '.D=' combinations were attained. Although capable of supplying a wide range of customization, one change in the data could necessitate several more attempts.

If you were clever you might achieve the desired result. Variables could be specially formatted, or text moved to obtain a specific look. But, only the simplest modifications could be made in this manner. The more complex the application, the less likely creative formatting would accomplish the effect.

The ANNOTATE Facility

The ANNOTATE facility differs from these methods in that it can be made 'automatic'. With the ANNOTATE facility, additional graphics may be placed on plots in relation to actual data values, screen cell coordinates or any mix of the two. Additional special systems permit absolute or relative coordinates and percentage of area specification.

ANNOTATE in a broad sense, may be considered a controlled access path to the graphics subroutine library. It permits you to specify commands to the graphics procedure for interpretation, execution, and eventual display on the generated plot.

The commands are contained in a SAS data set. Special variables in the data set pass information to ANNOTATE for interpretation and execution.

You may annotate your graphs by requesting the desired commands (via designated variables) in a SAS data set, that is then specified as the ANNOTATE= option to the graphics procedure.

e.g.

```
PROC GMAP
DATA    = sample data
MAP     = sample map
ANNOTATE= company.logo;
ID id;
CHORO   id / discrete nolegend
        ANNOTATE = map.detail;
RUN;
```

The ANNOTATE= option is available in both the PROC and the statement lines. The data set specified in the PROC statement will be applied globally to all interior statements in the procedure. This is useful in applications where a company logo is to be placed on each plot generated. The statement specification is active only for the statement which invoked the ANNOTATE= option, and its output is placed on the graph in addition to any ANNOTATE output generated from the PROC statement data set.

Each ANNOTATE data set in use is processed twice in generation of its output- once, before any graphic output is produced by the procedure, and again after all output has been generated. This division of processing permits you place annotations on the graph in relation to the order of SAS graph9cs generation. If labelling PROC GPLOT area fills, for example, you will want to place the text 'AFTER' the area is drawn on the graph. Otherwise, the text will be covered up when the area is filled in.

Applications

The special applications needs of SAS/GRAPH users may now be addressed through a common interface method within the SAS environment.

The immediate uses of the ANNOTATE facility center around accurate data relative positioning of text labels, particularly in mapping applications. ANNOTATE uses may include:

- Placement of city names on maps
- Drawing of lines between points
- Labelling termination of curves
- Highlighting data minimum, maximum, or inflection points
- Special presentation graphics

The facility is designed to be flexible enough that an entire graphics procedure may be generated from a DATA step, yet sufficiently simple to profit non-programmers.

Supported Functions

The ANNOTATE facility carries out its instructions through a user-specified series of commands and parameters in a SAS data set. Commands are the observations in the data set, parameters are the particular variables in the observation. The parameters (variables) have special reserved names which are indicative of their use.

A simplified description of an ANNOTATE observations is;

Variable	Description
X	The X or horizontal coordinate
Y	The Y or vertical coordinate
FUNCTION	Particular feature you want to accomplish
other variables	Additional Information

X/Y Coordinates

X and Y specify the coordinates on the graph to which the FUNCTION will be applied. They may be considered in any of a number of available reference systems. These coordinate systems translate/transform user-specified values to the appropriate screen locations.

Panel two (2) illustrates the reference systems currently available in the ANNOTATE facility. For our discussion here, let's assume that percentage ranges are 0% (minimum) to 100% (maximum), and non-percentage ranges are displayed minimums and maximums. Odd numbered systems will be considered 'percentage systems', even numbered systems will be considered 'value systems' or 'non-percentage systems'.

System numbers one (1) and two (2) are the data systems. The shading in panel two indicates the plot area referenced by these systems. Note that the 'data area' is enclosed by the axes lines. For PROC GSLIDE, which does not have distinct data areas, the data system is the same as the window system described below.

System numbers three (3) and four (4) are the screen systems. The shading in panel two indicates the plot area referenced by these systems. These systems most closely resemble the .M= and .D= values currently used by the GSLIDE procedure and TITLE, FOORNOTE, and NOTE statements.

System numbers five (5) and six (6) are the window systems. The shading in panel two indicates the plot area referenced by these systems. These systems are defined as the previously described screen systems, less the area required by TITLE and FOOTNOTE statements. If no TITLE or FOOTNOTE statements occur in a given procedure, then the window systems are identical to the screen systems.

All of the systems numbered one (1) through six (6) are absolute, that is, a given value is measured from the minimum value. Separate systems exist in which a value is measured from the last point referenced. To obtain the values for these relative systems, we add six (6) to the absolute system values, using hexadecimal notation (10='A', 11='B', 12='C') when the value exceeds nine (9). Using this numbering, the relative screen value system is 'A' (4 + 6 = 10 = 'A').

These reference systems may be used in any combination to specify screen locations in the ANNOTATE data set. The (X,Y) coordinate pair is internally transformed by the ANNOTATE facility to mate the requested location to its formal screen designation. The X and Y variables need not be referenced with the same system value. X can be referenced as 'DATA VALUE' and Y referenced as 'window percentage' in one observation, and both coordinates as 'screen value' in the next.

The ANNOTATE facility internally tracks two coordinate pairs through processing of observations to assist in this transformation process. These values are obtained from specified (X,Y) pairs and the SYSTEM specifications that accompany them in each ANNOTATE observation. These internal variables are used in supplying default values for missing values of X or Y in subsequent observations or as additional information specifiers. Some commands use these internal coordinates as start points.

