Potsherds Into Printouts

The Ban Chiang Computer Project

JOHN HASTINGS

As a new recruit to the Ban Chiang lab I was astonished at the huge quantity and variety of material being studied and the number of people working on it. The awesome amount of information being amassed, I was told, was being 'put in a computer,' and I was assigned to an experienced volunteer who showed me how to code artifacts. I never imagined that the next three years would involve me with nearly every aspect of the computer project.

The work of coding data for the computer was complicated and difficult at first, and when I began to understand it better it seemed downright crazy. Mostly it was numbers. Condition of Artifact, for example, was coded '3' for 'whole and intact,' '4' for 'complete but reconstructed,' and so on to '7' for 'fragmentary.' The computer wouldn't take anything for granted and there were about two dozen items to be coded for each object: where it was found, what material it was made of, its measurements, whether it was polished or faceted or scratched or worn. Even colors became numbers by comparing them with standard color chips in a Munsell color book.

The coding routine we had to follow was rigorous and tedious, but I came to realize later on that in fact this was one of the reasons for using a computer: it imposed a consistent and systematic structure for data being recorded by many different individuals, in our case volunteers and students with diverse backgrounds. Detailed instructions prescribed precisely what information was required for each class of artifact, and if a coder overlooked a category the computer would report the omission so it could be added later.

The supervisors were forever having us check each other's work. The coding sheets were checked, the punched cards were checked, even the final printouts of artifact descriptions were checked against the actual objects. I thought this was overdone until I tried checking some of my own work. No matter how careful I had been there were always some mistakes to be corrected. By the time we were finished checking, the records were considerably more accurate and complete than they would have been with a less thorough manual system.

Except for the coding I didn't know much about the computer system until one day Chet Gorman greeted me with the preposterous suggestion that I should take over running the computer. The graduate student who had been doing it was leaving and there really wasn't anybody else. I plunged in, working mostly by trial and error, studying incomprehensible technical manuals, asking questions of anyone I

1
Co-director Pisit Charoenwongsa studies computer reports with the author. The Thai Fine Arts Department has available in Bangkok a computer similar to the one being used by The University Museum, on which it will later be able to run the Ban Chiang data and program tapes.

could find at the computer center, and after a few months it somehow began to make sense.

At this point Chet announced that he wanted me to teach him to use the computer, casually ignoring my protests that I only half understood the thing myself. We decided that he would undertake to ask the computer, which contained data on all the objects from Ban Chiang, to print a list of only those which had been found in burials. After a blackboard session to explain the instructions to the computer which would be needed, I took him to a keypunch machine in the DRL building. Chet wanted to do everything himself, with his own hands, so he sat down and with only a little help managed to punch the instructions into a set of IBM cards. Finally we went to the computer, fed in the cards, and waited as the moment of truth approached. We soon received a printout saying the run had failed; Chet hadn't made any errors, but I had given him the wrong tape number. Ignominy! Fixing the mistake, we resubmitted the cards, and to our mutual amazement everything worked. Chet had his list of burial objects.

Early in the project the contents of the field bag logbooks had been fed into the computer, so it knew exactly where, both horizontally and vertically, each of the thousands of bags of sherds had come from. The wonderful machine then listed all the bags from the vicinity of each burial to help start reconstruction of the pots found in them. While the coders had been working on the small finds, other tireless volunteers had been reconstructing 300-odd pots.

When the time came to describe these pots for the computer I learned one more aspect of the project—creating the coding instructions. We needed to code dozens of measurements and a description of the shape (how can you describe shape accurately in numbers, or even words?) along with some account of the surface decoration. Most important, I had to learn the kinds of archaeologically meaningful information that would be needed for studying the pot's cultural significance: character of the clay and what inclusions might be in it, evidence of the manufacturing methods used, the context in which the pot was found. There seemed no end to what might be recorded. Some of our ideas had to be dropped because the data would be too imprecise to be of value. Others had to be simplified: painted or incised designs came to be coded merely as Curvilinear, Geometric, or Lines. The scheme I devised for coding the shape could give a fair idea of the form of the pot, but numbers are a poor substitute for seeing the pot itself. Fortunately we would not have to depend on the computer as our only reservoir of information. Every pot had been carefully photographed, and many of them also drawn to show clearly the designs on them, and these visual records would nicely complement the measurements and detailed observations we would put into the data bank.

