

ENGINEERING RESEARCH INSTITUTE  
UNIVERSITY OF MICHIGAN  
ANN ARBOR

Informal Memorandum 5

LANGUAGES FOR ANALYSIS OF CLERICAL PROBLEMS

ARTHUR W. BURKS  
IRVING M. COPI  
DON W. WARREN

Note: This material is issued as an informal memorandum because it is tentative, it has not been completely checked, and the exposition is unpolished. Administrative circumstances prevent further work on the material. Consequently, it is necessary to present the results so far achieved in this form rather than in a complete regular report.

Project 1828

BURROUGHS CORPORATION  
RESEARCH CENTER  
PAOLI, PENNSYLVANIA

November, 1954

ENGN  
UMR 1023

## TABLE OF CONTENTS

Section	Page
PREFACE	iii
1. PREVIOUS WORK	1
1.1. Description	1
1.2. Discussion	4
2. AN ALTERNATIVE APPROACH	5
2.1. Defining the Problem	5
2.2. Modifying the Program	6
2.3. The Hierarchy of Languages	8
3. ABSTRACTION LANGUAGES	9
3.1. The Application of Logic	9
3.2. An Intuitive Abstraction Language	11
3.3. A More Rigorous Abstraction Language	14
FIGURES	19

## PREFACE

This report summarizes the work done by the Ann Arbor group in attempting to extend the results obtained in 1950 by the Burroughs Research Laboratory's Methods Development Section under the direction of James E. Myers. Those who participated in the follow-up work herein **described** had the following aims: first, to formulate the original problem in sharper terms; second, to uncover some of the general questions involved; third, to construct for the clerical processing of business data a sequence of languages of which the last could readily be coded for machine use; and fourth, to explore the directions in which accomplishment of the first three aims might suggest improved methods of processing business data and new or better uses of the results of such processing. In the course of intermittent work on the problem it was discovered that these aims were considerably beyond the reach of a small group with such specialized backgrounds and at such a distance from the source of the problem data. However the problem was thought through rather carefully. Although no solution can be offered, a number of suggestions will be made. These are concerned mostly with the sequence or hierarchy of languages in general and the formulation of the "Abstraction" language in particular. They are presented in the hope that they may be of use to anyone undertaking further work on the problem.

## Informal Memorandum 5

### LANGUAGES FOR ANALYSIS OF CLERICAL PROBLEMS

#### 1. PREVIOUS WORK.

##### 1.1. Description.

The aim of the Methods Development Section in working on this problem, as described in Burroughs Research Memorandum Number 14-J.E.M., November 15, 1950, was "...to produce a number of processed statements of any problem...which...represent successively greater degrees of abstraction and summarization; each to serve the needs of groups with their own specialized decisions to make." The "statements" in question are alternative ways of formulating, in varying detail, an accounting procedure designed to accomplish a given aim. Four such statements were "roughed out", and were labeled "Symbolysis", "Preliminary 'Machine' Scheduling", "Abstraction", and "Character Handling Parameters." The different amounts of detail to be contained in these four statements were characterized by Myers as "Great", "Considerable", "Little", and "None".

The first of these statements, the Symbolysis, was based directly on detailed reports made by observers in the field. Those observers were assigned to investigate and report on the actual day to day and desk to desk processing of information contained in invoices, credit reports, customer statements, etc. From these detailed reports was constructed a large flow chart with special symbols representing the operations of sorting, posting, checking, etc., performed on those invoices, credit reports, customer statements, etc., as they were moved through the business offices. The large flow chart constitutes the Symbolysis, and was intended to embody an overall but detailed picture of the actual clerical procedures in use by the business concern being studied (the DuPont Paint Stores, investigated by the Methods Development Section in 1950).

The second statement, the Preliminary "Machine" Scheduling, was a first estimate of the types of machines which could be utilized to perform the various operations described in the Symbolysis. Such a scheduling of machines, with approximate specifications of their capacities and of the order in which they would receive, process, and transmit information to other machines, clearly provides an alternative description of the way in which (the same, or similar) data could be routed and processed in the establishment under study. It was the intention of the Methods Development Section

that this second statement should be less detailed than the first, and somewhat tentative, (as signalled by their use of the word "preliminary" in the name of this second statement).

The third statement, the Abstraction, was intended to be a flow chart like the Symbolysis, but different from the latter in three respects. First, it was to contain considerably less detail. Second, it was to emphasize more the abstract or logical aspects of information processing, and to be much less concerned with the particular methods actually used for manipulating and transmitting information. (It was not clear to what extent the actual procedure of the business concern being investigated was to be adhered to, or to what extent alternative procedures were to be allowed. Such alternatives might include a change from monthly billing to cyclic billing, or dispensing with certain forms sent to the customer, or the like.) Third, whereas the notation used in the Symbolysis is closely related to business usage, the Abstraction makes use of special symbols similar to those of mathematics and logic in representing the information and its processing.

The fourth statement, the Character Handling Parameters, was intended to be a very concise but quantitative description of the amount of clerical work performed in the business concern investigated. The aim was to have it consist of only a page or two of very condensed verbal and numerical material which could be utilized by top management of the Burroughs Corporation, as for example in making decisions about the feasibility of continued research and in estimating the cost and saving of mechanization. The Methods Development Section felt that for such an adequate fourth statement to be prepared, it would be necessary to devise some basic unit capable of being used in measuring all different kinds of clerical work. Such a unit would be analogous to the binary digit as the simplest unit of (any kind of) information, or possibly to the therblig as a basic unit of (any kind of) physical labor.

Since most of the work of the Ann Arbor group was focused on developing languages appropriate to descriptions of clerical operations along the general line of the third statement, and since Section 3 of the present report will be concerned entirely with such Abstractions, we now enlarge on our summary of the third statement as conceived by the Methods Development Section.

The Abstraction is a large chart (e.g., a standard D-size drawing) with an attached glossary. The chart contains representations of four different kinds of elements: Inputs, Outputs, Maintained Records and Permanent Data, and Correspondences and Computations. Each element of any of these kinds is represented on the chart by a rectangle or box.