These internal pairs are designated (XLAST, YLAST) and (XLSTT, YLSTT). The pair (XLAST, YLAST) refers to the last position of the "pen" at the completion of a non-text related function (like MOVE or DRAW), while (XLSTT, YLSTT) tracks the last position of the "pen" during the LABEL function. The coordinate pairs are updated automatically and may be manipulated by the utility type commands (like PUSH or POP) available in the ANNOTATE facility, but are not directly user-specified.

Below is a list of some of the available ANNOTATE commands and a short description of each.

COMMENT

Performs no actual function, but is provided as a documentation aid. Text of the comment is carried in the TEXT variable.

MOVE

Permits movement to a specific point on the graphics page without the drawing of a line. MOVE is most often used to prepare for a DRAW command, and updates the (XLAST, YLAST) coordinates.

DRAW

Used to draw a line on the graphics page. Segment drawn is defined as from (XLAST, YLAST) to the (X,Y) specified in the DRAW command. You have control over color, style and thickness of the drawn line. DRAW updates the (XLAST, YLAST) coordinates.

POINT

Special MOVE/DRAW to a specific point on the graphics page. On a CRT device, POINT illuminates a single screen pixel. POINT updates the (XLAST, YLAST) coordinates.

LABEL

Places text on the graphics page. Users may specify the color, size, font, base angle, and rotation of the characters displayed. The LABEL function updates the (XLSTT, YLSTT) coordinates.

POLY

The POLY function designates the beginning of a POLYGON DEFINITION in the ANNOTATE observation stream. Users may define the fill pattern and color, as well as the perimeter line type in this observation. The POLY function is used in conjunction with the POLYCONT function to define and fill areas on the graphics page. POLY and POLYCONT do NOT update the (XLAST, YLAST) coordinates.

PIE

Facilitates the drawing of pie slices on the graphics page. You may specify the color, fill pattern, arc angle, radius and edge lines of the slice being drawn. The PIE function maintains (XLAST,YLAST) as the centre of the PIE defined.

PIEXY

A utility command to return a location on the perimeter of the slice drawn. This is useful when positioning text around a pie (circle) on the page. The PIEXY command assumes that a pie slice has previously been defined (drawn). Unpredictable results may occur if a slice has not been drawn and the PIEXY function is invoked. The location is returned in the (XLAST,YLAST) coordinate pair.

FRAME

Allows you to specify a background color for the data area of the graphics page. This function works in addition to the CBACK= option. (See GLOBAL statements for details)

BAR

(X,Y) defines the upper-right hand corner of a bar (rectangle) whose lower-left hand corner is (XLAST,YLAST). You may specify the color and fill pattern of the bar.

OTHER VARIABLES

The remaining variables in panel one (1) serve as additional information sources for the requested functions. Some of these variables assume new meanings for different functions, though the same variable name is used. We will call these dual variables. One such dual variable is STYLE. STYLE may represent a character font if text is to be produced (LABEL) or a pattern to be used in an area fill (POLY).

PANEL 1

VARIABLE USE BY FUNCTION

M	D	P	P	P	P	P	S	L	T	C	F	B	C	P
O	R	O	O	I	I	U	O	W	A	X	N	R	A	O
V	A	L	L	E	E	S	P	A	B	T	T	A	R	M
E	W	Y	Y	.	X	H	.	P	E	2	L	M	.	M
.	.	.	C	.	Y	.	.	.	L	C	2	E	.	E
.	.	.	O	N	T	.	.	N
.	.	.	N	T	X	.	.	T
.	.	.	T	L	T	.	.	.

VARIABLE NAME

X	X	X	X	X	X	.	.	.	X	.	.	.	X	.	X
Y	X	X	X	X	X	.	.	.	X	.	.	.	X	.	X
XC	X	X	X	X	X	.	.	.	X
YC	X	X	X	X	X	.	.	.	X
XSYS	X	X	X	X	X	.	.	.	X	.	.	.	X	.	X
YSYS	X	X	X	X	X	.	.	.	X	.	.	.	X	.	X
HSYS	X	X	X	X	X	.	.	.	X	X
WHEN	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
FUNCTION	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
POSITION	X
SIZE	.	X	.	.	X	X	.	.	X	.	.	.	X	.	.
ANGLE	X	X	.	.	X
ROTATE	X	X	.	.	X
STYLE	.	.	X	.	X	.	.	.	X	.	.	.	X	X	.
COLOR	.	X	X	X	X	.	.	.	X	.	.	.	X	X	X
LINE	.	X	X	.	X	X	X	.
TEXT	X	X

PANEL 2

EXPECTED VALUES

ABSOLUTE	'1'	DATA	%	0% TO 100% of AXIS AREA
	'2'	DATA		AXISMIN to AXISMAX
	'3'	SCREEN	%	0% TO 100% of PAGE WINDOW
	'4'	SCREEN CELL		00 to EDGE of PAGE WINDOW
	'5'	WINDOW	%	00 TO 100% of PLOT WINDOW
	'6'	WINDOW CELL		00 TO EDGE of PLOT WINDOW
RELATIVE	'7'	DATA	%	0% TO 100% of AXIS AREA
	'8'	DATA		AXISMIN to AXISMAX
	'9'	SCREEN	%	0% TO 100% of PAGE WINDOW
	'A'	SCREEN CELL		00 TO EDTE OF PAGE WINDOW
	'B'	WINDOW	%	00 TO 100% of PLOT WINDOW
	'C'	WINDOW CELL		00 TO EDGE of PLOT WINDOW