DATE: 11/17/81 BAN CHIANG SMALL FINDS PAGE NO: 237

CATEGORY

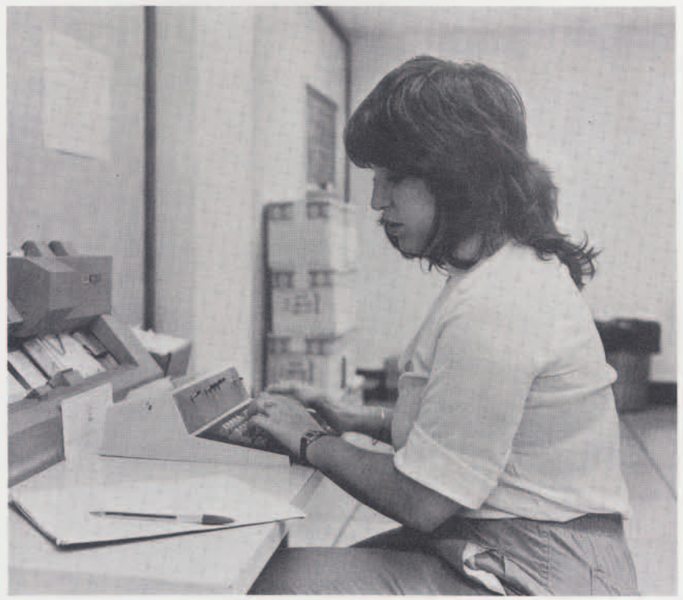
	SERIAL:	00163130
001	SITE NAME:	BAN CHIANG BCES
008	SMALL FIND NUMBER:	0762
010	BAG NUMBER:	2834
011	MATERIAL:	BRONZE
012	SMALL FIND CLASS:	BRONZE ARTIFACT
013	CONDITION OF ARTIFACT:	INTACT
020	SQUARE:	D6 / D7 BK.
030	LAYER NUMBER:	27 S.
050	BURIAL DESIGNATION:	076
055	ARCHAEOL. ASSOC./LOCATION/POSITION:	BELOW BURIAL, RIGHT SIDE, CRANIUM
080	RECORDER/ANALYST:	JVH
102	NUMBER OF PIECES OR FRAGMENTS:	001
170	WEIGHT (GM):	0091.0
310	MAXIMUM DIMENSION (CM):	13.6
315	DEGREE OF CORROSION:	AT LEAST SOME METAL PRESENT
316	CONSERVATION:	CONSERVED
317	PREPARATION FOR LABORATORY ANALYSIS:	SAMPLED, MOUNTED OR PREPARED
319	OBSERVATIONS:	NONE
321	CONSERVATION LAB NUMBER:	2011
520	TENTATIVE SUBCLASS:	SOCKETTED POINT
521	ORIGINAL LENGTH OVERALL (CM):	15.5
524	ORIGINAL LENGTH OF BLADE (CM):	10.0
540	MAXIMUM WIDTH OF BLADE (CM):	03.9
543	MAXIMUM THICKNESS OF BLADE (CM):	00.3
544	MIDRIB:	PRESENT
545	SHAPE OF BLADE:	SEE DIAGRAM NO.5
550	OUTSIDE DIAMETER OF SOCKET (CM):	02.3
551	INSIDE DIAMETER OF SOCKET (CM):	01.9
569	COMMENTS ON METAL ARTIFACT:	SOCKET IS ELIPTICAL IN CROSS SECTION 2.3 X 1.7 CM. MIDRIB IS CONTINUATION OF SOCKET AND IS UNIFORMLY TAPERED FROM THE SOCKET HOLE TO THE VERY TIP OF THE BLADE, AND IS THE SAME ON BOTH FACES OF THE BLADE. BURIAL IS FLEXED AND ON RIGHT SIDE.