(1) Inputs. Each box labeled as an input represents a type of message or report (e.g., an invoice or a credit report) which originates outside of the clerical operation being studied. The box which represents a particular message-type will contain symbols representing the various items of information that usually appear in messages of that type. Arrows between symbols or bracketed groups of symbols are used to indicate logical or mathematical relations between the items represented by those symbols (e.g., between "b" and "s" where "b" represents the name of a branchstore and "s" represents the region of that branch store's location, and between "ψy" and "z" where these three letters represent respectively the unit price of a product, the number of units sold, and the dollars charged for the product/transaction ).

(2) Outputs. Each box labeled as an output represents a type of message or report (e.g., a customer's statement or a statistical report) which originates inside the clerical operation being studied and is sent for use outside, either by customers or by other departments of the firm. Output boxes, like input boxes, contain symbols representing items of information and arrows indicating relations among those items.

(3) Maintained Records and Permanent Data. The boxes bearing this third label represent storages of processed information (e.g., ledgers and account books) for use within the clerical operation being studied. These boxes too contain symbols and arrows showing the mode of organization of the data they contain.

(4) Correspondences and Computations. One large box has this fourth label. Its different parts contain sequences of symbols which are labeled to represent the kind and amount of data transferred from inputs to outputs, from inputs to records, from records to outputs, and the various computations performed in the course of such transmission of data.

Boxes of these four types are connected in the Abstraction chart by arrows representing lines of flow of information. Each arrow is labeled to show the quantity of information flowing along the path represented by that arrow. Each of the boxes also has its own quantitative labels indicating the number of alphameric (alphabet letters or numerical symbols) characters contained in a typical message or entry of the type represented.

Information is presented in an Abstraction by means of special symbols which represent familiar types of items from business records (e.g., "θ" represents invoice number, "e" represents the customer's name and address, " $\sum_d$ " represents the sum of all of a customer's remittances for date d, etc). In addition to these arbitrary symbols for standard items, verbal notes are sometimes inserted for non-standard operations. The use of arbitrary symbols, which effects a considerable saving in space, requires that the Abstraction chart have attached to it a glossary explaining what meanings are attached to those symbols.

## 1.2. Discussion.

The members of the Ann Arbor group are in strong general agreement with the approach described above. On a few specific parts of that program, however, they have some reservations and criticisms.

First, it seems doubtful that the same chart can succeed in presenting an overall view of a large-scale clerical operation in its entirety and at the same time contain great detail. The initial survey may result in hundreds of pages of reports describing the actual clerical activities performed. If the Symbolysis based on those reports were to include anything approaching complete detail, it would be tremendously unwieldy and of doubtful utility.

Second, it seems premature to prepare a machine schedule, even a preliminary one, before working out an Abstraction. The Methods Development Section apparently regarded the Abstraction as a device for modifying a machine schedule already at hand, whereas it ought rather to be viewed as a prerequisite for devising a machine schedule. A closely related point will be presented in some detail at the beginning of Section 2. (Of course it may be that business considerations demand the earliest possible machine scheduling; if so, it must be regarded as highly tentative and subject to later modification when the Abstractions have been worked out.)

Third, the Abstraction presupposes the development of a notation or language which will show the interrelations of the items symbolized more adequately than does the notation used by the Methods Development Section. This remark is not intended to labor the fact that the previous work was incomplete, but to point out a specific gap that must be filled.

Fourth, although the Abstraction is a fundamentally sound and fruitful device, the notion of "Correspondences and Computations" seems too broad a category to be useful, and might better be either replaced by or divided more explicitly into a number of subcategories, such as computation, checking, posting, searching of records, etc.

Fifth, the "Character Handling Parameters", whose statement constitutes a very useful phase of the program, has been so conceived as to involve the extremely difficult problem of devising an adequate unit of clerical activity. Even the narrower notion of a unit of computation remains undefined, despite its having been given considerable thought by many computer experts. It is perhaps not even a well formed or significant problem as posed in Memorandum Number 14-J.E.M. For since operators with different types of desk calculators must perform different numbers of operations to accomplish the same computation, the



number of units of clerical operation required for a given computation is surely relative to the equipment used. We shall discuss this general problem further in the following section.

## 2. AN ALTERNATIVE APPROACH\*

### 2.1. Defining the Problem.

The problem initiated by the Methods Development Section is that of devising a general technique for mechanizing the clerical operations of a business concern. Certainly the first step in accomplishing any such mechanization is to observe the clerical operations actually performed by the concern's employees speaking ordinary business English and using standard office equipment, e.g., typewriters, files, ledgers, and probably desk calculators. But subsequent steps of the mechanization will differ according to different ways in which the task is conceived. The task can be viewed in any one of (at least) three ways.

From one point of view, subsequent steps should be the following. The actual clerical routine observed should be described in more abstract terms, as by an Abstraction. Then a schedule should be constructed for business machines: (a) into whose internal language messages expressed in ordinary business English can be coded, and (b) which will perform all of the clerical operations just as they were performed by human beings, but more quickly, and, in the long run, more economically. The assigned task, from this first point of view, is simply to mechanize the existing clerical procedure for the sake of higher speed and lower cost.

From another point of view, after the first step the procedure should be somewhat different. Instead of describing the existing clerical routine in abstract terms, the second step should consist in first fixing the existing types and quantities of the inputs received and outputs delivered by the clerical staff, and then searching for the most economical method, however different it might be from the one actually in use, of flowing information from those fixed inputs to those fixed outputs. The ultimate aim from this second point of view is twofold: first, to improve or rationalize the clerical procedure of accomplishing a fixed and definite assignment, and, second, to mechanize that improved procedure.

Before examining the task from still a third point of view, it should be noted that the desirability of a specified output depends

---

\*The "alternative approach" here presented is based upon the previous work of James E. Myers and the Methods Development Section, with many parts of which it is in agreement.

upon its cost, and that its cost is a function of both the cost of the inputs required for its production and the accounting procedure being used. The first functional dependence is obvious; the second may be explained by considering that the given output may be a very inexpensive by-product in one accounting procedure, but a very costly additional step in a different one. Hence to optimize the accounting routine of a given business one ought ideally to consider all available inputs together with the cost of each, and all possible outputs together with the value of each. These considerations lead naturally to the third avenue of approach to our problem.

