

DECLASSIFIED
 E.O. 13526
 Authority
 By L.S. HAVA, Date 12-15

CONFIDENTIAL

Declassified by D. Janosik
 Deputy Associate Director for Policy and Records
 on 13 Oct 2010 and by MTB
 MPRO-104 HARVEST BULLETIN NO. 2

| | Act | Info |
|--------------|-------------------|------|
| 18 July 2960 | <i>[initials]</i> | |
| 3 | | |
| 31 | | |

Return to MPRO-302

SUBJECT: SUPERSTETH GENERAL LITERAL DIAGNOSTIC PROGRAM

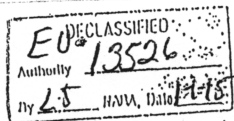
The SUPERSTETH General Literal Diagnostic program was one of the first programs selected to be written for HARVEST by MPRO-104. In October 1959 one programmer was assigned to the project, and in December 1959 an additional programmer was assigned. The project is well under way with most of the various phases of SUPERSTETH written and debugged. It is expected that SUPERSTETH will be available as soon as HARVEST becomes operational.

The source material for the project is "SUPERSTETH Diagnostic Programs" PROD-03 Informal no. 18, dated 25 August 1958. The numbering used later in this report is the same as that used in the source material. Conferences have been held with a PROD-03 representative to answer programmers' questions.

Each of the SUPERSTETH routines was analyzed as to type - computational, frequency counting or matrix formatting - and as to information needed to perform that routine. A new rebus has been designed to permit sponsor selection of routines and where necessary, parameters and thresholds. A supervisory program has been planned to control operational runs of selected routines and to control variable parameters and thresholds.

At the outset, several assumptions had to be made. Since the I/O package had not been confirmed, it was assumed that the data stream was in memory and the statistics computed would be saved. The program was designed to handle 50,000 character streams and any alphabet size up to 64. Streaming instructions were written to take the various frequency counts. Calculations for streaming instruction setups have been written and debugged. Since the streaming instructions themselves could not be debugged for some time, it had to be assumed that the necessary count had been taken and that the individual frequencies were already in memory.

CONFIDENTIAL



~~CONFIDENTIAL~~

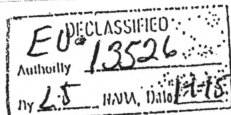
For speed of operation, normalized single precision floating point mode was used in computing the various statistics. Routines to float the counts have been written and debugged. The computed statistic was converted to a decimal unsigned number rounded to four decimal places. For ease of programming, some routines were combined into one routine. There were five such sets; the routines combined are indicated later in this report.

Work is still in progress to program the polygraphic repeat and derived stream phases. For the polygraphic repeat phase, it has been planned to utilize two generalized file operators — offset and sort — being written by other programmers in MPRO-104. For the derived stream phase, streaming instructions are being written.

The following phases of SUPERSTETH, including their supervisory control routine, have been written and debugged:

1. Monographic Tests
 - 1.4 Monographs by Position-in Group
 2. Digraphic Tests
 - 2.5 Formation Digraphs within Group
 - 2.6 Formation of digraphs as selected by sponsor
 - 2.6.0.1 Formation of combined digraphic distributions
 3. Monographic Local Roughness
 5. Width Writeouts
- 18 square roots based on alphabet size.

~~CONFIDENTIAL~~



~~CONFIDENTIAL~~

The following are the computational routines that have been written and debugged:

MONOGRAPHIC

- 1.2.1 Delta I C and sigmage
- &
- 1.2.2 Chi-squared and sigmage

MONOGRAPHIC BY POSITION-IN-GROUP

- 1.4.2 Delta I C, sigmage; Chi-squared, sigmage for each column in group
- &
- 1.4.3 Total Chi-squared and sigmage for all columns in group
- 1.4.4 Chi-squared on column vs. column and sigmage
- 1.4.5 Chi-squared on column vs. column at slides and sigmage

DIGRAPHIC

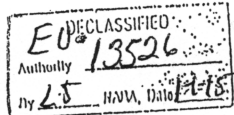
- 2.1.2 Delta I C and sigmage
- &
- 2.1.3 Chi-squared and sigmage
- 2.2.1 Chi-squared for rows sums vs. a flat distribution and sigmage
- &
- 2.2.2 Chi-squared for column sums vs. a flat distribution and sigmage
- 2.2.3 Chi-squared for row vs. column sums and sigmage
- 2.2.4 Chi-squared for rows vs. column sums at slided and sigmage
- 2.3.1 Chi-squared for rutting and sigmage
- 2.3.2 Chi-squared for symmetry and sigmage
- 2.3.3 Chi-squared for bias and sigmage
- 2.3.4 Chi-squared for comparing distribution along a diagonal against that expected from row and column sums and sigmage

WIDTH WRITEOUTS

- 5.1 Average columnar delta I C on widths and sigmage
- &
- 5.2 Monographic delta I C and sigmage for each column of width

3.

~~CONFIDENTIAL~~



~~CONFIDENTIAL~~

Width Writeouts (Continued)

- 5.3 Chi-squared on column from width writeouts and signage
- 5.4 Chi-squared on column vs. column at slides from width writeout; signage

The following routines have also been written and debugged:

- 1.1.1 Monographic distribution in frequency order
- 2.4.1 Poisson selection of the digraphic distribution
- 2.4.2 Selection of those digraphic that occur "m" or more times

The following streaming instructions have been written:

- 1. Monographic Frequency Count
- 1.4 Monographic by position-in-group frequency count
- 2. Digraphic Frequency Count
- 3.1 Coincidence count on stream offset against itself
- 3.2 Subtotals of monographic coincidence count according to position-in-group
- 5. Counts on each column in width
- 8.1.1 Form the delta-n stream
- 8.1.2 Form the delta-n stream with subtraction reversed
- 8.3.1 Form de-fibonacci stream

As the I/O package has still not been confirmed all input and output instructions, including formatting of headings, counts or statistics, remain to be written.

When debug facilities for streaming become available the streaming instructions will be checked out.

Any questions concerning SUPERSTETH on HARVEST call Wallace Rossmiller or Lillian Hall on extension 4624.

~~CONFIDENTIAL~~