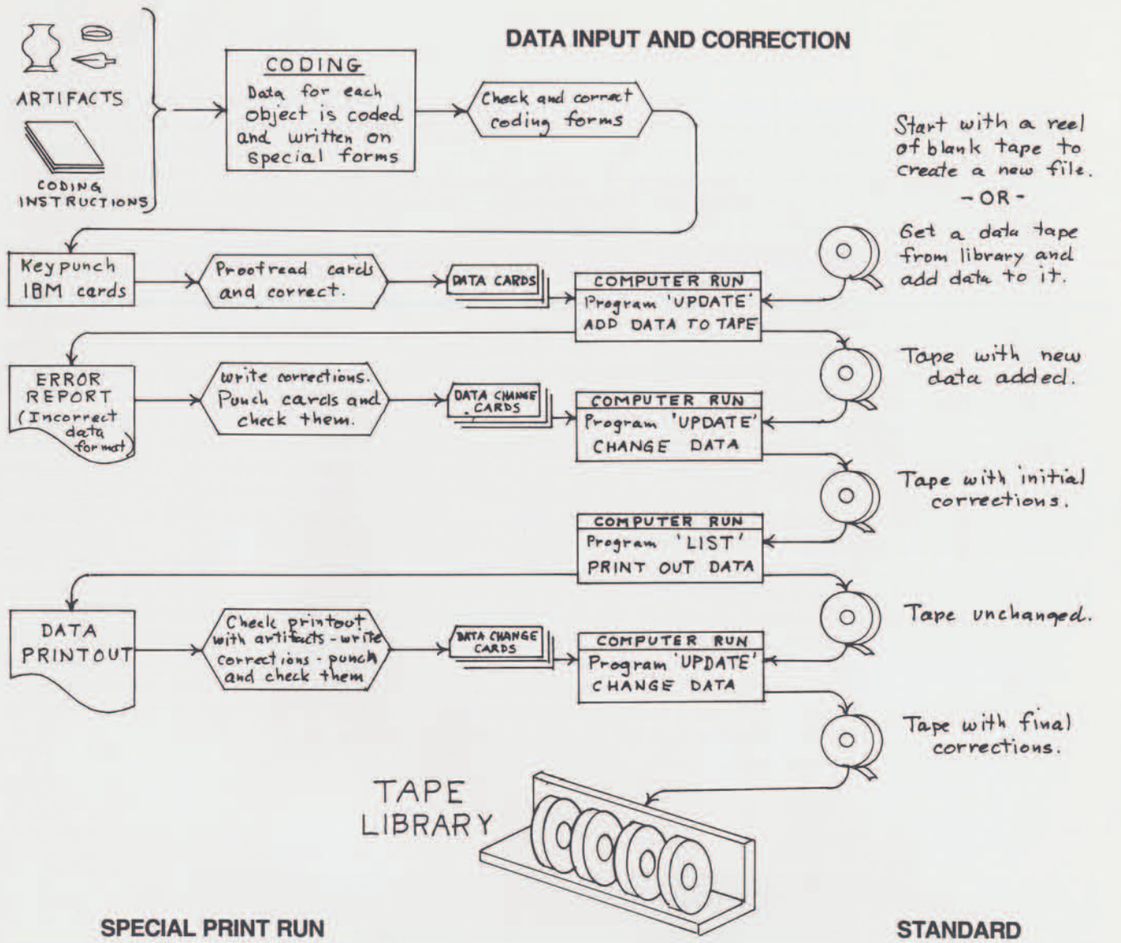
The pottery data is now in the computer, and adding to the data bank will continue. Next will be information on the burial skeletons: age, sex, orientation, and so on. But the data in the computer is now complete enough to be useful. A printout of the entire file constitutes a convenient, easy to use catalogue of the Ban Chiang material. Duplicate printouts can make the data available at the Thai Department of Fine Arts in Bangkok or elsewhere. A printout of the metal artifacts is compact and handy for studying that class of object, and preparation of the Ban Chiang Travelling Exhibition has been facilitated by a special listing of the exhibit objects. A printed index of artifacts arranged by type quickly reveals, for example, how many clay rollers were found and where.

Chet Gorman intended the Ban Chiang data bank to be a model for computerizing Southeast Asia archaeological data. His vision was of a computer with many sites coded in the same system, with which one could search for patterns and similarities and perform statistical analyses on the entire area.

2
Example of printout for one artifact. The computer tapes hold data on 4919 such 'small finds,' 341 pots, and 6289 bags.

3
Data from the coding sheets is punched into IBM cards on a keypunch machine by Cheryl Applebaum.





HOW THE COMPUTER WORKS

The Ban Chiang Project uses a system of computer programs called SELGEM (Self Generating Master), developed by the Smithsonian Institution for cataloguing information about large collections of objects. When cataloguing artifacts, various kinds of data are entered for each: class of object; provenience (square, quadrant, layer, bag number); description and measurements; perhaps even comments. Each of these kinds of data is given a "category number" which simplifies coding and makes it easy to find data in the computer. For instance, kind of material is entered as category 011, so to make a list of bronze objects the computer is instructed merely to check category 011 of each artifact and print out those which say "bronze."

Categories have been established to permit describing the various kinds of Ban Chiang artifacts. Many are rigidly specified (maximum dimension is always entered as centimeters with 2 digits, decimal, digit; 4 cm. must be entered 04.0) to permit accurate searching and statistical analysis, while others are flexible, even allowing free-text comment of unlimited length. Data can be added to the computer at any time, not only to enter additional artifacts but also to add new categories of data. Thus the data bank can be built progressively as information becomes available or new needs become apparent.

While Ban Chiang SELGEM is a large data base running on a mainframe IBM 370 computer, it is surprisingly economical. The data is stored on magnetic tapes at a cost of five cents a day for 4,000,000 characters, and computing is done not through an on-line terminal but by low-cost 'batch processing,' usually at night. The computer can select artifacts with any specified property (such as everything from square D4), sort them into any desired order (such as by layer), and print the result in various ways. Reports with custom-designed formats are possible, and coded data can be translated so that it is printed in meaningful words. For statistical analysis SELGEM can transfer its data to the statistical 'program packages' which are usually available on large computers.

4
Flowchart of typical computer input and output. Much of the data entry procedure is concerned with finding and eliminating errors to make the final data as reliable as possible. Data are stored on magnetic tape and can be selected or reorganized by the computer for printed lists and reports.

5
Once the information is in the computer, printouts can be produced and the work of data analysis can begin. Here the author is preparing a computer run to count the number and weight of sherds found at each layer of the excavation.

Bibliography

Schauffler, William
1979
"Computerized Data Base Management Systems in Archaeological Research: A SELGEM Case Study," *MASCA Journal* 1[2]: 50-55.

Credit
All photographs by David Gladstone.

John Hastings is currently a Research Associate at The University Museum and has been in charge of the Ban Chiang computer project for the past three years. An engineer, whose company manufactured gold leaf and did vacuum metallizing, he has been a longtime member of the Museum and on his retirement became associated as a volunteer with the Ban Chiang Project.