From this third point of view the task appears still more complex. Here after the same first step, neither the existing procedure (as in the first case) nor the existing input-output schedule (as in the second case) is regarded as fixed. A larger or smaller number of different inputs and outputs should be considered, together with alternative methods of flowing information between them. The problem as conceived from this third point of view is not simply to mechanize a given routine, nor to mechanize it after first rationalizing it within the limits set by traditional conceptions of the "proper" number and format of inputs and outputs. The utilization of electronic digital computing machines in its clerical operations permits a business concern to supply its other departments (e.g., sales, research, development) with more data which can be statistically processed and delivered in more convenient format. The potential advantages of such new and improved outputs, together with the ability of an electronic machine to handle more input data, or to obtain the same amount of output data from less input data, should be exploited to as high a degree as possible for full return on the investment in such equipment to be realized. From this third point of view the task is not to get the same old job done more quickly and more cheaply, but to have a bigger and better job done in less time and for less money. The problem is not merely one of mechanization or rationalization of clerical means, but includes also the rationalization of clerical ends. It is a problem in what might well be called "clerical engineering."

## 2.2. Modifying the Program.

From the third point of view discussed in Section 2.1 it might be desirable to modify the general program designed by the Methods Development Section. How such a modified program could utilize the conception of a hierarchy of languages will be explained in Section 2.3. Here we shall consider some possible modifications of particular parts of the program which are suggested by the criticisms made in Section 1.2.

First, in the initial field work phase of the task, the investigators should be encouraged to do more than merely observe and record the operations of the clerical department of the business being studied. They could well make inquiries in other departments of that business to learn what other output data it might be desirable for the clerical department to supply. And the observers should be encouraged also to include in their reports critical suggestions for combining or eliminating various steps of the operation.

Second, the detailed information collected ought to include not only the quantity of information contained in each message, and every step at which the message is handled as it is processed through the clerical operation, but also the precise amount of its information that is used at each handling, and the (average) time, delay, and cost (man plus training plus machine) of each handling. This information could be used in connection with the Character Handling Parameters statement, where in place of the unavailable "unit of clerical work" we might have the unit cost per clerical operation with specified equipment. The latter could be more meaningfully compared with the cost for mechanized operation of all or part of a specified routine for accomplishing the same purpose.

Third, the Symbolysis might better serve its intended function if the flow chart (or charts) were supplemented by a prose description and analysis of the operation charted. That written account could contain complete details, which would permit the flow chart(s) to be less detailed and more schematic. It would be desirable to draw the chart(s) and write the description in such a way as to permit a maximum amount of easy cross reference between them. In preparing the charts it would be helpful to have the flow paths labeled to permit easy identification of confluent and branch paths (e.g., paths A, B, C join to make path ABC, and path D branches into paths D<sub>1</sub>, D<sub>2</sub>, D<sub>3</sub>, ...). And in preparing the prose account serious attention should be paid to the observers' suggestions for combining, dividing, or eliminating steps. If some of those suggestions seemed particularly fruitful, they might well lead to the preparation of alternative Symbolyses (although in general the working out of alternative procedures might better be done in preparing the Abstraction statements).

Fourth, the preparation of several alternative Abstractions should be undertaken. In ~~them~~ the greatest freedom should be allowed, not only in the method of flowing information from inputs to outputs, but also in considering changes in the inputs and outputs themselves. Only those items should remain constant which are absolutely beyond control, such as the necessity of preparing tax forms or of accepting payment by check. These alternatives can be checked with supervisory personnel for the selection of an optimum, but all reasonable alternatives should be retained for final decision in the light of machine specifications when the actual mechanization is to be determined.

### 2.3. The Hierarchy of Languages.

Clerical operations involve the reception, processing, and transmission of completely specific descriptive and numerical information. Ordinary business English is the language in which these operations are traditionally performed, and is adequate also to describe those operations. The internal language of electronic digital computers is also adequate for the performance and description of clerical operations. But it is not at all a simple matter to effect a translation from ordinary business English to the internal language of an electronic office machine. Moreover, the problem of translation is further complicated if it is decided that we want not just a mechanization of the existing clerical routine, but the mechanization of a more rational clerical routine for the accomplishment of somewhat more ambitious purposes.

To effect such a translation from ordinary business English to the internal language of electronic computing machinery, it seems advisable to construct a hierarchy of languages. The "statements" proposed by the Methods Development Section are expressed in different languages of this hierarchy.

At one end of the proposed hierarchy of languages is ordinary business English, and at the other end is the internal language of electronic digital computing machinery. Before discussing the intermediate (and auxiliary) languages, it will be helpful to compare and contrast the two ends from two points of view, or to place them relative to each other along two different dimensions.

Ordinary business English is as far as any other language of the hierarchy from the completely coded internal language of the machine, which may be expressed by saying that it is minimal with respect to degree of mechanization, whereas the internal machine language is maximal with respect to degree of mechanization. Both of these end languages are altogether concrete, adequate to express the completely specific descriptive and numerical information present in the clerical operation. We may say, therefore, that both languages are minimal with respect to degree of abstraction. If we diagram the various languages of the proposed hierarchy on a graph, as in Fig. 1, whose vertical axis rises to represent increasing abstraction from clerical detail, and whose horizontal axis goes right to represent increasing mechanization, then both end languages will be located at the horizontal axis, ordinary business English at the extreme left and the machine's internal language at the extreme right.

