CGTM # 58
Jan. 1969

PEG--On-Line Date-Fitting Program, A Summary

by

Lyle B. Smith
Stanford Linear Accelerator Center
Stanford, California

PEG is a program used for the on-line interactive solution of data-fitting problems. The method of least squares is used to accomplish the fitting. The man-machine interaction is designed to make PEG useful even to persons having no programming experience. This is accomplished by using "interaction by anticipation" which means that all possible desires of a user are anticipated and choices are presented on a graphical display device (CRT) allowing user selection of options rather than user specification of actions. Figure 1 shows a view of the console employed by the PEG system.

PEG has been written almost entirely in FORTRAN IV with only a few assembly language routines used for handling the graphical display device. This choice of language has simplified the implementation and gives the program some degree of machine independence due to the widespread use of FORTRAN IV.

An IBM 360/75 and later an IBM 360/91 computer with an IBM 2250-II display unit (CRT with lightpen and keyboard) have been used for the implementation of the PEG on-line fitting system. The program is currently running on the 360/91 under the OS/MVT operating system. The operating system allows 150K (K=1024) bytes (byte=8 bits) of core storage to be used by programs using the 2250 display. Such programs are given a high priority in a multiprogramming environment, thus appearing to have sole use of the central processor.

Because of the limited amount of core storage, the PEG system involves 63 overlay segments in order to provide the desired range of capabilities.

Figure 2 shows the overall flow of control in PEG as well as summarizing the capabilities of the program. Control is passed between the various displays corresponding to the boxes of the flowchart in Figure 2 by means of lightpen commands. In what follows we describe briefly the functions of the displays. In Smith [1969a] a more detailed description of the use of PEG and criteria for the design of such interactive systems are set forth.

Introduction

The introductory display shows remarks which briefly describe the program and give some general information.

Choice of Function

This display allows selection of a fitting function. The currently available choices are:

a) Orthogonal polynomials

b) Spline functions

c) Fourier approximations

d) User defined functions

Choice of Data Mode

The methods of entering data into the system are:

a) Data cards

b) Keyboard

c) Residuals of previous fit

d) Data of previous fit

PEG - (ON-LINE DATA-FITTING)

