C
O
P
Y

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK
Department of Fine Arts and Archaeology

January 31, 1950

Professor Carl W. Blegen
University of Cincinnati
Cincinnati, Ohio

Dear Carl:

Here is a question which hardly concerns me personally; but it was set before me as a desirable opportunity, of which the specialists might well take advantage if they could work out a plan in common. I am here merely setting the question before you and Miss Kober, in preliminary form, to ascertain if you regard it as practicable and desirable.

In short, International Business Machines (which as you know has a close connection with, and a laboratory at, Columbia) would like to place their facilities at the disposal of some project in the solution of unknown languages, analogous to decoding operations during the war. When the idea was first presented to me, I was rather stymied (knowing little of the possibilities of IBM), and as the nearest available victim passed the buck to H. J. Spinden and the Maya language. Spinden failed to suggest a plan for the language, and came up with the Maya chronology, which would be nothing more than a matter of tables if it were not for the astronomical correlations which complicate it. In the course of a conference, I felt that the opportunity should not be lost for the Minoan-Helladic language(s).

There is a graduate student here, Miss Jane Henle, who has been working for some time on the general problem of Carian, Lycian, Cypriote, etc. and for the past half year on Minoan. In other words, she is acquainted with the material and the problems. We went over to the IBM to acquaint ourselves with the machines and the possibility of correlating them with the problems. It seems to me that if the machines can be made to do the work of compiling all the statistical information on a uniform system, coded at Cincinnati, New York, and London (or Oxford), with all the possibilities of sorting it out in various ways and for various purposes, a tremendous amount of time can be saved.

I enclose a sample of a standard code card which might fit the purpose; or special cards might be printed, of the same format but with a different group subdivision, adapted to the Minoan-Helladic requirements. As you can see, the enclosed card takes 80

(over)

characters arranged in 20 groups (each of 4); thus the code possibilities would run from 0001 to 9999, on a numerical system; but you would hardly need to know 9999 facts about your inscriptions. So the cards might be printed with groups of 3, the numerical code running from 001 to 999, which would be practical. Since, however, you are dealing with a language, an alphabetic code might be more sympathetic, though in this case 456,976 combinations on groups of 4, or 17,576 combinations on groups of 3, would both seem needlessly expansive. I would, therefore, suggest an alphabetic code in groups of 2 (AA-AZ, BA to BZ, etc.) which would permit 676 types of information to be coded, probably all that you would need. And it would have the advantage that it could be easily kept in mind, being composed of 26 series of 26 items, all in alphabetic sequences.

The alphabetic 2-letter code would seem more satisfactory also in view of the relation of the problem to the technical requirements of the machine. Presumably each card would represent one object (tablet, sealstone, or sherd); and it would probably be desirable to use, say, four different colors for pictographs, Cretan linear A, linear B, and mainland. Additional colors would permit the inclusion of Cypriote, etc. It would evidently be necessary to reserve the first 6 spaces for the catalogue or list number of the object, e. g., P317 (Pylos 317), F23 (Phaestos 23), or K1279 (Knossos 1279). Since the cards are limited to 80 spaces, the reservation of 6 spaces for the catalogue number would leave space for 37 items (2-letter code) regarding any object; therefore you might need to run over to a second or third card, so that they would be numberd, say K1279A, K1279B, and K1279C, using at a maximum all 6 reserved spaces. Three cards, for instance, could thus record 111 facts about one object. This would be an additional reason for using the 2-letter code, since, with groups of 3 or 4 numbers, there would be space for only 24 or 18 items, respectively, after reserving the space for the catalogue number.

The numerical code, if used, consists merely in punching single digits on the key punch machine, presumably in groups of three, e.g., 007 or 741. The alphabetic code consists in two punches for each letter, e.g., Y1=A, Y9=I, X1=J, X9=R, 02=S, 09=Z; if you touch Q on the typewriter, then the machine automatically makes two superposed punches, X (at the top of the card) and 8. Thus QG would actually appear on the card as X8-Y7.

