"Problems, and A Program"

(EXPERIENCE WITH THE DATATRON AT GENERAL PETROLEUM CORPORATION)

ARCH B. JOHNSTON

Mr. Johnston, currently Manager - Data Processing Research, has been with General Petroleum Corporation in Los Angeles since 1937, except for service during World War II.

After his return from service he fulfilled assignments in the Market Operations Department, Comptrollers Department, where he became Assistant Manager of Systems in 1954. Shortly thereafter he began an intensive study in the field of electronic computers in business.

Mr. Johnston is responsible for development and installation of all business programs on Datatron, and continuing research toward new electronic equipment and better use of present system.

Mr. Chairman, members of the Third Electronic Business Systems Conference. It is a pleasure for me to have an opportunity to meet with a group of people who have the same interests as I, and this is especially true when we can meet in San Diego. I suspect, if all Angelenos were very truthful, they would agree that they have asked themselves more than once what could be done to get their job set up in the San Diego Area.

I am here this morning to bring you an account of how (and a little about what) we are doing at General Petroleum with our computer. The time is necessarily limited, so that I have done a lot of soul searching in recent weeks trying to decide exactly what phase of our work would be most interesting and, I hope, helpful to you. Only a small part of the activity that we have had under way for about a year could be discussed at best.

And so I have made a decision with which I hope you will agree. In brief, I am going to by-pass the conventional tales of how we got into computer, I am going to ignore the interesting but not very significant facts about how we selected personnel, arranged the lay-out of the Computer Room, and assorted subjects. I am even going to avoid the always ripe subject of why we, at General Petroleum, selected the DATATRON instead of one of the other very fine medium-scale data processing systems. (There *are* several others, you know!)

I am going to "branch around" these aspects for two reasons. First, they have been adequately covered in sundry articles and papers (and copies of these materials are being made available to you here today to take away and peruse at your convenience) and second, I think that the strongest contribution I could make before you today would be to give you a current and factual account of the things most speakers before electronic groups have studiously avoided in the past. Specifically, I will talk to you about the PROBLEMS in electronic data processing, about the things neither equipment manufacturers nor other users of electronic machines seem able or willing to discuss openly. In doing this, I hope that I will not discourage those of you who are under way in some sort of electronic effort, nor frighten off those of you who may be in the early investigative stages of a computer study.

"Problems"

In talking about these problems, I want to make clear two things. One is my firm belief that, despite the "problems", no business enterprise the size of my Company (and this perhaps should include smaller companies) can exist competitively in the future *unless* they interweave an electronic system of some sort with the daily decision-making routine of their operations. We have already seen this in the oil industry, and the result is that there, today, exists almost no medium or large oil company that is not in computers to some degree. Many of you in the aircraft industry know the truth of these words, far better than we, for history will show that the first and strongest computer users were aircraft people.

Now, about "problems." These can generally be divided into three categories:

1. The Equipment
2. The Programming
3. The Input.

It may seem that I've overlooked "output" but, in our case at least, we've found that overcoming the other three types of problems results in pretty fine "output" —no problems there that can't be solved by simple control measures.

About Equipment. These are problems almost one hundred per cent related to "Maintenance." Theoretically, every element of every computer system is pretty fool-proof (or so the designers thought) and, if the

hardware performed as advertised there would be no "equipment" problems. Unfortunately, this is not the case and this is not alone the situation with the Datatron. My association with users of other gear reveals that our problems are common problems.

In our system, we have four major elements of hardware. I mention these in order of their historic age intentionally:

1. The Input-Output Boxes
   (I.B.M. 089, 523, 407, Flexowriter)
2. The Central Computer
   (Datatron Model 205 with 4000-word Magnetic Drum)
3. The Magnetic Tape System
   (Three Data Readers and Control Unit)
4. The Buffering-Editing System
   (Cardatron with one Input and two Output Stations).

In discussing system reliability, I would like to treat each of these elements separately.

The *Auxiliary Machines* for card reading, card punching, line printing, and character printing are extremely reliable. They are not 100%, but they will, in aggregate, measure greater than 99% "up"—and that's good!

Regular, but infrequent, attention keeps the electro-mechanical gear humming. Resident I.B.M. service engineers, normally busy in our Tabulating Department, also wet-nurse the machines hung onto the computer, and they do it willingly and well.