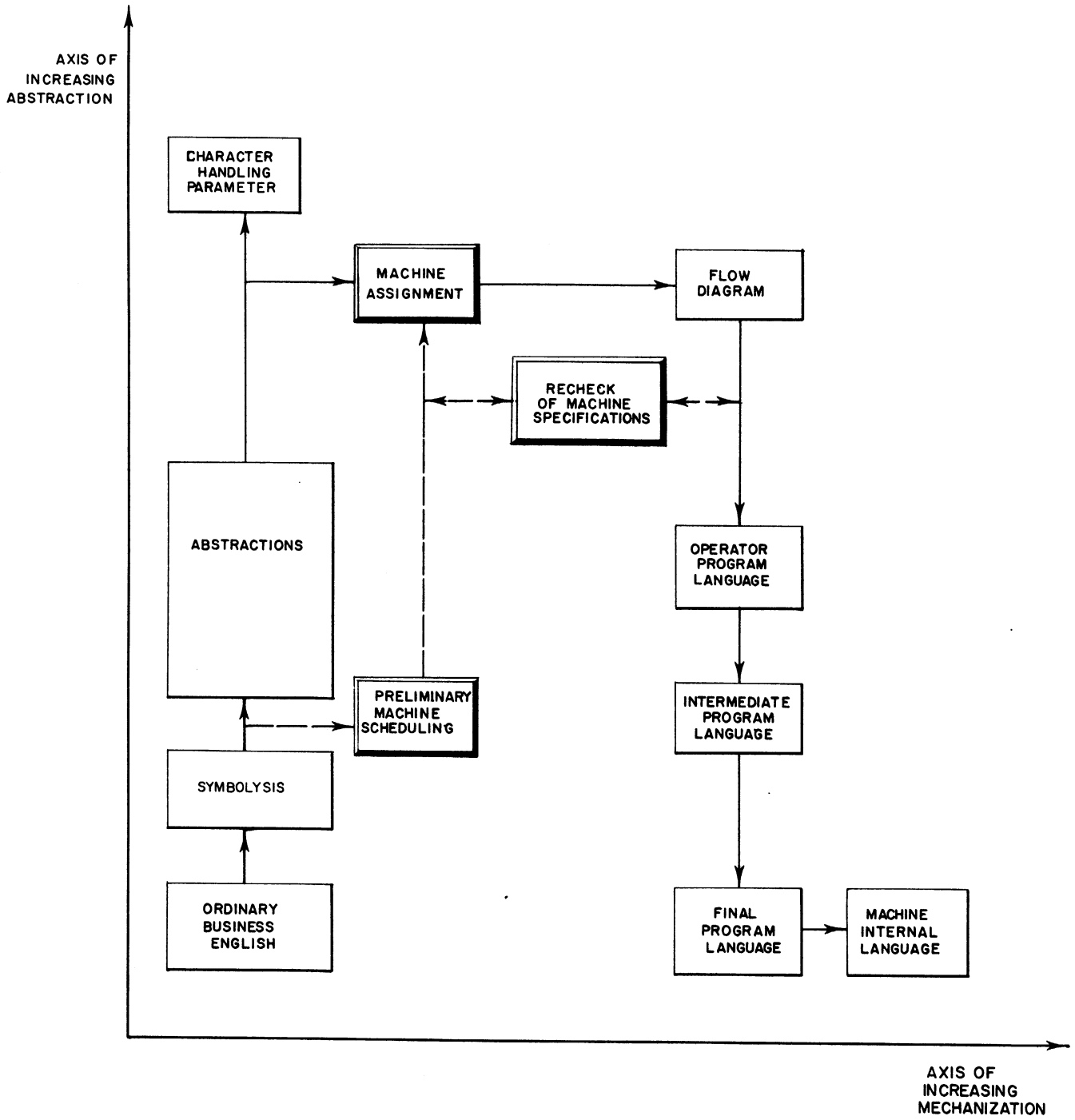


Fig. 1

From the lower left of the diagram, we proceed up to the more abstract language of the Symbolysis, and on up through a set of Abstraction languages. The path divides above the latter, one branch going up to the still more abstract Character Handling Parameters statement, the other turning for the first time in the direction of increasing mechanization. A decision having been made, on the basis of work with the Abstractions, as to what clerical routine is to be mechanized, the next step is the Machine Assignment, which is not a language of the hierarchy on the same footing with the others, but rather a necessary stage in passing from the Abstraction to the Flow Diagram. The three boxes with double walls are not proper languages within the hierarchy, but rather auxiliary engineering languages. Their distinctness is also suggested by the use of broken lines to indicate their connections.

The notion of a Flow Diagram as a usual first step in preparing a problem for machine solution is sufficiently familiar in computer work to require no explanation here. But the next box, labeled Operator Program Language, which is lower in degree of abstractness, and closer to complete mechanization, is not so familiar a concept. All that is intended here is a verbalization of what is represented by the flow diagram, to the end of permitting an easier translation of it into the Intermediate Program Language. In the direction of greater mechanization and lesser abstraction is the Intermediate Program Language, in which the routine for the machines is specified more precisely and more technically, but in a somewhat less esoteric and more convenient dialect than can be fed directly into the machine.\* The Final Program Language is that in which the routine in all its specific detail is coded for insertion into the machine, which then translates the routine into its own Internal Language.

### 3. ABSTRACTION LANGUAGES.

#### 3.1. The Application of Logic.

The notations and techniques of symbolic logic can be used in various different ways to describe clerical operations. Their introduction provides a seemingly unlimited number of different possible modes of expression. Although some appear to lend themselves better than others to the preparation of an Abstraction, the Ann Arbor group was unable to decide which application of symbolic logic led to the ideal Abstraction language. The following techniques are

---

\*See Burks, A. W., An Intermediate Program Language as an Aid in Program Synthesis, Project Engineering Research Institute Report 1828 (Burroughs Corporation, Research Center, Paoli, Pennsylvania), Ann Arbor, January 25, 1951.

presented not as final answers, but as indications, for purposes of future research, of what the application of logic can contribute to the development of Abstraction languages.

The various applications proposed are embodied in two different languages. The first, presented in 3.2, requires the user to resolve certain notational ambiguities intuitively, on a common-sense basis. The second, presented in 3.3, utilizes a notation that is enlarged in the direction of greater self-sufficiency and rigor. The first may be more useful in describing alternative accounting procedures, while the second is a closer approximation to the language used in programming a problem for a computing machine. Both languages are applied to the same clerical problem: an abbreviated version of the DuPont Paint Stores operation. Attention is confined to the problem of representing the data organized at four of the most important points of the accounting process. No consideration is given to other important features of the Abstraction, such as correspondences, computations, and the interconnecting flow-paths.

