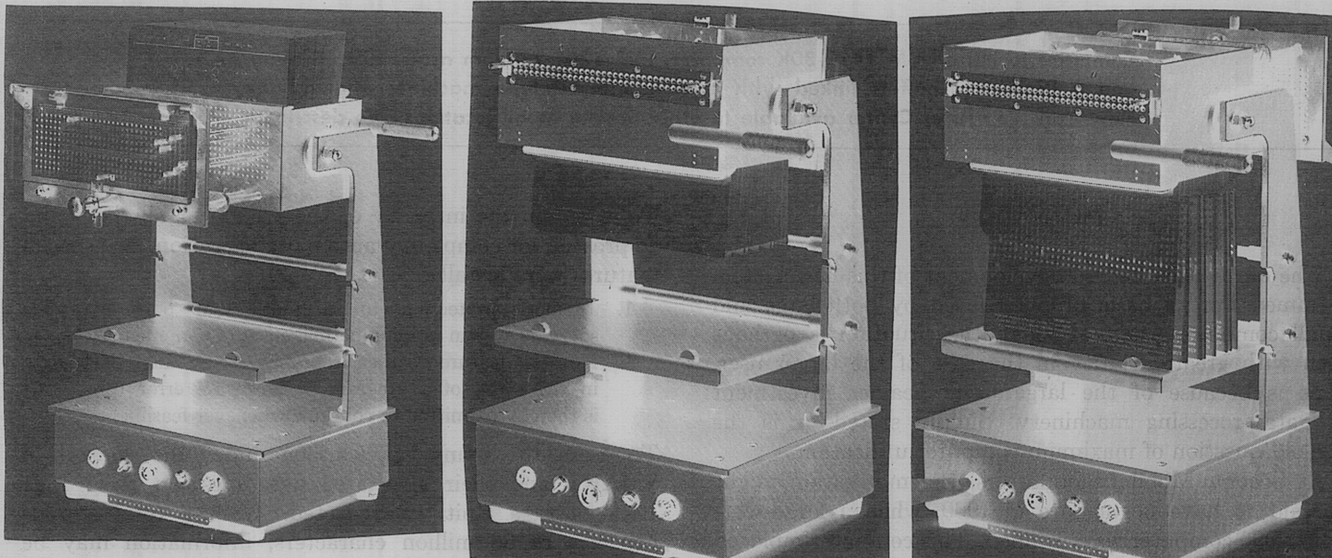

Figure 9. Operation of the selector from programming sheet

First stage: A programming sheet placed on the selector and needled; a pack of cards inserted in the selector. Second stage: The selected cards protruding downwards. Third stage: The selected cards completely separated.

## APPLICATIONS

This information system answers varied questions relating to a single notion or any possible combination of notions according to the needs of the moment. Here are a few questions to illustrate the capacities of this tool:

How can I condense a carbonyl group with a methylene group to form an olefin double bond?
Which methods can be used to convert a carbonyl compound into an olefin?
By which methods can olefin double bonds be produced?
How can acetophenones be halogenated in the ring, and how in the $\omega$-position?
How can C-alkylation be effected without simultaneous alkylation of hydroxy group?
Which reaction can be effected with metal hydrides?
How can I prepare *tert*-butyl chloride?
How can the ester group of peptides be saponified without hydrolyzing the peptide bridge?

## ACKNOWLEDGMENT

The author thanks Labaz, of Brussels, for having allowed publication of this work; the editors and publishers of "Organic Syntheses" and "Organic Reactions," "Cahiers de Synthése Organique," and "Namenreaktionen der organischen Chemie" for so readily offering their invaluable collection of procedures; and Professor Karrer of Zürich for his energetic promotion of this work.

# SMART (Socony Mobil Automatic Real Time) Computer System*

H. B. WILSON
Computer-Communications Department, Socony Mobil Oil Company, Inc., New York, New York

Received December 9, 1965

Until only recently, the phrase "real time" among computer users had come to be associated with highly specialized government applications, for example, in the guided missile and space satellite projects. The extremely large investment in computer hardware, system software, and communication facilities for such projects precluded the application of this concept to commercial areas for many years.