The *Datatron Computer,* next oldest in point of time, is almost as solid. Reliability is greater than 99.9% and "up-time" of the main frame itself has averaged better than 95%—a tribute, I think, to good design initially, and to built-in marginal checking circuitry which permits effective daily preventive maintenance. P.M. time on main frame alone need not exceed one hour a day—with an occasional six or eight hour stint (perhaps every two weeks) to do a thorough check-out. We have 1,527 hours on our computer now, and have had no major problems. Our drum is intact. Once or twice, ElectroData has checked head alignment, but no trouble has been encountered. Our spare parts costs are within the range originally estimated and virtually all tubes, diodes, etc., that require frequent replacement are readily available at our local radio parts house. No main frame "down-time" has exceeded five hours and normally will be less than one hour.

The *Magnetic Tape System,* although installed with our computer, is a newer development and, while a number of "scientific" users had had tape for as long as a year or more before us, we were one of the first users to attempt business data processing involving heavy volume, sustained operation of tape. Our *Average* experience with the Tape System to date is that it has been about 98% reliable, but only about 80% "up." Fortunately, the Maintenance Engineers are now getting the measure of things and, while we continue to have the same small family of troubles, they are generally quickly recognized and corrected through component replacements or (more often) readjustment of Schmidt triggers or gain levels. Down-time on tape system now seldom exceeds three or four hours, in extreme cases, and quite often is more like 30 minutes. "Up" time on tape is getting better, but it is still "wanting" in both reliability and accuracy.

Half the battle seems to be getting good tape in the first place, properly calibrating it, and extensively testing it. The greatest weakness of the Datatron System is the cost of preparing tape for use. We find that, where the system problem demands absolutely perfect tape, 10,000 blocks each channel, we must spend at least two hours, on-line, to adequately test the tape by calibrating, writing test patterns, reading back, etc. When a bad block is encountered, additional time may be spent to recalibrate. This requirement for large numbers of perfect tapes seldom exists in installations doing engineering or scientific work—so many of our problems are not known to some Datatron Users.

But, it is almost always true that business data processing, particularly where there is a philosophy for integrated systems, will require lots of tape, and will require it to be perfect. We, at General Petroleum, have over 100 reels in our tape library now—and it is slowly growing as we develop new computer applications.

It is my understanding that ElectroData is now considering both design improvements to the tape system and a plan to supply Datatron Users with precalibrated, tested, and guaranteed tape at a reasonable price. If they are able to do these things, they will have contributed a needed service and improved their competitive position. The economics of "do-it-yourself" tape testing (not too bad in our case where we own the equipment) are pretty grim in a rental installation, and an unreadable block, however infrequent, causes significant lost time to redevelop.

The *Cardatron System,* installed less than four months ago, is a real pleasure. When it works, it's unquestionably the greatest thing on the intermediate market. It does all that was promised and increases the efficiency of the system dramatically. And it works very well for us.

Our experience since July has been that the Cardatron is about 99% reliable and about 90% "Up." "Up" time would be better yet, except that most malfunctions are new ones and the maintenance crew doesn't yet have that intuitive sense of what remedial action should be taken. This nets out to an extended "hardware debugging" session that is as much "training as "unscheduled maintenance."

This brief summary will tell you a little about our equipment problems. We are sometimes discouraged, short range—but we regain our perspective when we

hear that others have no bed of roses either, or when a ten-hour program runs without failure and gives a fast and successful result. Long range, we feel secure—the Datatron 205 is getting better as engineering improvements come along and, most important, ten to fifteen G. P. employees are learning a Helluva lot about the operating problems in Data Processing—the "hard way."