In other words, the process would consist in typing in code the contents of each object, on one or more cards. The code might be numerical or alphabetic, but, for reasons of simplicity and adaptability as cited above, the alphabetic 2-letter code permitting 676 possible items would seem to be the best. The construction of a code would be up to the specialists, and should include everything that you wish to describe, everything that you wish to ascertain through statistical method, such as individual symbols for all the characters, for determinants, for numbers, initial and final letters of words, etc. This is all very rough, but may suggest ideas.

There might perhaps be 2500 cards, at the rate of one per object, and so actually many more. There could be as many duplicate sets as desired, duplicate cards being struck off on the reproducing punch, and sets deposited at Cincinnati, Columbia, Brooklyn, Oxford, etc. etc.

Printed lists of all the information distributed over the cards can be reshuffled and printed in a single line for each card, by another machine, at the rate of 80 lines per minute, with a minimum of effort; and such printed lists can be struck off again and again in accordance with different problems and theories of arrangement, so that the entire contents of all existing objects can be compared, on sheets of paper instead of examining the cards themselves. And of course a collating machine will pick out automatically all the cards with one or another characteristic, at the rate of 20,000 cards per hour. This would seem to open the possibility not only to the rapid tabulation of facts which are apparently now being done tediously by hand, but also to the rapid comparison of all the characteristics of linear A with linear B, and of the pictograph with both, with untold possibilities.

Miss Henle is about to experiment with coding and recording a few Pylos tablets and a few Knossos tablets, in the IBM laboratory. But the formulation of a standard code would depend upon a collaborative effort, and then, if it is to be of any practical use, upon the compilation and interchange of material, so that as many investigators as might be desired, scattered from Cincinnati to Oxford, might simultaneously concentrate on various aspects of the problem.

I should like to make a preliminary statement to those interested in the support of such a project. If you feel that you have any ideas regarding it, I should be most happy to receive them.

Very sincerely yours,

William B. Dinsmoor (signed)

WBD:D

(over)

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK
Department of Fine Arts and Archaeology

February 3, 1950

Professor Carl W. Blegen
University of Cincinnati
Cincinnati, Ohio

Dear Carl:

I trust that you received my preliminary letter of January 31, regarding the utilization of IBM facilities as preparation for a wholesale onslaught upon the Minoan code. Pending your answer, and to give you more basis for judgment, I may add that since writing to you Miss Henle has been receiving some training in the IBM laboratory and that with their staff she has worked out the principles of a more practical code, numerical (4-figure), and on the scheme of giving on each card the place and catalogue number of the object (tablet), obverse or reverse, line, position of word in line (all these for identification purposes), and then the word itself spelled out. In short, the project can have more practical use, evidently, if each card has only one word. This may amount to 100,000 cards. But quantity is almost immaterial in such a process, since the collating machine can sort all 100,000 cards in five hours, and those desired for any purpose printed off in a few minutes. The experts prefer a numerical code on the ground that the collating machines capable of utilizing this are more widely available than those adapted to alphabetic codes. The first experimental tablets have been coded, and they look like monthly salary statements; Miss Henle can read them at sight, but she stutters when she comes to the actual Minoan word and has to resort to a pencil.

Sincerely yours,

William B. Dinsmoor (signed)

WBD:D

C
O
P
Y

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK
Department of Fine Arts and Archaeology

February 6, 1950

Professor Carl W. Blegen
University of Cincinnati
Cincinnati, Ohio

Dear Carl:

Bulletin no. III in re IBM. Please correct one figure in second letter (Feb. 3); hastily glancing at the printed IBM cards used for the experiment, divided into 4-figure groups, I wrote that the more practical code proved to be "numerical (4-figure)" --but I meant "numerical (3-figure)." That is, the experimental code runs from 001 to 999.