At Socony Mobil Oil Company headquarters, however, this advanced concept has been turned inward, as it were. Late in 1962, the SMART computer-communications system was put into operation which enabled Socony personnel at remote locations to transmit data to the computer center, direct the computer to perform certain computational and/or retrieval functions, and to receive "answers" in a matter of seconds. The objective was to improve computer utilization by significantly reducing the "turn-around time" previously existing between the moment when an "order" for services was received at the computer center and the time when it was fulfilled.

* Presented before the Division of Chemical Literature, Symposium on Equipment for Processing Information, 150th National Meeting of the American Chemical Society, Atlantic City, N. J., Sept. 13–16, 1965.

**Computer facilities (IBM 1410-80K computer, IBM 1301 random access disk memory, and a Digitronics buffer) linked with conventional teletype machines, making the New York Computer Center available to other Socony Mobil locations, are described.**

## MOTIVATION

The motivation behind development of this new system stemmed from a basic problem faced by both large and small computer users, although the situation is perhaps seen with greater clarity in the case of the larger user. This is because of the large users heavier investment in data-processing machinery. Simply stated, it is the age-old question of maximum computer utilization.

At Socony Mobil headquarters, the motivation might be said to have begun late in 1960, when a large-scale IBM 7090 computer was installed. This computer replaced the slower IBM 704 which had previously been the tool applied to the solution of technical and scientific problems. At about the same time, installation of smaller IBM 1401 data-processing systems began and the IBM 650, which had hitherto solved commercial-type data processing problems, was released.

Sophistication in the use of computers grew with the 7090 as we began to place this system in operation on commercial as well as technical problems, with the 1401 computers largely acting as satellites for data reduction and input/output purposes.

Early in 1961 it became apparent that the use of computers as management tools in our company would increase at a rapid, sustained pace. It appeared to Computer Department management that the solution to the problem of handling an expanding workload might not lie entirely in the realm of additional machines and operating staff.

A study was undertaken to devise ways to use existing equipment more intelligently, and more automatically. An important outgrowth of this study proved to be that "turn-around time" was often more critical than actual computer "run time." If something could be done to provide for more automatic operation with less human intervention, a greater volume of work could be handled in a given time period for a given dollar cost.

Progressing in this direction the concept of a "real-time system" developed and a research project was undertaken to devise such a system for future use. Throughout 1961 this research project was pursued.

## HARDWARE SELECTION

At the heart of the system, a standard IBM 1410 computer, with "real-time" linkages, was selected. This permitted the release of one installed 1401 system, thereby reducing system costs and yet maintaining upward compatibility of the job library created for the 1401.

The initial 1410 configuration contained the following standard features:

1. 40K main memory
2. Interrupt or trapping ability to facilitate priority processing
3. Dual input/output (I/O) channels with the ability to "overlap" I/O and main memory cycles

To maintain the integrity of the system at all times and to provide for completely automatic operation, two special features were installed:

1. Memory protection—to safeguard that portion of the software system residing permanently in core memory
2. Nonstop feature—causes a branch to an analysis routine in the event of a program or hardware error (processing is thereby permitted to continue whenever feasible)

To store the many applications under the control of SMART and their associated data files, an IBM 1301 Disk Storage Unit was selected. With a total storage capacity of 56 million characters, information may be accessed in from zero to a maximum of 180 milliseconds.

The selection of a communications buffer, in retrospect, proved to be a most difficult undertaking. Buffering equipment announced in early 1961 was either found to be deficient for our purposes or excessively priced. As a result, a contract was awarded to the Digitronics Corporation to construct a communications buffer to Socony specifications.

Some of the more important features of this unique buffering equipment are:

1. Capacity to accept messages from up to 64 remote terminals simultaneously
2. Unit line transfer from or to the computer (each terminal has enough core memory to accommodate the longest line)
3. Automatic code conversion (Baudot ↔ BCD)
4. Real-Time Clock and an Interval Timer which cycles each $\frac{1}{10}$ of a second
5. Transmission echo check for informing the 1410 of the disposition of output

Completing the basic hardware configuration are the remote I/O devices linking SMART and the real world. With close to 200 domestic TWX and international TELEX units installed in company locations throughout the world, for normal business messages, their selection as the I/O medium was obvious.

## SOFTWARE SYSTEM

Although originally conceived in total, the present, highly sophisticated level of software development was actually achieved through a series of well-defined steps over a period of two years.