About Programming. Since the programming and problem analysis phases of Data Processing are in my personal portfolio, I am *"delighted"* with our results to date. (They may not be many, but—we think—they are good.) Seriously, aside from the fact that it takes us twice as long to finish a job as anyone (not in EDP) thinks it could, our programs are shaping up well. A few of our random (and controversial) conclusions and techniques may interest you:

1. The Datatron command repertoire is fully adequate for all work we have tried to do.
2. Program timing can be estimated very closely before actual running — using published times from Datatron Manuals.
3. Actual programming assistance from manufacturer's sales technical representatives is not worthwhile beyond point of initial training, as they can't be made adequately familiar with the problem as rapidly as trained Company Systems people can be schooled as programmers.
4. Best potential business problem Analyst-Programmers come from ranks of Systems and Tabulating Analysts.
5. Thirty days' training will make these men *think* they are "first class' programmers, but
6. They aren't—until they have logged at least six months' heavy experience.
7. Auto-monitors and tracing routines are unnecessary for the debugging of business problems. Short service routines for Drum Dump, Tape Dump, and Print Out Selective Locations are adequate.
8. Problems involving heavy tape work require programmed checks against:
   a. Skipped block(s) on multiple block reads or writes.
   b. Incorrect data transfer (individual block sum checks where accuracy of results must be 100% —with check sum partially circulated).
   c. Shifted information in Word 20, where accuracy not much less than 100% will do.
9. Other tape problem program "musts" are:
   a. Tallies for all Tape Search Commands.
   b. "Restart" procedures for all types of tape failures.
10. Input data sum checks (or equivalent safeguards) should be included to control punched card input, since neither I.B.M. 089 nor Cardatron are infallible.
    a. A substitute input control, sometimes useful, is to immediately *output* (where punch is available) card—then, at end of run, compare cards off-line.
11. Output control is desirable—and sometimes mandatory—for same reasons.
    a. One technique is to write each "card image" to a work tape as soon as card write order has been executed (usually from Loop 4)—then, at end of run, input all output cards, comparing with tape and punching second output card in case comparison fails (theory being that lightning will not strike twice in same place!).
12. All programs should be as "self-loading" as possible—set-up time is too long (and errors too many) when human intervention is heavily involved.
13. Relative address coding is essential—Program changes and error corrections can't be *efficiently* made when program is originally coded in real address.
14. Debugging time may be bountiful and somewhat uncontrolled at first, but eventually must be rigidly disciplined and limited to short sessions.
15. Program documentation is extremely important, if programmers are ever to be divorced from past completed work.

About Input. Input data (for most business problems) is usually punched cards and the need for controls is too obvious, too well known, and too widely practiced to need much mention to machine accountants. But, integrated systems for computers usually begin long before the punched card—with the basic documents of the business, be they Sales Tags (about which I'll talk later), or Shop Orders, or Time Cards.

And it behooves the Analyst-Programmer to apply all the techniques of work simplification and forms design to these early phases. If data is *easy* to record at point of origin, it will be less prone to distortion enroute to the Computer. A few hints from our experience include:

1. Good, sensible, workable number systems (codes) designed *into* basic forms to be "self-coding" wherever possible.
2. Minimization of codes (allowing computer to expand where necessary) so that people originating data have as little to do as possible.
3. Adequate training of clerical and keypunch personnel *before* the demand is on them to produce volumes of data—for a new computer program, or a conversion from a punched card system which used a different intermediate format.
4. Built-in, programmatic checks against the potential kinds of input errors like:
   a. Invalid codes
   b. Inadvertent blank columns
   c. Double punched cards.

"Purification" of input can't be over-emphasized. Next to actual unscheduled maintenance, our largest contributor to "lost time" is incorrect input data. This is a far more serious danger than all computer system errors combined.

"And a Program"

Now, in the field of applications, I would like to briefly outline one program—a Petroleum Marketing Statistical Program—that is really several programs.

These programs do not concern themselves with the dollars and cents that go to make up the Profit and Loss Statement, but rather with the statistics of sales that enable our Management to determine the effectiveness of the Marketing effort. To be helpful, this information is needed yesterday; that is to say, if Sales Management is to make effective decisions, the information on which these decisions is based must be current. Because of this and the large volume of data to be processed, we thought this problem would be a good computer application. Our experience has thus far proved this to be the case.

In order to understand our objective, let's examine the statement we must produce. Each line is numbered and this corresponds to a product or group of products. Total lines for like products are indicated. The vertical columns are:

This Month Sales
This Month Last Year Sales
Percentage Gain or Loss
This Year to Date
Year to Date Last Year
Percentage Gain or Loss

and then the troublesome but all-important column "Sales Objective."

We must produce approximately 3,000 of these statements each month for varying levels of management. They record sales by classes of trade, within these levels. (Branch—about 300, District—about 30, Division—3, and Company Total.)