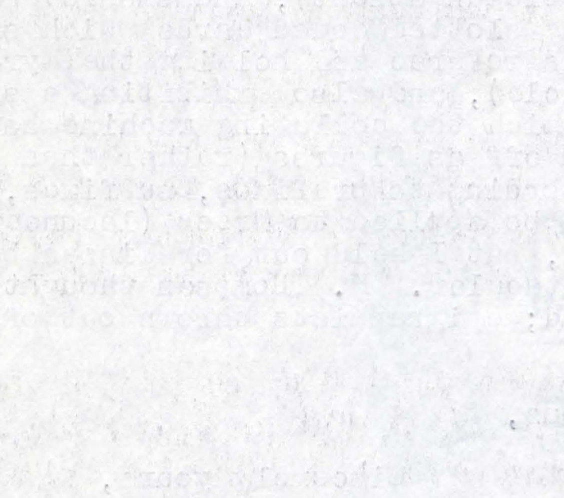
I asked Miss Henle to come to the Archaeological Club meeting yesterday and to report on her three-day training at the IBM laboratory. The usual group was present, including Rodney Young (but unfortunately C. W. Blegen absent), and she gave a very good account of herself, showing slotted coded cards which she read off at sight (Ingholt serving as referee and holding the Myres table of signs from Miss Kober's article), and also exhibiting a sheet of paper about 8 1/2 feet long on which the collating machine had shuffled the cards and printed them off as figures (rather than slots), grouping them automatically according to prefixes, suffixes, etc. Raubitschek thought it ought to be applied to Greek (lacunae in inscriptions), Nock to Etruscan, but I held out for Minoan because the backers want something spectacular. H. Thompson thought that the social implications were bad; epigraphists thrown out of their jobs by button-pushers.