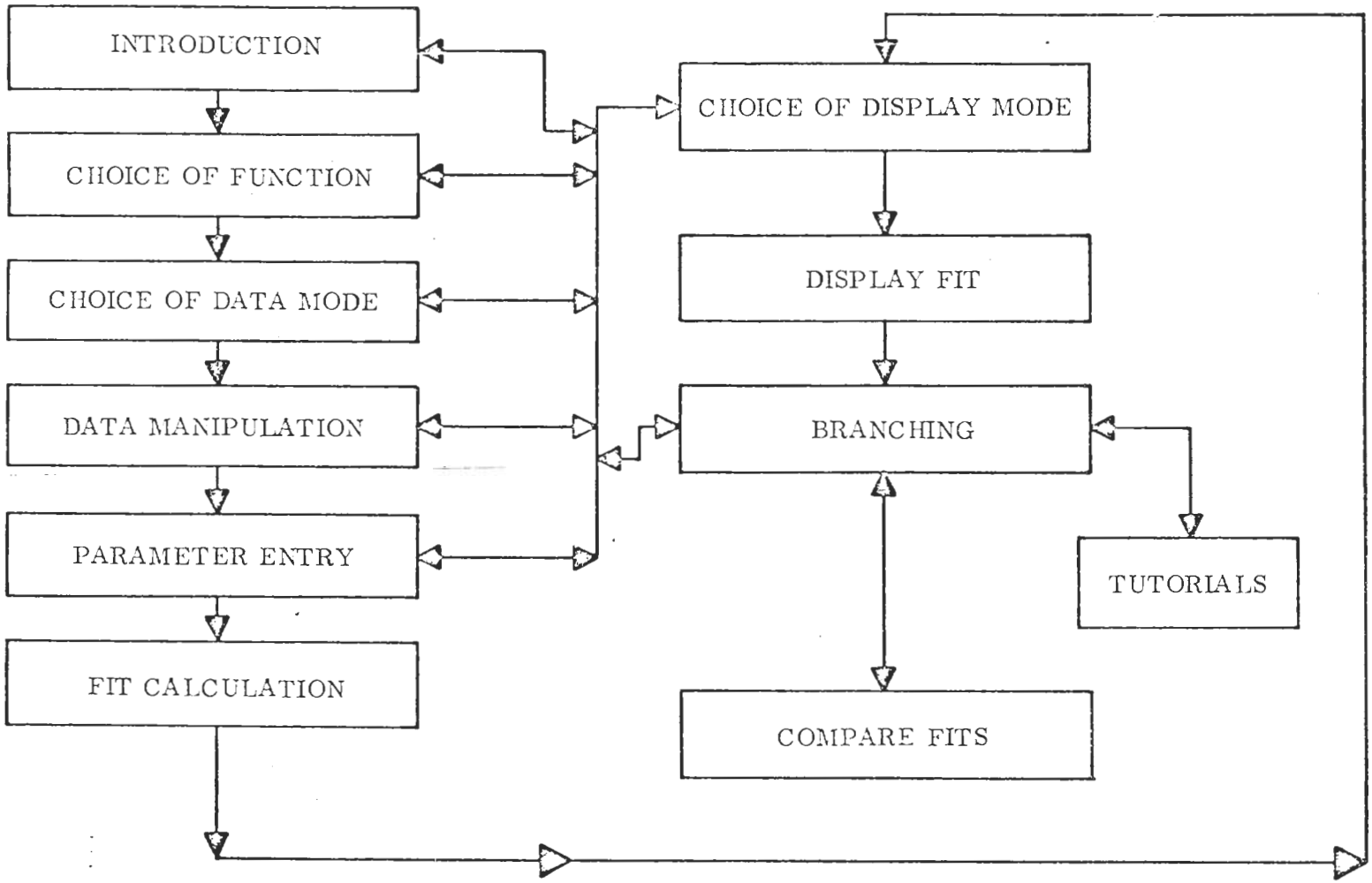

Figure 2

Several sets of data may be included in the deck of cards submitted to the computer when a run is to be made. These sets are then read by PEG. One set is read each time "data cards" is selected on this display. Data can also be entered on-line through the 2250 keyboard. The residuals of a computed fit are saved for possible use as data and the previously used data can be recalled for refitting by a different function.

Data Manipulation

There are several displays involved with manipulating the data. The capabilities included here are:

a) Correction of data points

b) Selection of a subset of the data

c) Transformation of the data (either or both coordinates can be transformed in various ways)

d) Point deletion

Parameter Entry

The display allowing parameter entry varies with the fitting function.

a) Orthogonal Polynomials - the degree is entered

b) Fourier Approximation - the degree is entered (in this case degree n means n cosine terms, n sine terms and a constant term)

c) Spline Functions - the degree and the number of joints are entered. This is followed by entry of initial values for the joint positions and a decision as to whether the joints are to be held fixed or allowed to vary.

d) User Defined Functions - the number of parameters and their initial values are entered and a decision is made as to whether or not the parameters are to be varied automatically in finding the least squares fit. If not, PEG allows the user to examine the fit for a given set of values and change them manually (through the keyboard).

If automatic variation is chosen the program varies the parameters using a minimization routine to minimize the sum of squares of residuals.

