DOCUMENT RESUME

ED 042 203 CG 005 739

AUTHOR	Lambert, Nadine M.; And Others
TITLE	Anecdotal Processing to Promote the Learning Experience. Stress of School Project Report.
INSTITUTION	California Univ., Berkeley.
SPONS AGENCY	Public Health Service (DHEW), Arlington, Va.
PUB DATE	Feb 70
NOTE	41p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$2.15
DESCRIPTORS	Behavior, *Computer Oriented Programs, *Data Processing, Educational Improvement, *Individual Characteristics, *Information Systems, *Learning, Research Tools, School Organization, Systems Development

ABSTRACT

The Stress of School Project attempted to find new ways by which the school can help children learn more effectively. The resulting information system, APPLE (Anecdotal Processing to Promote the Learning Experience) was designed to process anecdotal data obtained in schools and integrate it with quantitative measures of pupil characteristics. Many features of APPLE are oriented toward building and maintaining magnetic tape and disc files of a variety of types of educational data. The system is also generalized so that other school systems may use it. The basic features of the APPLE system, the APPLE storage and retrieval program and the APPLE supporting programs are specified. Types of behaviors incorporated in the system include: (1) academic-intellectual; (2) social-emotional; (3) physical appearance; (4) general; and (5) other. The APPLE data structure and the way data are organized into files give the system its potential as a research tool and its generalizability to a variety of administrative needs. The files are discussed quite thoroughly. (KJ)

The Stress Of School Project

Anecdotal Processing to Promote the Learning Experience

A report of the Stress of School Project
University of California, Berkeley

Nadine M. Lambert, Project Director
Carolyn S. Hartsough, Assistant Project Director
Fred L. Converse, Systems Programmer
Emily K. Converse, Programmer

This research was supported by USPHS,
NIH Grants MH 14,605-01 and 02.

ANECDOTAL PROCESSING TO PROMOTE THE LEARNING EXPERIENCE

APPLE

February, 1970

This research was supported by USPHS,

NIH Grants MH 14,605-01 and 02

ANECDOTAL PROCESSING TO PROMOTE THE LEARNING EXPERIENCE[1]

Nadine M. Lambert, Carolyn S. Hartsough

Fred L. Converse, Emily K. Converse

University of California, Berkeley

Background

Most of us on the educational scene are aware of the tremendous amount of available or readily obtainable information with which to plan learning experiences for children. However, it is remarkable that in the Moon Age, a scarce few school districts employ even the most rudimentary data processing systems for collecting and reporting data or for conducting systematic studies of the outcomes of school programs. Moreover, the data which districts have available in some machine sensible form often are not integrated into a single record-keeping system so that analysis of the data across sets of records can be accomplished readily.

Large or wealthy districts have the personnel and budget resources to allocate to the development of district-wide computer record-keeping and reporting. The smaller, poor districts have none of these resources and they either continue to handle school record-keeping by hand or utilize computer service centers for unit jobs such as test scoring and reporting, grading, and attendance accounting. Meanwhile the information necessary for planning and evaluating children's learning experience is not available to the teacher and knowledge about the child in school during one year is usually lost by the following year. Until now there has been no available method for recording and processing information about the experiential aspects of the child's transactions with the school nor for matching this information with administratively gathered data.

Many attempts have been made to develop data processing and information retrieval programs for educational uses which have the potential of general applicability of school systems (Goodlad, Caffrey, O'Toole, & Tyler, 1965). The failure of schools to adopt general programs of the type that have been developed and tested in extensive field trials is a subject worthy of investigation in its own right. But a cursory review of the problem of adapting any set of data processing systems to individual school requirements for record-keeping, reporting, and evaluation leads to the conclusion that the problems of adaptability lie in the constraints which are placed on the school system user, and in the lack of generalizability of the specific retrieval programs to individual school district's needs. For example, data processing programs which require the school district to report scores of a particular type, for specific tests, and in a specific format leave no freedom for the school district to select the type of standard scores they prefer, the number of tests to be reported, and, of course, the form in which the tests are to be summarized. Futhermore many data processing programs developed by individual school districts or state agencies are closed systems specifying the types of data which can be handled and making no provisions for either adding to or modifying the data input.