The initial version of SMART became operational during the fourth quarter of 1962. This version, which is described in more detail later, provided only in-house utilization (within Socony headquarter offices). During this period programmatic considerations involving on-line, man-machine communication were tested and refined. SMART was then opened to domestic TWX and international TELEX users for remote computing services the second quarter of 1963.

With what now amounted to a direct line from any location in the Mobil organization to the New York Computer Center, the second major modification involved

SMART in the area of pure data communications. The ability to accept, edit, and otherwise prepare Linear Programming models, Fortran source language, etc., for immediate computer processing resulted in an average decrease of 50% in total job "turn-around" time for these important areas of application. An additional side benefit occurred in that other data communication facilities were released, thus reducing costs.

The last major modification to SMART was brought about by IBM's announcement of the 1410 Operating System in 1963. The SMART System in combination with the operating system now permits maximum utilization of the computer since regularly scheduled jobs may run concurrently with "real-time" processing.

## SYSTEM OPERATION

For a better understanding of the interrelationships of the equipment and programs making up the real-time system, the operations of the system have been charted and are divided into three sections: (1) input, (2) processing and operational control, and (3) output. Figure 1, a flow diagram, shows the movement of data through the system.

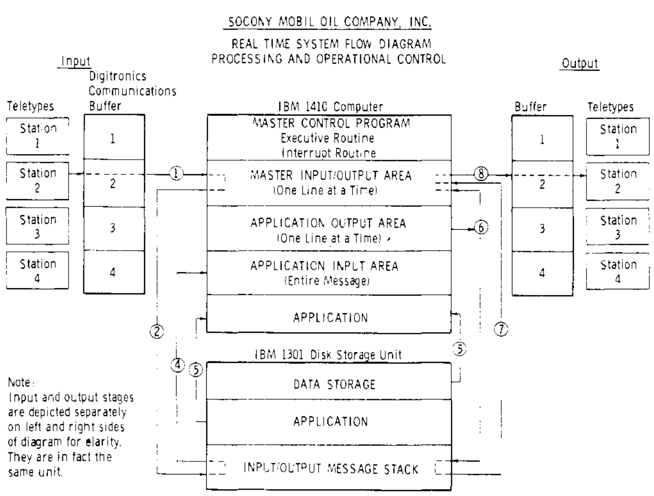


Figure 1.

**Input.** Data is entered into the system from remote stations *via* the Teletype machine (TTY), a line at a time. A line of data is variable with a maximum of 75 characters. A complete transaction for an application is also variable with a maximum of 25 lines (1875 data characters).

Input lines may be sent from all remote stations simultaneously. The Communications Buffer (C/B) holds all characters of each line until an "end of line" signal is received. When the TTY operator generates this signal as the last character of each line, the C/B then interrupts the 1410. The 1410 accepts the interrupt, stops what it is doing, and branches to the INTERRUPT routine. The function of the INTERRUPT routine is to monitor and provide for orderly movement of data to and from the remote stations. The INTERRUPT routine moves the line from the C/B to the 1301 disk storage, repeats the input line at originating TTY, and returns to the operations the 1410 was performing before this interrupt.

All lines of a message from a specific remote station are assembled in an input area within the 1301 until the TTY operator indicates to the 1410 by a control character that the message is complete. The interrupt routine then indicates to the executive routine that a complete transaction exists in the 1301 input area.

While the 1410 is performing these operations the C/B is independently accepting data from the other remote stations. Steps 1 and 2 on the flow diagram (Figure 1) show the data flow thus far described.

**Processing and Operational Control.** The Master Control Program provides processing and operational control through the use of the INTERRUPT routine and the EXECUTIVE routine. While the INTERRUPT routine provides for the orderly movement of individual input/output lines, the EXECUTIVE routine provides for the orderly movement of the entire input transaction from the processing stack through completion of processing and creation of output.

The EXECUTIVE routine takes each transaction from the input stack in the 1301 and moves it to the 1410 memory on a "first come, first serve" basis. It then moves from the 1301 the application and any associated data file, necessary to process the transaction into memory, and turns operational control over to that program.