Both languages use letters from the early part of the alphabet, a,b,...,h, as descriptive variables; letters from the latter part of the alphabet, p,q,...,w, as numerical variables; and letters from the middle part of the alphabet, k,m,n, as descriptive-numerical variables (e.g., invoice numbers). The letters i,j,x,y,z are reserved for use in our metalanguage, and the letters l and o are not used at all. The following assignment of meanings to these letters and to the special symbols  $\Sigma$ ,  $L$ , and [ ] constitutes our complete glossary for the first Abstraction language, and is used, with some additions, for the second also.

a	remittance identification
b	branch store (or department)
c	customer (name and address)
d	date
f	conditions of sale
	$f_1$ cash
	$f_2$ charge
	$f_3$ dealer
g	name of goods (products)
h	product class
k	identification of sale (invoice number)
q	quantity (of goods sold)
r	dollar amount of remittance
s	dollar amount of sale (tax included)
$\bar{s}$	unit price of product
$\underline{s}$	extended price of product ( $=\bar{s}xq$ )
t	dollar amount of tax

	$\sum^u$	dollar amount of discount
	$\sum x$	the sum of all quantities $x$
$\angle_i(x_1, \dots, x_n)$		a list or sequence of n-tuples $(x_1, \dots, x_n)$ of the category determined by $i$ and order by $x_1$ (e.g., chronologically or alphabetically)
$[x_1, \dots, x_n]$		a sequence of n-tuples $(x_1, \dots, x_n)$ similar to $\angle_i$ but where the order is not significant.

### 3.2. An Intuitive Abstraction Language.

In the problem at hand--and also for the complete Paint Stores Problem as reported by the Methods Development Section -- all input data can be classified either as sales data or as remittance data. A variable such as  $k$  can occur in the context or either type of input, in the report of a sale or in the report of a remittance for a sale. It is necessary to be able to determine the proper context for each occurrence of such a variable, that is, to determine the other variables to which it is related. To permit this determination we specify the format of the two types of input data, which amounts to specifying the kind and arrangement of variables which appear in the two contexts.

For each of these contexts we define a relation to show the interdependence of the variables implied by that context. The sales relation is represented by  $R_k$ , where

$$R_k(k, b, c, f, d, [g, h, \bar{s}, q, \bar{s}], t, u, s)$$

asserts that  $k$  is the invoice number for the sale by branch  $b$  to customer  $c$  under condition  $f$  on date  $d$  of products  $g$  each of product class  $h$ , at unit price  $\bar{s}$  in quantity  $q$  for extended price  $\bar{s} \times q = \bar{s}$ , with tax amounting to  $t$  and discount amounting to  $u$ , where the total amount on the invoice is  $s$ .

The remittance relation is represented by  $R_a$ , where

$$R_a(a, r, b, c, d, [k, s, u])$$

asserts that  $a$  is the remittance of  $r$  dollars to branch  $b$  by customer  $c$  on date  $d$  covering a set of invoices of which invoice  $k$  is in the amount  $s$  after the discount  $u$  is taken.

Any occurrence of a variable will be in just one of these two relations, and that will be indicated by the context. Sometimes subscripts will be used to show the ranges of variables, as in the second line of Form II of Fig. 2, where the subscript  $a$  attached to  $k$  and  $s$  indicates that only invoices and amounts covered by the listed remittance  $a$  are permissible values for these variables (the presence of  $a$  indicates that the relation is  $R_a$  rather than  $R_k$  in which  $a$  does not occur).



I $k, c, b, d, f$ $[g, h, \bar{s}, q, \bar{s}]$ $\sum \bar{s} + t - u = s$
---

FORM I: INVOICE

II $b, d$ $[c, a, r, [k_a, s_a], \sum s_a, u]$		
RECEIPTS		
TODAY(d): $\sum s_a$	$\sum r$	$\sum u$
M.T.D.: "	"	"
SALES		
TODAY(d): $\sum_{f_1} s_k$	$\sum_{f_2} s_k$	$\sum_{f_3} s_k (= \sum_{i=1,2,3} \sum_{f_i} s_k)$
M.T.D.: "	"	"
DAY'S MONEY: $\sum r + \sum_{f_1} s_k$		

FORM II: DAILY CREDIT REPORT

Fig. 2

Form I of Fig. 2. represents an invoice. The first line of symbols (here and on subsequent forms) represents the fixed data: invoice number  $k$ , customer's name (and address)  $c$ , name of branch store  $b$ , date  $d$ , and condition of sale  $f$ . The second line of Form I represents a set of entries each consisting of a product  $g$  of class  $h$  at unit price  $\bar{s}$  in quantity  $q$  for extended price  $\bar{s}$ . Relation  $R_k$  clearly applies since the variables represent only sales data, not remittance data. Hence the products listed were sold to  $c$  under condition  $f$  by  $b$  on  $d$  and recorded on invoice numbered  $k$ . The last line shows that the invoice amount  $s$  is produced by summing the extended prices  $\bar{s}$ , adding the tax  $t$ , and subtracting the customer's discount  $u$ . The volume and route of flow of data should be shown on this form and those that follow, but since the authors have no particular recommendations as to how it should be represented it is omitted throughout.

Form II of Fig. 2 represents the Daily Credit Report submitted every day by each branch to the central accounting office.\* The first line here too shows all fixed data: the name of the branch store  $b$  and the current date  $d$ . The second line represents a set of entries each of which consists of the name (and address) of a customer  $c$ , the remittance  $a$  made by him on date  $d$  in amount  $r$  covering one or more invoices  $k_a$  whose amounts  $s_a$  total  $\sum s_a$  with discount  $u$  already taken. The remaining lines provide a break-down of the day's and month's business. The third line shows for date  $d$ : first, the sum of the amounts of all invoices for which payment was received; second, the sum of the amounts of all remittances received; and third, the sum of all discounts taken in connection with those remittances. The fourth line gives the month-to-date totals for the three categories above and can be computed by adding the figures from line three to the corresponding figures from line four of the previous day's report.