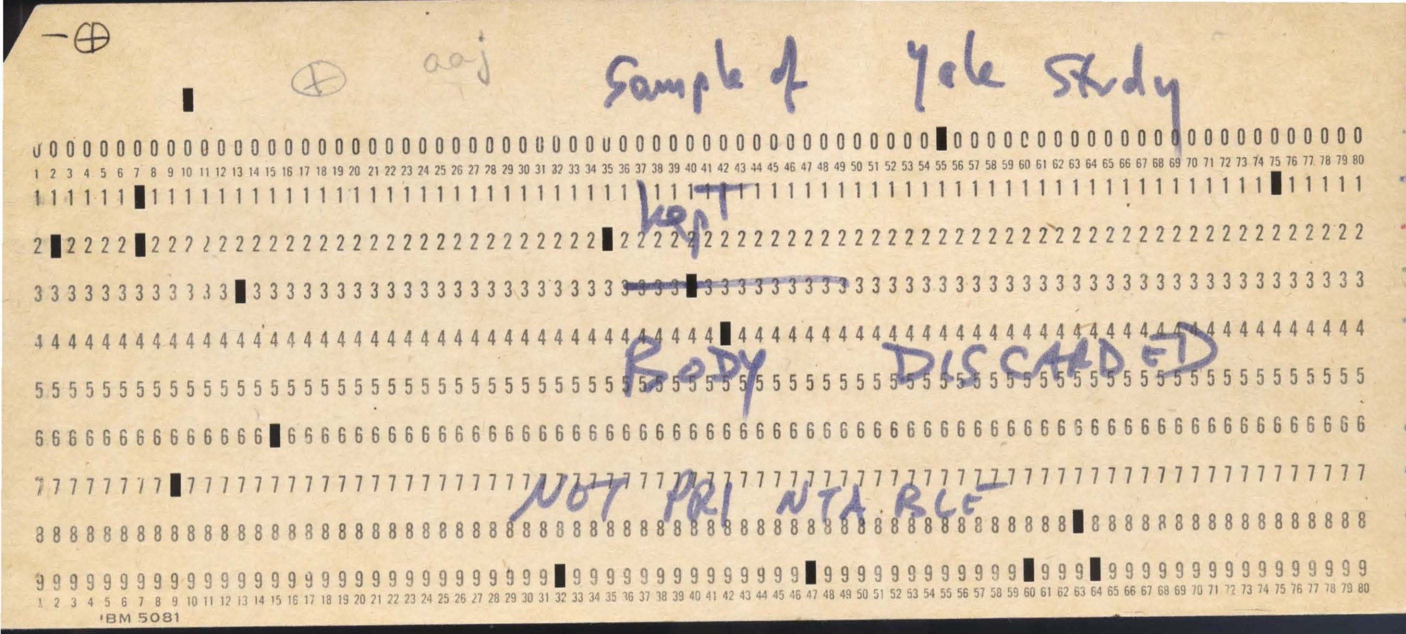
Awaiting your reactions,

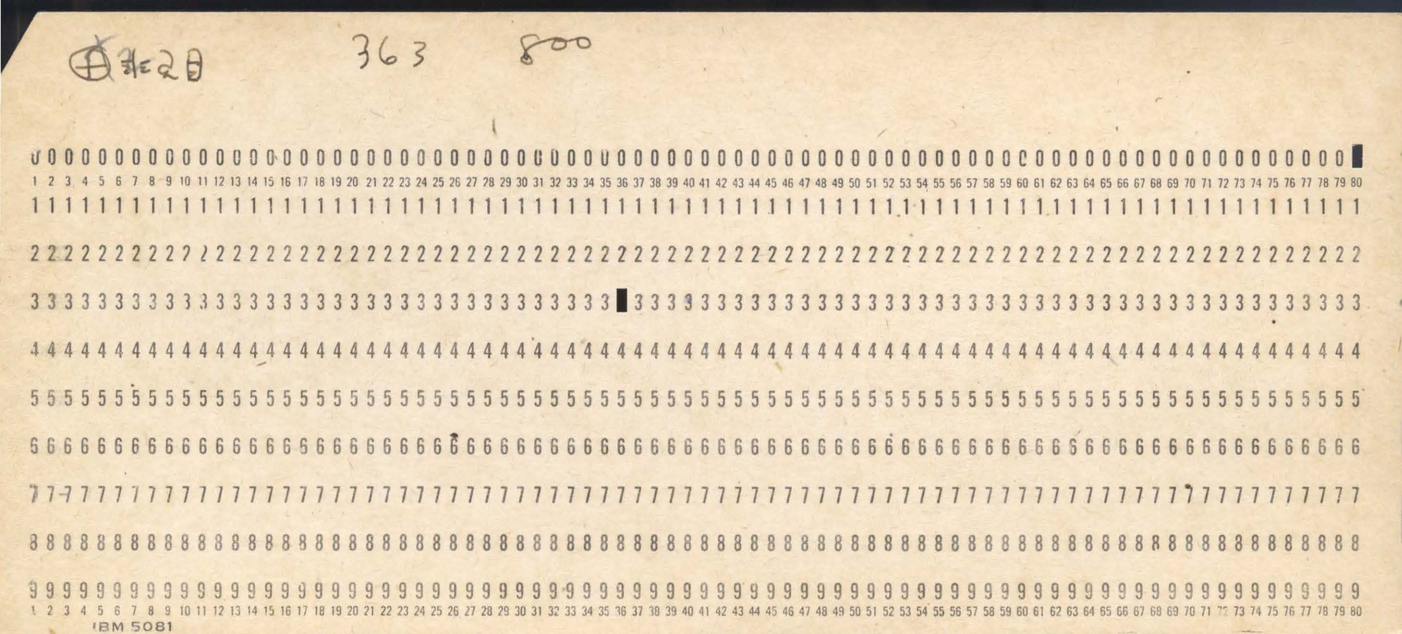