In the present state of technology many system analysts assert that general programs cannot be general enough to have relatively universal application. Therefore, they recommend that the goal of general purpose information systems should be abandoned. Those who dare step where angels fear to tread and suggest the feasibility of developing a set of general data processing and information retrieval programs cannot ignore this problematic past in proposing ideas for future development.

Several initial considerations are basic to the development of generalized programs. One stumbling block to building an integrated record-keeping system is the problem of establishing a dossier code which not only will follow the child through school, but which also can be used to alphabetize the many rosters which schools use to report information to staff, such as the roster of class members for whom test scores are being reported. Storage of records is still another hurdle, but with the increasing availability of disc storage, and remote access terminals, this problem is one which appears fairly easily resolved. Adaptability of the programs to the principal computer hardware systems such as IBM, Honeywell, Burrough's, RCA, or Control Data and the inter-face of the retrieval programs with existing computer software are primary considerations in the creation of general information processing and retrieval programs with maximum adaptability. And this list would not be complete without pointing out the need to maintain different levels of confidentiality of records so that the privacy of the child and the family can be ensured as data are reported and summarized for various educational uses.

Requirements for Development of the APPLE System

The Stress of School Project attempted to find new ways by which the school can help children learn more effectively. Among the objectives of this project was the study of teacher strategies for meeting the needs of children risking failure in school. The attainment of the project objectives required the development of tools for integrating the processing of quantitative and qualitative information in a pupil personnel data base. The resulting information system, APPLE (Anecdotal Processing to Promote the Learning Experience), was designed to process anecdotal data obtained

in schools and integrate it with quantitative measures of pupil characteristics and behavior.

Generalizability of the APPLE System

Early in the development of the system for the Stress of School Project it became apparent that the evolving computer programs should not become specific to the project goals, but should become a comprehensive system for handling pupil personnel data which would be readily applicable to a wide variety of school systems and their individual information processing requirements. The effort to make the programs general rather than specific in no way prevented the project research staff from attaining the project goals; rather this requirement of generalizability enabled us to maintain an open system of record-keeping into which could be incorporated any unexpected source of information about pupils which had not been contemplated earlier.

Types of Pupil Personnel Data to be Processed

Many features of APPLE are oriented toward building and maintaining magnetic tape and disc files of the following types of educational data.

1. Data for monitoring learning and for decisions relevant to modifying the learning environment of the individual to adapt to his readiness. APPLE should be of greatest value to education in this area where rapid entry of measures of the pupil's progress, digestion (elimination of "noise" and detection of significant parameters), and presentation of easily comprehended diagnostic data displays permit prompt adaptation of the setting within which his learning should advance.

2. Data of traditional educational administration such as those legally required for state reports (identification, admission, attendance, withdrawal, grade status, immunization, guardianship, addresses) and/or

for protection of health, safety, and welfare of the child (telephone numbers, places of employment of responsible adults, reports of examinations of hearing, sight, medical, and dental examinations), and/or for guidance and reporting to parents and community (test results, achievement, and citizenship marks). The connectives "and/or" above are included to remind us that much of these data are processed to meet multiple needs. Recording it, storing it, integrating it in various combinations, reporting it, checking it, correcting it . . . are expensive tasks. If the system is adopted by a school district, APPLE should have its economic justification in this area alone.

3. Additional data gathered for program evaluation and for research to improve the efficacy of the educational process. APPLE's cost of operation must be low to be administratively attractive for application to data in area 2 above, however, the initial costs of developing a system with research potential as well as information retrieval capabilities have been borne by the Stress of School Project. The APPLE system provides for data retrieval and data reduction as a by-product of administrative data processing by educational agencies.

Problems of Diversity of Data

The wide variety of types of data processed by each district and the common practice in most districts of using fragmented procedures for analysis and reporting greatly handicap the districts in using the data in integrated ways or in communicating it to other districts as the pupil moves. Rarely, for example, are test scores of pupils tested in one year readily available in subsequent years for comparison with later test data, report card data, or attendance accounting data.