The sales figures in the fifth and sixth lines of Form II are not concerned with remittances at all, and the relation  $R_k$  applies as indicated by the subscripts  $k$ . The invoice amounts for the day's sales are accumulated on line five in three categories defined by the three conditions of sale: cash ( $f_1$ ), charge ( $f_2$ ), and dealer ( $f_3$ ). The expression in parentheses at the right is an alternative symbolization of the same data. In line six the running totals within each of these categories appear. The last line reports the day's income as the sum of remittances and cash sales.

---

\*The type of Daily Credit Report symbolized is patterned after that of the DuPont Paint Stores, and is quite different from that which appears in the Abstraction prepared by J. E. Myers and the Methods Development Section.

Form V of Fig. 3 (the Form Numbers are taken from the Abstraction prepared by J. E. Myers and the Methods Development Section) represents the Customers Ledger. The first line contains two constants,  $d^*$  and  $f_2$ , the second of which indicates that only regular charge sales are entered as data here. The first item,  $d^*$ , is a constant only in the special sense that it is fixed at any given time that the ledger is being used; it represents the date to which the ledger has been (or is being) brought "up to date". For simplicity we may assume that the ledger is posted completely each day.

The single complex expression in the second line of Form V indicates the way in which the data presented are organized. The two bracketed sets of symbols represent the two different types of entries: first, sales for each of which are entered the invoice number  $k$ , amount  $s$ , and discount  $u$ ; and second, remittances for each of which are entered the remittance identification  $a$  and the amount  $r$ . The rightmost listing operator and the  $d$  beside it indicate that these entries are listed by date for each date  $d$  from the first of the current month (1 inst) to the present date  $d^*$ . (They could equally well be listed by date from the date of the customer's last statement, if that date is different from 1 inst, as in the case of cyclic billings. For definiteness, however, we shall assume that the billing is monthly with statements sent out on the first of each month.) For any given date  $d$  there may be entries of only one type, or there may be no entries at all. These dated entries are organized by customer  $c$ , and together with the list of sales and remittance records for each customer  $c$  is entered that customer's balance from the previous month.  $d < 1 \text{ inst} \sum (s-r)$ , which is the sum of all invoice amounts less remittances, whose dates  $d$  are earlier than 1 inst. And those entries already subclassified by date and customer are finally grouped by branch  $b$  with which the customer has his account and with which the transaction was made. More concisely, we can say that line two indicates that sales and remittance data are organized first by branch  $b$ , within each branch they are organized by customer  $c$ , and for each customer is listed his last month's balance and all subsequent sales and remittances in chronological order.

There is no indication on Form V of the physical format of its entries: a one dimensional magnetic tape would be as adequate a realization as the more conventional collection of ledger volumes. Note that the method of computing the balance  $d < 1 \text{ inst} \sum (s-r)$  is left unspecified also. Only the minimum requirement is laid down that the balance determined by all sales and remittances to the end of the previous month must be available. This minimum permits the possibility of computing that figure either monthly, daily, or anew each time the information is required.

V

$d^*(= \text{today}), f_2$

$$\triangleleft (b, \triangleleft (c, \sum (s-r), \triangleleft (d, [k, s, u], [a, r])))$$

$d < 1 \text{ inst} \quad 1 \text{ inst} \leq d \leq d^*$

FORM V: CUSTOMERS LEDGER

XI

$c, d^*(= 1 \text{ inst})$

$$\triangleleft (d [k, s, u], [a, r])$$

$1 \text{ ult} \leq d \leq d^*$

$$\sum (s-r)$$

$d < 1 \text{ ult}$

$$\sum (s-r) + \sum (s-r)$$

$d < 1 \text{ ult} \quad 1 \text{ ult} \leq d \leq d^*$

FORM XI: CUSTOMER'S STATEMENT

Fig. 3

Form XI of Fig. 3 represents the customer's statement. The first line contains the customer's name and address  $c$ , and the date of the statement  $d^*$ , which is assumed to be the first of the current month. The second line contains a listing, by date  $d$ , of all invoices and all remittances for the past month. On the third line the old balance from the previous month's statement is given on the left, and the new balance is computed on the right. Here again the abstractness warrants mention. A natural method of preparing each statement would be to copy the data directly from the up-to-date ledger. But it is not the purpose of the Abstraction to specify any such definite procedure. Hence the notation is so designed as to leave unprejudiced the method of the statement's preparation.

### 3.3. A More Rigorous Abstraction Language.

An alternative Abstraction language is presented in the present subsection. It utilizes all of the variables introduced in Section 3.1, and puts the same interpretation on them. In addition it contains a number of primitive function symbols, which are listed below together with their meanings.

$A(k)$	remittance for sale $k$
$B(a)$	branch to which remittance $a$ was made
$B(k)$	branch making sale $k$
$C(a)$	customer making remittance $a$
$C(k)$	customer to whom sale $k$ is made
$D(a)$	date of remittance $a$
$D(k)$	date of sale $k$
$D_1(d)$	date of first of month to which date $d$ belongs
$D_2(b,c,d)$	date of oldest sale by $b$ to $c$ for which no remittance has been made as of date $d$
$D_3(b,d)$	date of oldest sale by $b$ for which no remittance has been made as of date $d$
$F(k)$	conditions of sale $k$
$H(g)$	class of products $g$
$P(g)$	price per unit of product $g$
$Q(g,k)$	quantity of products $g$ sold in sale $k$
$R(a)$	dollar amount of remittance $a$
$U(c)$	tax rate applicable to customer $c$
$V(a)$	dollar amount of discount associated with remittance $a$ .

It should be remarked that discount is treated differently in this subsection than it was in Section 3.2. It is assumed here that the discount available to customer  $c$  is taken by him when he makes his remittance, but that no discount is subtracted in computing  $s$ .