Sincerely yours,

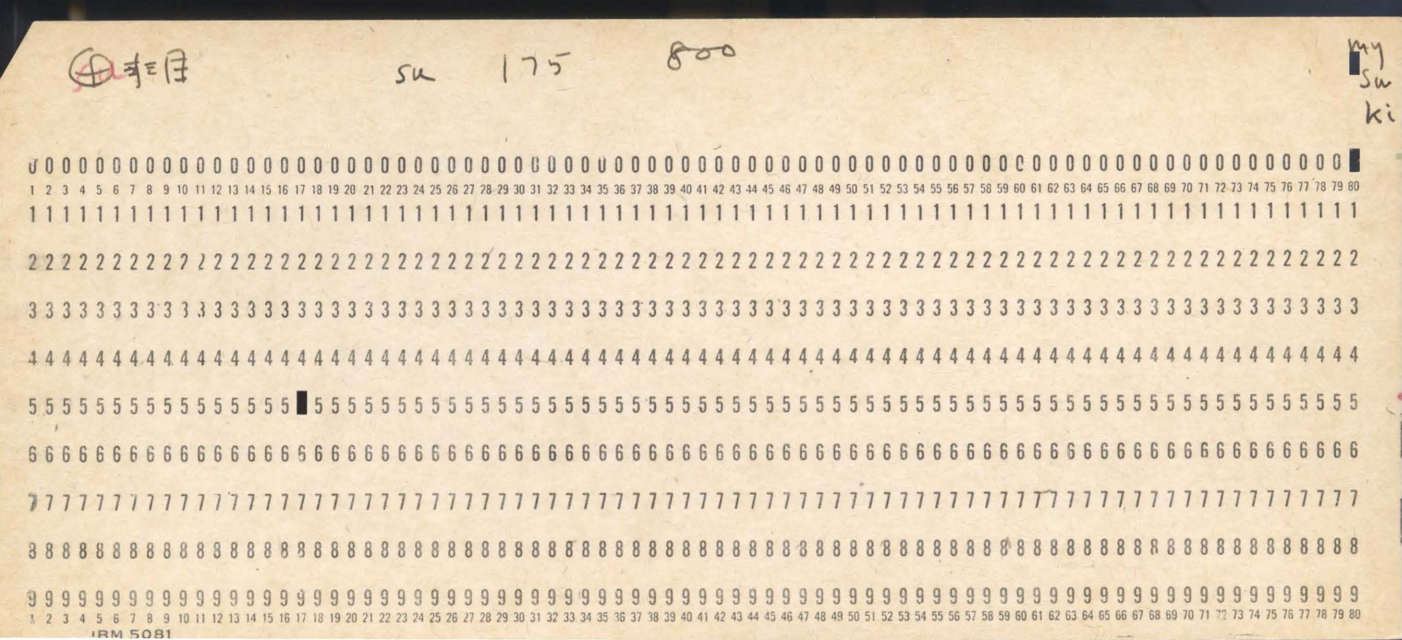
William B. Dinsmoor (signed)