Three separate grade level cohorts of subjects from each of four school districts in the East Bay Region of the San Francisco Metropolitan Area constituted the pupil sample for the Stress at School Project. Subject data collected from existing school records, classroom observations, teacher consultations, parent interviews, and project assessments comprise the epidemiological data by which school-based stress in children and relevant teacher strategies was to be studied. For APPLE, however, the characteristics of the record-keeping, data collection, and data processing systems of the school districts are of more direct interest. These four districts represent the wide range of procedures which are typical of the administrative record-keeping activities of California school districts.

Test data. For several years one of the four districts had maintained computerized test scoring and reporting services, and could integrate these data with grade and attendance information. Another of the four districts contracted with a local school testing service for test scoring and reporting services, but did not integrate attendance and grade reporting with test data even though a unique pupil identification number was assigned to every pupil in the district. The other two districts made minimal use of data processing systems. Except for reports of achievement and ability test scores required by the State Board of Education, all other elementary school data were processed by hand and there was no attempt to integrate information for an individual pupil other than what was available from teachers' reports and other data on each child's cumulative record.

Report card data. In the four participating districts, the forms of report cards were as varied as the record-keeping systems. Each district utilized different achievement and behavior classifications by which to evaluate children's behavior at the time of regularly scheduled report periods. Grades were recorded by hand and the record was maintained usually as a series of separate forms in the child's file. Parent conferences replaced grades in some districts, making an ongoing account of pupil progress impossible without provisions for processing anecdotal data. Often test score reports were transferred by hand from master lists onto cumulative records; and in others, pressure sensitive labels provided some guarantee that the information was actually entered into the pupil's school file.

Attendance data. Quite diverse practices characterized the attendance keeping system of the school districts as well. In two cases the daily attendance was entered into a pupil's card by the school secretary and summarized by pupil, grade, and school monthly. In other cases, the monthly record was kept on an IBM mark sense card which was prepunched with the pupil's name and ID number. But since attendance accounting reports were kept primarily for the computation of the district's average daily attendance (ADA) from which is computed the State school apportionment for each district, these data were kept separately and not made a part of the pupil's cumulative file. Attendance reports were not merged with other pupil personnel information. We therefore encountered considerable difficulty in coping with these record-keeping procedures so as to employ the data for the study of the causes or the effects of learning and behavior problems of children.

Teacher observations. In none of the participating districts was there any systematic record of teachers' observations of children, their comments about children's progress during the year or other qualitative information which might be of assistance to future teachers. While some useful information of this type might be found on some of the cumulative records of a single school, the records varied widely.

This diversity of form, content, and accessibility of the school records of the subject pupils in the participating districts provided an excellent sample of problems in information control, management and processing by which to develop and test the APPLE system as a set of general pupil personnel record-keeping programs. The variation in record-keeping practice among the districts required immediate adaptation of the APPLE formats to the data so that they could incorporate the diversity of information items within a data species. In order to utilize existing school district data on the project pupils we had to plan to provide for, develop, and test programs which could translate existing data formats into APPLE formats so as to eliminate the necessity of a great deal of custom programming when the system moves from the developmental to the operational phase and is employed in other school districts. And finally the lack of accessibility and the lack of complete data on the subject pupils amplified the need for a systematic economical method for processing pupil data in order to capitalize on ordinarily available information for the planning of individual and classroom learning experiences.

Features of the APPLE System and Programs

Thus the requirement of generalizability as well as the particular types of data to be processed have dictated the major features of the

APPLE system. Although largely oriented toward processing pupil personnel and program evaluation data, APPLE can be modified readily to handle "human" data in areas where anecdotal annotation is important as in medical, psychiatric, correctional, vocational or welfare areas. The basic features of the APPLE system, the APPLE storage and retireval program, and the APPLE supporting programs are specified below.

The APPLE System

1. The system is economically feasible as a basic file maintenance system for a school district.

2. The programs are in compatible modular design for user convenience and for easy program modification. Programs process tape and disc files in multiples of 80 character records. Tests of compatibility of programs have been made on CDC, IBM, and Burroughs computer systems.