In addition to the preceding primitive functions, our present Abstraction language contains a primitive dyadic relation  $\gamma(x,y)$ , where  $\gamma(b,c)$  asserts that customer  $c$  has a charge account with branch  $b$ . Terms from the glossary already presented can be used to construct the following defined relations:

$$\begin{aligned} |a,b,d| &=_{df} B(a) = b \text{ and } D(a) = d. \\ A(a,b,c,d) &=_{df} B(a) = b \text{ and } C(a) = c \text{ and } D(a) = d. \\ K(k,b,c,d) &=_{df} B(k) = b \text{ and } C(k) = c \text{ and } D(k) = d \\ &\text{and } F(k) = f_2 . \end{aligned}$$

We also construct four defined functions (in the second and third of the following definitions an expression of the form  $\hat{x}\{\dots x\}$  symbolizes the class of all  $x$ 's satisfying the condition  $\dots x$ ).

$$\begin{aligned} D_4(d) &=_{df} D_1 \{D_1(d)-1\} \text{ (First day of the previous month).} \\ T(k) &=_{df} \sum_{\hat{g}\{Q(g,k)>0\}} \{Q(g,k) \times P(g)\} \text{ (total of extended prices).} \\ W(b,c,d) &=_{df} \sum_{\hat{k}\{K(k,b,c,d'), d' \leq d\}} S(k) - \sum_{\hat{a}\{A(a,b,c,d'), d' \leq d\}} \{R(a) + V(a)\} \text{ (Current)} \\ S(k) &=_{df} T(k) + T(k) \times U(C(k)) \text{ (Invoice amount).} \end{aligned}$$

The present notation is intended to be quite rigorous, and very nearly self contained. It would be much less convenient to use than the language constructed in the preceding subsection, but it is considerably closer to the sort of language one would need to program clerical problems for an electronic digital computer.

Form I\* of Fig. 4 represents an invoice. Its first line contains the invoice number  $k$  and the values of four functions of  $k$ : its date  $D(k)$ ; the name and address of the customer to whom the sale was made  $C(k)$ ; the branch making the sale  $B(k)$ ; and the conditions of the sale  $F(k)$ . The second line contains a listing, by product, for each product whose sale is entered on invoice  $k$ , of entries each of which contains: the name of a product  $g$ , the class of that product  $H(g)$ ; the quantity of products  $g$  sold in sale  $k$ ,  $Q(g,k)$ ; the unit price for product  $g$ ,  $P(g)$ ; and the result of multiplying the quantity of product  $g$  sold in sale  $k$  by the unit price for product  $g$ ,  $Q(g,k) \times P(g)$ , which is the extended price. The third line contains the total of those extended prices,  $T(k)$ . The fourth line contains the result of multiplying the total of the extended prices by the tax rate applicable to the customer to whom the sale was made,  $T(k) \times U(C(k))$ , which is the tax for that sale. The fifth and last line of the invoice shows the invoice amount  $S(k)$  which was defined to be the sum of the total of the

I\*

$k, D(k), C(k), B(k), F(k)$

$\bigwedge (g, H(g), Q(g, k), P(g), Q(g, k) \times P(g))$   
 $\hat{g} \{Q(g, k) > 0\}$

$T(k) (= \sum_{\hat{g} \{Q(g, k) > 0\}} \{Q(g, k) \times P(g)\})$

$T(k) \times U(C(k))$

$S(k) (= T(k) + (T(k) \times U(C(k))))$

FORM I\*: INVOICE

XI\*

$b, c, D_1(d)$

$W(b, c, D_1(d) - 1)$

$\bigwedge (d', \bigwedge (k, S(k)), \bigwedge (a, R(a), V(a)))$

$d' \{D_1(d) \equiv \hat{k} \{K(k, b, c, d')\} \quad \hat{a} \{A(a, b, c, d')\}$

$d' < D_1(d) \cdot \{(\exists k) K(k, b, c, d') \vee (\exists a) A(a, b, c, d')\}\}$

$W(b, c, D_1(d) - 1)$

FORM XI\*: CUSTOMER'S STATEMENT

Fig. 4

extended prices and the tax for that sale,  $T(k) + T(k) \times U(C(k))$ . (The amount of data per invoice and the number of invoices per day should also be shown in some convenient form, and similarly for the other forms of this section as in 3.2.)

Form II\* of Fig. 5 represents the Daily Credit Report submitted every day by each branch to the central accounting office. Its first line shows the fixed data: the name of the branch  $b$  and the date  $d$ . Its second line contains a listing, by customer, for each remittance  $a$  made by that customer  $C(a)$  to that branch  $b$  on date  $d$ , of entries each of which contains four items: first, the name and address of the customer making remittance,  $C(a)$ ; second, the dollar amount of his remittance,  $R(a)$ ; third, a listing, by invoice number  $k$ , for each invoice covered by remittance  $a$ , of invoice numbers  $k$  and the sum of all invoice amounts covered by remittance  $a$ ,

$$\sum_k \overbrace{S(k)}^{\{a=A(k)\}} ;$$

and fourth, the dollar amount of discount associated with remittance  $a$ ,  $V(a)$ . The third line of Form II\* contains three entries which show the day's receipts: first, the sum of the amounts shown on all invoices  $k$  for which remittance was made to branch  $b$  on date  $d$ ,

$$\sum_k \overbrace{S(k)}^{\{B(k)=b \cdot D(A(k)) = d\}} ;$$

second, the sum of the dollar amounts of all remittances  $a$  made to branch  $b$  on date  $d$ ,

$$\sum_a \overbrace{R(a)}^{\{a,b,d\}} ;$$

and third, the sum of the dollar amounts of discount associated with all remittances  $a$  made to branch  $b$  on date  $d$ ,

$$\sum_a \overbrace{V(a)}^{\{a,b,d\}} .$$

The fourth line contains three entries which show all receipts for the month to date: each is the sum of entries of the type appearing above it for each day from the first of the current month to the present date  $d$ . The fifth line shows the day's sales in a listing, by type of sale, for all three types, of the sum of the invoice amounts shown on invoices recording sales of that type made by branch  $b$  on date  $d$ . The sixth

\*The dot between  $B(k) = b$  and  $D(A(k)) = d$  is the usual logical notation for conjunction; it may be read as "and".