WBD:D


A I J R S Z
Y
X
0
1
2
9

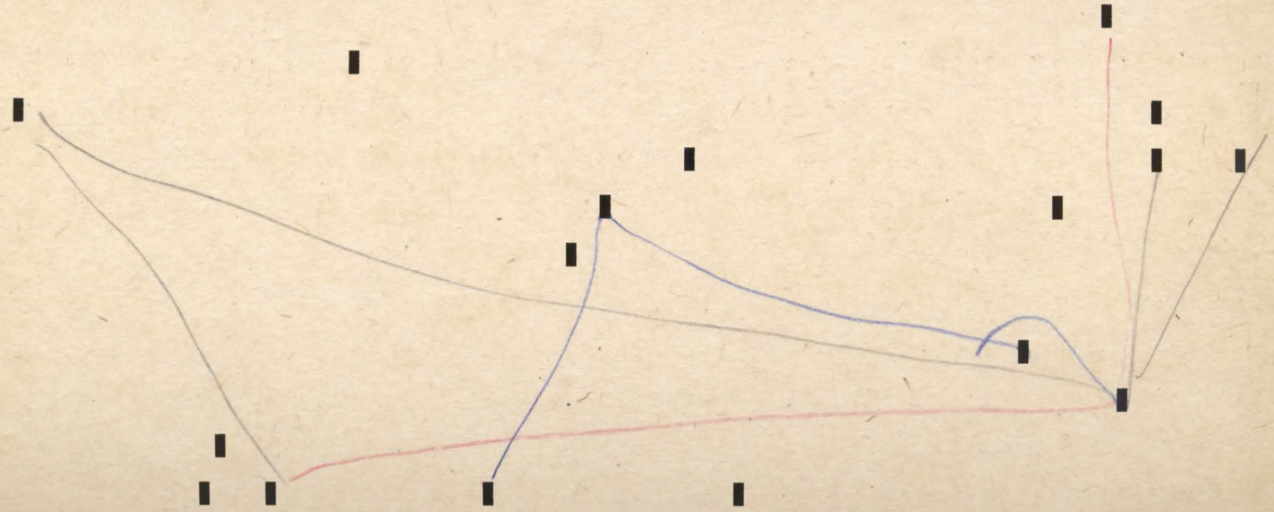

aaj
Sample of Yale Study
kept
BODY DISCARDED
NOT PRINTABLE
IBM 5081


363 800
IBM 5081


su 175 800
my
sw
ki
IBM 5081

17

2



⊕ 下
ev