3. APPLE can process substantial quantities of anecdotal annotation (unlike most educational data processing systems). However, such data as attendance, test scores and grades are handled more efficiently in arrays of fixed fields (unlike most language processing systems).

The APPLE Storage-Retrieval Program

1. The APPLE main program manipulates diverse forms of input files stored in cards, tape, disc, or paper tape, or keyed directly from remote console. The inputs can be copied, abstracted, matched, merged or otherwise processed to produce simultaneously several varieties of output files, including "disguised" files to protect confidentiality.

2. The system provides readable labeled reports of pupil personnel information in many different formats as appropriate to the type of data being processed.

3. For special research or administrative purposes, APPLE analyzes files according to user's specifications, abstracting lists of dossier codes of those pupils satisfying and those not satisfying the specified requirements or abstracting selected information about those pupils whose dossier codes match a previously established list.

4. Profiles of pupil populations can be derived by the capability of the program to provide counts showing how pupils do or do not satisfy data specifications and to provide summary counts showing the same for pupil groups.

5. In order to interface with standard statistical packages and utility programs, APPLE abstracts and reassembles items from the input files and from summary records to produce customized output files.

6. APPLE accepts data in <u>any</u> card format in order to incorporate existing school district records and with user's specifications will convert the data to an APPLE format. To a limited extent it can accept such irregular formats intermixed with regular APPLE formats.

7. APPLE identifies and alphabetizes pupils in school populations by the assignment of permanent unique dossier codes under rules which permit initial assignment of codes to 100,000 pupils in each of 100 cohorts and permits subsequent expansion to up to ten times that number.

8. To facilitate addition, replacement, or deletion of individual lines, APPLE assigns sequence numbers to each line of the file.

<u>The Supporting Programs</u>

1. The "sort package" provides for copying selected (or all) APPLE images into larger disc records together with a fifty-character key synthesized from (a) items found in the image, (b) items "remembered" from

previous images, and (c) items converted by user supplied tables to permit "as if" sorting (e.g., as if "aggression" and "atn-getng beh" were both "3"). The disc records are then sorted, copied to tape in original 80-character format, and listed.

2. The remote access capability of the system is provided by "TTY retrieval" which lists on the teletypewriter all records of those pupils whose file numbers have been teletyped to the computer. Lists are provided of the file numbers scanned, showing those not sought, those sought and found, and those sought but not found.

Processing Natural Language Data in the APPLE System

Conceptual Basis for Processing Natural Language Data

The APPLE system owes a substantial debt to the work of Eiduson in the development of the Psychiatric Case History Event System (PsyCHES) (1966). From an empirical analysis of psychiatric case histories, she determined that the total information pertaining to a patient could be perceived as a series of events which took place in the life of the patient and his close family members. Events were conceived of as actual happenings as well as psychological phenomena such as fears, fantasies, etc., which could not be observed, but had to be inferred. The structure of an event was considered to be similar to an English sentence and was comprised of the words, numbers, and symbols which were typical of events of particular kinds. These parts of the information contained in the event came to designate the standard format fields which identified or defined an event in the PsyCHES program. PsyCHES retained the natural language in which the case history was reported, and required no checklists, ratings, or codes to reduce the natural language into a priori categories.

This organization of descriptive information relative to the history, interviews, observations, and medical and psychological data collected on the patients in the Hudson research provided us with the basis for conceptualizing the management of anecdotal information available on children in schools. Case study and guidance reports following individual psychological examinations are the principal sets of data which are analogous to psychiatric case history files. However, such information was available on less than 5% of the school subjects in the Stress of School Project. Classroom observations, teacher reports, principal contacts, parent reports, and a large variety of other incidental information comprised the typical sorts of anecdotal data which combine to form the "event" information of children in the school setting. Thus we conceived of an event as in the PsyCHES system--anything that happened to a child, was observed to happen, reported to happen, or any information which was pertinent to the child in school, especially the interaction of the child with his instructional program, his teachers, and his peers.