The application performs the operations dictated by the transaction, creates the output, and returns operational control to the EXECUTIVE routine. If the remote station designated to receive the output is busy, the EXECUTIVE routine puts the output in the 1301 for later transmission. If the remote station is not busy the EXECUTIVE routine moves the first line to the C/B for transmission and puts the remainder of the output in the 1301 output stack. The EXECUTIVE routine then checks the input stack for the next transaction. Steps 3, 4, 5, and 6 of the flow diagram show the data flow through processing and operation control.

**Output.** When a TTY operator signals the 1410 that an input message is completed, the 1410, before allowing entrance of the next input message, determines if there is output for that remote station. If there is output the 1410 moves a line to the C/B, and the C/B transmits the line to the remote station. After the C/B has sent the last character of a line to the remote station, it sends an interrupt to the 1410. When the 1410 is interrupted, it moves the next output line to the C/B for transmission. All output stored for a specific remote station is transmitted before an input message can be sent. If no more output is found, the 1410 allows input to be sent. Steps 7 and 8 of the flow diagram show the flow of output data.

## APPLICATIONS

Up to this point applications have been mentioned only in passing. Certainly these are the foundation for any successful computer system. SMART presently contains over 50 diverse production applications, and their number is constantly increasing.

Engineering, mathematical and/or statistical, as well as financial applications have been incorporated.

Many of the engineering applications are indigenous to the petroleum industry such as the Flash Calculation which will analyze the amount of vaporization and liquid volume that will result for certain petroleum components when pressure and temperature are varied.

Several data-processing applications are also included in SMART. An example of this type application is the Manpower Retrieval Program. Given a set of "requirements" (in terms of education, experience, skill, job level, and other characteristics), this program will search the entire manpower inventory file and select those employees fulfilling all of the requirements.

SMART manuals have been prepared containing the necessary operating instructions and detailed write-ups for every production applications. At last count well over 400 such manuals have been requested by company employees throughout the world.

## SYSTEM STATISTICS

To evaluate system performance in terms of dollars and cents, SMART automatically maintains utilization statistics by remote location and application. Also maintained is the actual elapsed computer time to produce a solution.

Although solution times may vary from a fraction of a second to several minutes, using 1964 statistics the average solution time, for the 20,000 problems submitted during that year, required 12 seconds on the 1410 with a resulting company cost of 35 cents. This amount is particularly significant when viewed in comparison to the several man-hours or man-days required to arrive at comparable solutions by manual methods.

## CONCLUSION

The primary advantages of SMART may be summarized as follows:

A. Operating Efficiency
   1. Total job time has been significantly reduced.
   2. Maximum utilization of the computer is achieved since regularly scheduled jobs may be running concurrently.
   3. Computer operator errors and job setup time are virtually eliminated.

B. Remote Capabilities
   1. The large scale computing capabilities of the New York Headquarters have been linked, through existing telephone facilities, to any point in the worldwide Mobil organization.
   2. As many as 64 different company locations may use these facilities simultaneously.

C. Time Current Information
   1. Up-to-date operating information is maintained on the computer's mass memory and obtained by company management within seconds.

It is for these reasons, as well as the direct system savings, that we are convinced the SMART system has contributed to our ultimate goal—company profitability.

---

# Chemical Substructure Searching with Linear Notations*

BEATRICE A. MARRON, GLORIA R. BOLOTSKY, and STEPHEN J. TAUBER
Center for Computer Sciences and Technology, Institute for Applied Technology,
National Bureau of Standards, Washington, D.C. 20234

Received February 8, 1966

**A technique for doing chemical substructure searching directly on linear notations, an input format, is described. Some of the implications and limitations are discussed.**

Among the requirements for a comprehensive chemical information storage and retrieval system is the ability to search a file of chemical structures both for individual compounds and for classes of compounds defined by any desired substructure. There is presently greater emphasis on searching chemical structures than related information both because of the very basic nature of structure information, and because it lends itself to mechanical searches.

Chemical structures are usually shown in written communications as networks of atoms and bonds, but linear notations have an advantage for computer input because they consist of sequences of symbols. Here is described an experimental system developed at the National Bureau of Standards to perform chemical substructure searches on Hayward linear notations. The general method should be applicable to other linear notations.

## INPUTS

The basic unit of information in the system is the chemical structure. Each structure as it enters the system is assigned an eight-digit identification number.

* Presented in part at the 1965 Congress, International Federation for Documentation, Washington, D.C., Oct. 12, 1965.