In preparation for this application, Sales History by Branch, Class of Trade, and Product Line Number, by months for the year of 1956, was supplied by Division and Comptroller's Tabulating Units. This information, recorded in an aggregate of over three hundred thousand I.B.M. Cards, was fed to the Datatron under control of a "one time" Program (603). This "one-time use" program developed District and Division level sales and stored data, for all "lines" of all levels of statements, on magnetic tape. Resultant tapes (one for Central, two each for Northwest and Southwest Divisions) represented the electronic counterpart of the Card History Files required under the previous I.B.M. punched card method.

The magnetic tape history files were later "up-dated", one month at a time, so that the history recorded, as we started actual operations with July 1957 business, was for the period July 1956-June 1957, inclusive.

Actual processing for one month follows the pattern outlined below:

1. Division Accounting Offices (Comptroller's Tabulating in the Southwest Division) forward Plant Sales Summary Cards and Distributor Sales Cards to Electronic Computer Unit, Los Angeles. Volume per Division 12,000 to 18,000 cards.
2. Cards are processed "off-line" to convert product code to line number. They are then further summarized by line number and class of trade and reproduced—six "images" per card. Resultant cards (approximately 2,000) are recoded to Computer Branch Codes and sorted to that sequence. They are then ready for input to the Datatron. Approximately six man hours per Division are required on auxiliary machines to prepare input.
3. Computer Program (601) now inputs the "current month" cards through the 089 Collator. This program reads all the cards for a Branch, accumulates on the magnetic drum each of the six records per card to "line" totals by the eleven different primary class of trade groupings required at District level. Total lines are then developed. These groupings are next accumulated into a District assembly but are retained in original form on the Branch assembly area of the drum. They are then consolidated into the eight statements required at Branch level and these "current month" columns for each Branch statement are written on magnetic tape. When the last Branch in a District is written to tape, the program develops up to 18 statements for the District, transfers values to the Division assembly area, and writes District statements on a second tape. When the last District for a Division is written, the program develops Division total statements and writes these on the second magnetic tape.
   Time required for this program is approximately one minute per Branch, per District, and for Division totals, or approximately two hours to completely develop "current month" data for each Division. Magnetic tapes resulting from Program 601 are the Input to Program 602.
4. Computer Program (602) next inputs the tape from 601 and the "history" tape (which records sales for each of the preceding twelve months, and year to date sales this year and last year for approximately 1,000 different "location-class of trade" entities). This program develops each line of each of approximately 1,000 statements, one line at a time, calculating percentage increase or decrease for "month" and "year to date." As fast as a line is developed, a card is punched, with line number and all information needed for the final report.
   Sales Objectives are received from Division Offices and "posted" to the History Tapes quarterly (monthly in case of Northwest) by a fast program just prior to running of Sales Analysis.

Program 602 automatically determines whether or not "objectives" are present and, if so, calculates percentage attained. The objective and percentage are then incorporated in the line output card.

Another automatic function of Program 602 is to develop and store on the History Tape the appropriate data required for the Quarterly Comparative Sales Reports prepared electronically for Home Office Lubricants and Process Products Department. (This report is run after completion of Sales Analysis for the quarterly month under control of Program 604.)

As this program processes individual lines, it simultaneously "up-dates" the History Tape, substituting "current month this year" sales for "current month last year" and changing "year to date" sales.

Time required for this program is approximately six hours per Division.

5. Output cards from Program 602 are next merged with Statement Heading Cards and final reports are run on the Datatron 407 Printer at a speed of 150 lines per minute. Total printing time is about three hours per Division. (A program revision now in process will eliminate the heading card merging operation, supplying this information from tape in correct sequence during the 602 program running.)
6. As soon as printing for a Division is completed, forms are machine burst, carbon removed, and distributed.

The above description of procedures is, of course, a gross over-simplification of the actual programs, but it may convey the general pattern of electronic data processing that is involved. As a measure of actual complexity, Programs 601 and 602 each contain over one thousand machine "commands" and each utilizes all of the 4,000-word magnetic drum for program, tables, and data assembly areas.

At one time or another during the Sales Analysis Runs, every component machine in the Datatron System is in use.