$\Pi^*$

$$\sum_{\hat{a}\{\alpha, b, d\}}^{b, d} (C(\alpha), R(\alpha), \sum_{\hat{k}\{a=A(k)\}} \sum S(k)), V(\alpha)$$

$$\sum S(k) \quad \sum R(\alpha) \quad \sum \hat{a}\{\alpha, b, d\} \quad \sum V(\alpha)$$

$$\sum_{d'=d} \hat{k}\{B(k)=b \cdot D(A(k))=d\} \quad \sum_{d'=d} \hat{a}\{a, b, d\} \quad \sum R(\alpha) \quad \sum_{d'=d} \hat{a}\{\alpha, b, d\} \quad \sum V(\alpha)$$

$$\sum_{d'=d} D_i(d) \quad \sum_{d'=d} S(k) \quad \hat{k}\{B(k)=b \cdot D(A(k))=d\} \quad \sum_{d'=d} D_i(d) \quad \sum_{d'=d} R(\alpha) \quad \sum_{d'=d} \hat{a}\{\alpha, b, d\} \quad \sum_{d'=d} V(\alpha)$$

$$\sum_{i=1,2,3} (f_i, \sum_{\hat{k}} S(k))$$

$$\sum_{i=1,2,3} (f_i, \sum_{d'=d} \hat{k}\{B(k)=b \cdot D(k)=d \cdot F(k)=f_i\})$$

$$\sum_{i=1,2,3} (f_i, \sum_{d'=D_i(d)} \sum S(k))$$

$$\sum_{i=1,2,3} (f_i, \sum_{d'=D_i(d)} \hat{k}\{B(k)=b \cdot D(k)=d' \cdot F(k)=f_i\})$$

$$\sum R(\alpha) + \sum S(k)$$

$$\sum_{\hat{a}\{\alpha, b, d\}} \hat{k}\{B(k)=b \cdot D(k)=d \cdot F(k)=f_i\}$$

FORM  $\Pi^*$ : DAILY CREDIT REPORT

Fig. 5

line shows the month's sales to date in a listing, by type of sale, for all three types, of the total sales from the first of the current month to the present date  $d$ , each day's sales indicated by the sum of the invoice amounts shown on invoices recording sales of that type made by branch  $b$  on that day. The seventh and last line reports the day's income as the sum of the total dollar amounts of remittances received by branch  $b$  on date  $d$  and the total invoice amounts shown on invoices recording cash sales made by branch  $b$  on date  $d$ .

Form V\* of Fig. 6 represents the Customers Ledger. Its single line indicates a listing, by branches, for every branch  $b$ , of listings, by customers, for every customer  $c$  having a charge account with branch  $b$ , of listings, by date, for every date  $d'$  to the present date  $d$  from the date  $D_2(b,c,d)$  of the oldest sale by branch  $b$  to customer  $c$  for which no remittance was made as of date  $d$ , such that either  $d'$  is the date of an invoice  $k$  recording a charge sale by  $b$  to  $c$ , or  $d'$  is the date of a remittance  $a$  made by  $c$  to  $b$ , of the following four types of entries.\*

The first entry is the date  $d'$ . The second entry is a listing, by invoice number, for every charge sale made by branch  $b$  on date  $d'$  to customer  $c$ , of the invoice numbers and the dollar amounts of those sales. The third entry is a listing, by remittance identification  $a$ , for every remittance made to branch  $b$  on date  $d'$  by customer  $c$ , of remittance identification  $a$ , dollar amount of that remittance,  $R(a)$ , and dollar amount of discount associated with that remittance,  $V(a)$ . The fourth entry is customer  $c$ 's current balance with branch  $b$  as of date  $d'$ . It should be noted that the present form is more comprehensive than the corresponding one of Section 3.2. In Form V\* track is kept of all unpaid bills for each customer, whereas Form V refers instead only to the last month's balance of each customer.

Form XI\* of Fig. 4 represents the customer's statement. The first line contains the name of the branch,  $b$ , the customer's name and address,  $c$ , and the date  $D_1(d)$  of the first of the current month. The second line contains the customer's balance as of the last day of the month before the immediately preceding month,  $W(b,c,D_4(d)-1)$ . The third line contains a listing of three types of entries, by date, for every date  $d'$  from the first day of the previous month  $D_4(d)$  up to but not including the first day of the current month, such that either  $d'$  is the date of an invoice  $k$  recording a charge sale by  $b$  to  $c$ , or  $d'$  is the date of a remittance  $a$  made by  $c$  to  $b$ .

---

\*The symbol  $(\exists -)$  in the expression under the third listing operator is the usual logical notation for existential quantification, and may be read as "there is at least one - such that"; and the symbol  $\vee$  is the usual logical notation for disjunction, and may be read as "or".

$$\begin{array}{c}
 \boxed{\text{V}^*} \\
 \wedge_b (b, \wedge_c \{r(b,c)\}, \wedge_{d'} \{D_2(b,c,d)\} \wedge_{k'} \{K(k,b,c,d')\} \wedge_{a'} \{a\} \wedge_{c'} \{c\} \wedge_{d''} \{d\} \wedge_{k''} \{k\} \wedge_{s(k)} \{s(k)\}, \wedge_{(a,R(a),V(a))} \{a\}, \wedge_{\{a,b,d',c(a)=c\}} \{a,b,d',c\}, \wedge_{(b,c,d''')} \{b,c,d'''\})
 \end{array}$$

FORM V\*: CUSTOMERS LEDGER

Fig. 6

The first entry is the date  $d'$ . The second entry is a listing, by invoice number  $k$ , for every sale made by branch  $b$  to customer  $c$  on date  $d'$ , of the invoice numbers and the dollar amounts of those sales. The third entry is a listing, by remittance identification  $a$ , for every remittance made to branch  $b$  by customer  $c$  on date  $d'$ , of remittance identification  $a$ , dollar amount of that remittance,  $R(a)$ , and dollar amount of the discount associated with that remittance,  $V(a)$ . The fourth and last line of Form XI\* contains the customer's current balance  $W(b,c,D_1(d)-1)$ .

UNIVERSITY OF MICHIGAN



3 9015 02653 5198