Fit Calculation

The least squares fit is calculated by various methods.

a) Orthogonal Polynomials - use Forsythe orthogonal polynomials (see Forsythe [1957]).

b) Fourier approximations - use Goertzel's algorithm as described by Ralston [1965] for the case of equally spaced points. For unequally spaced points or if weights are associated with the data a straightforward solution of the normal equations is employed.

c) Spline Functions - use a straightforward solution of the normal equations derived from an elementary spline representation of the spline function. This is a linear problem in the case of fixed joints. If the joints are allowed to vary the problem is non-linear and a direct search algorithm (see Hooke and Jeeves [1961]) is used to minimize the sum of squares of the residuals by varying the joint positions.

c) User Defined Functions - use the direct search algorithm or an interactive minimizer to minimize the sum of squares of the residuals, as the parameters of the given function are varied. User defined functions are included by writing a Fortran function which evaluates the desired function and conforms to a specified name and calling sequence. This option to define one's own fitting function gives flexibility to the PEG system. The interactive minimizer which can be used instead of or after using the direct search fit is an on-line Gauss-Seidel type minimization routine which displays graphs allowing a user to alter parameter values manually (by lightpen command) as a minimum is sought.

In Smith [1969a] the numerical methods used in the fitting are discussed in more detail. Weaknesses of the methods are pointed out and alternative methods are suggested. The modularity of the PEG system makes it easy to add or substitute different algorithms to perform the desired calculations.

Choice of Display Mode

After the fit has been calculated this display allows selection of one of seven different displays showing graphically and numerically the results of the fit. These displays graph the data with the fitting function superimposed, show a table of residuals, graph the residuals and show the numerical values of the parameters of the fitting function. Either half scope or full scope graphs can be chosen as well as a graph showing an extrapolation of the fitting function.

Display Fit

At this point the previously selected display mode is shown on the CRT. The seven different displays available here are:

a) Graph data and fit and show the residuals numerically.

b) Graph data and fit and show the coefficients numerically.

c) Graph data and fit on full scope.

d) Graph the residuals and show the residuals numerically.

e) Graph the residuals and show the coefficients numerically.

f) Graph the residuals on full scope.

g) Graph the data and fit extrapolated.

Branching Display

As shown by the flowchart in Figure 2 this display can be reached from most other displays in the PEG system. This allows a user to transfer control back to a previous step in the fitting process (for example to choose a new

degree or a different fitting function) as well as to escape from the various displays by transferring control to the branching display. Since the tutorial displays can be reached from this branching display, the latter capability provides a user with on-line assistance during a session at the console. From almost anywhere in the fitting process he can "branch out" to read tutorial information and then return to where he was.

Compare Fits

This display is reached from the branching display which has an option to save a fit for comparison purposes. The comparison display shows the degree, the maximum residual, and the sum of squares of the residuals for each of the built-in functions, orthogonal polynomials, spline functions, and Fourier approximations. Thus if the same data were fit by these three different functions a comparison of the "goodness of fit" could be made by this display.

Tutorials

At this point a list of several tutorial displays is displayed for user selection by lightpen. After a tutorial display has been selected, displayed and read, control is transferred back to the display showing the list. At this point a different tutorial can be selected or control transferred to the branching display. The currently available tutorial displays are explanations of the following:

a) Orthogonal Polynomials

b) Fourier Approximation

c) Spline Functions

d) User defined functions

e) How to enter numbers from the keyboard

A more detailed description of PEG is given in Smith [1969a]. A survey of interactive graphical systems and their relation to PEG is given in Smith [1969b]. This survey discusses existing and proposed systems and compares their language levels to that of PEG. A comparison of the solution of a non-linear least squares problem among the various systems is also given.

REFERENCES

Forsythe, G.E. [1947]. Generation and use of orthogonal polynomials for data-fitting with a digital computer. SIAM Journal, Vol. 5, No. 2, (June), 74-88.

Hooke, R. and Jeeves, T.A. [1961]. "Direct search" solution of numerical and statistical problems. Journal of the ACM, Vol. 8, No. 2, 212-229.

Ralston, A. [1965]. A First Course in Numerical Analysis, McGraw-Hill Book Co., New York.

Smith. L.B. [1969a]. The use of man-machine interaction in data-fitting problems. SLAC Report No. 96, Stanford Linear Accelerator Center, Stanford, California.

Smith. L.B. [1969b]. A survey of interactive graphical systems for mathematics, SLAC-PUB-540, Stanford Linear Accelerator Center, Stanford, California