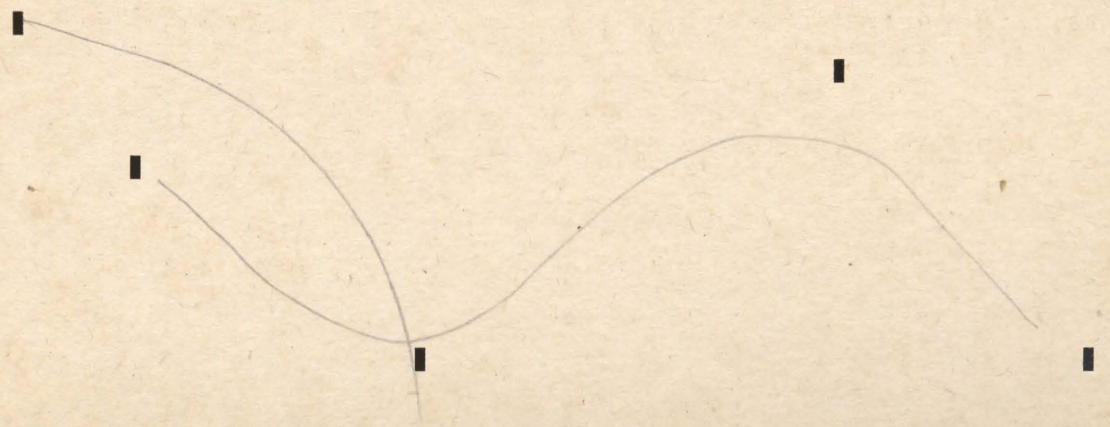
IBM 5081

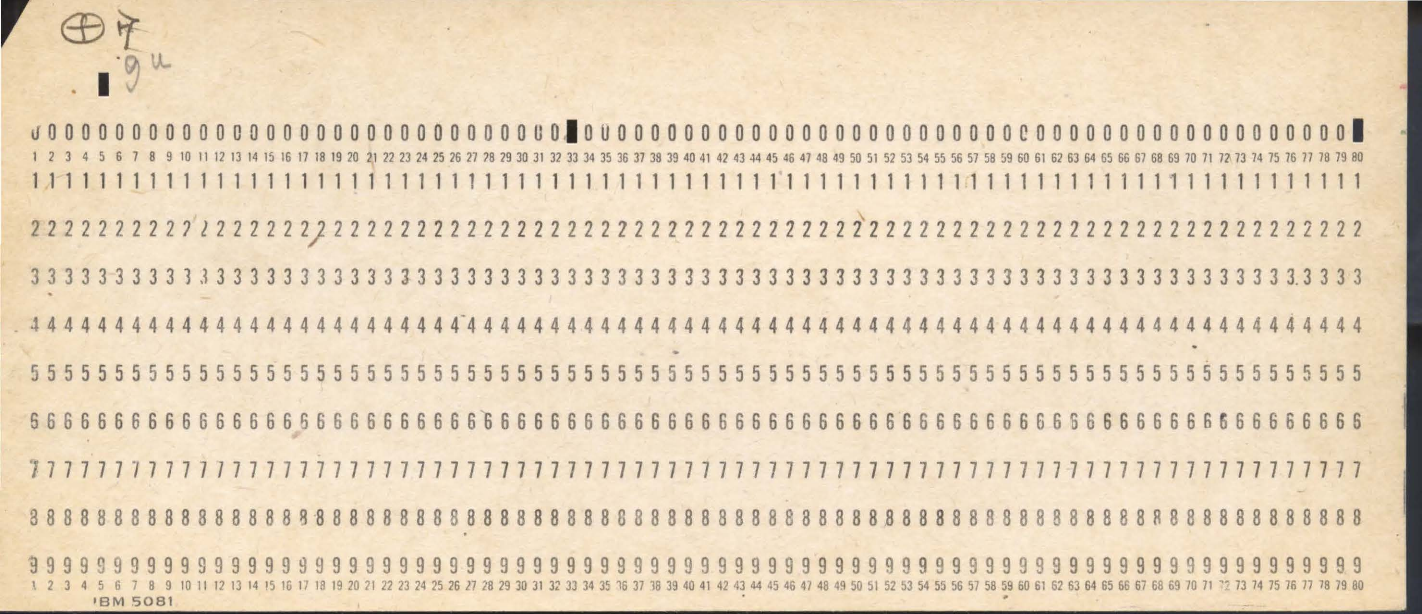
– ⊕ 下
ev

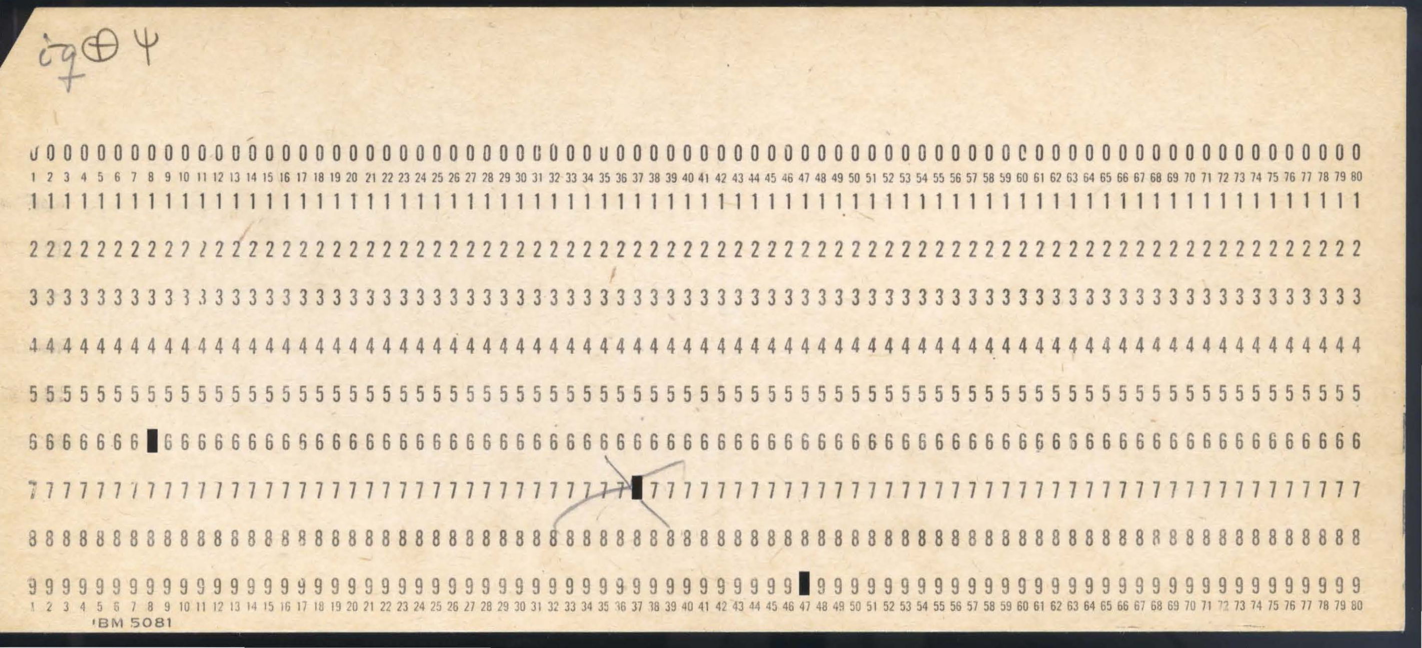
IBM 5081

⊕ ち    gd 751 800    gd ty

IBM 5081




IBM 5081


IBM 5081