Observations as Pupil Personnel and Research Data

Observation as a method of obtaining information has a long history in educational and psychological research. While observations lack precision and objectivity and present problems in the estimation of reliability as compared with data such as test scores, rating scales and other psychometric devices, only observational methods offer the possibility of capturing a picture of the child's behavior in its natural context. Ignoring temporarily the issues frequently raised about observer contamination or influence, observations provide the opportunity to collect and maintain an ongoing systematic record of the behaviors of children and

teachers as they happen in the classroom. Several investigators have attempted to record classroom behavior using observation techniques. Some observation check lists or categories of behavior which were developed (as in the work of Wrightstone, 1935, and Rechless and Smith, 1934) employed free observations as a means by which to collect samples of behaviors observed in schools. The term "free observations" usually refers to those observations made without describing what the observers are to record. These free observation records were then studied to develop a categorical system for an observational rating scale which became the research tool. Observation categories also have been created from a factor analysis of ratings of classroom behaviors (Seton, Collins, and Koo, 1965).

Still other school observation systems have been designed for specific uses. Ryan (1959) trained observers to rate teacher behaviors on a scale of bipolar dimensions such as "understanding" vs "aloof". Flanders "interaction analysis" is an observational tool used to classify the verbal behavior of teachers and pupils as they interact in the classroom (Amidon, 1966). This system employs ten categories of verbal behavior, seven "teacher talk" categories, two "pupil" categories, and a tenth category of "silence, noise or confusion". The observer makes a tally every three seconds in one of the ten categories. This system allows greater possibility for establishing reliabilities, but, like several of the procedures described above, there is no record of what actually occurred to check the accuracy of the observational record.

Medley, Schluck, and Ames (1968) developed a procedure for coding teacher-pupil behaviors onto a mark sense answer sheet which offers the investigator a check list method of observation which is immediately

ready for data processing input. OScAR (Observation Schedule and Record) is their classification scheme for examining the verbal behavior of teachers and pupils in the classroom. OScAR has the limitation of being a closed rather than an open system for observation (since the number of entries is finite) and is also limited by the fact that the nature of the entry is lost since the observer only classifies, and is not required to describe the observation.

Several investigators attempted to use observers to provide a running account of the life of a child through the use of anecdotal or free observational techniques. Barker and Wright's book (1955) describes in detail their use of free observation in an ecological study of behavior of children in a small midwest town. Flory (1968) employed a transcription from a relatively unstructured interview to determine the nature of the child's play experiences at home as reported by the mother. He analyzed the transcription to sort responses into categories describing various aspects of play activities. A system for processing natural language observations into categories for computer input and retrieval has been developed by Caldwell, Honig, and Wynn (1968) under the acronym APPROACH (A Procedure for Patterning Responses of Adults and Children). In this system little attention is paid to the details of the observation procedure in terms of purpose of the observations, position of observer, types of behaviors recorded, or context in which observations were made.

From the foregoing description of uses of observation as a method for studying children in various contexts and as a research tool, it is apparent that few investigators have attempted to record actual behaviors, incidents, or anecdotes concerning individual children. The emphasis in

most observational research has been upon establishing reliability in rating or categorizing behavior and upon the development of scales and observation methods which allow data to be easily manipulated and analyzed.

We considered that the data obtained by a relatively free observational method (using psychologists, teachers, principals, or parents as the reporters of behavior) offered a valuable and unique source of information about children in school. Therefore, we decided to attempt the development of a system in which observers in the classroom would observe and record children's behaviors along with the context, antecedents, and consequences of those behaviors. We have also directed our attention to developing methods for efficiently and accurately recording and transcribing observations in such a way that they will become, when processed, more than a source of information about individual children, but a record of the child's progress through school, and an account by which the nature of children's experiences in school can be studied and compared.

Development of the Lexicon of School Observations

In order to define the event phenomena common to children's school experiences a major effort of the Stress of School Project was to study empirically the nature of behavior described in observations of children in the school setting and the nature of the interactions of the teacher with the child. We analyzed over 10,000 free observations of more than 2000 children in 75 classrooms in 10 schools to develop the "Lexicon of School Observations". Observers employed no a priori categories in making observations. Their only instructions were to make a record of descriptive statements of what they believed to be important information about children in school and to include in the record a statement describing

the antecedents and consequences of the behaviors as well as the interaction of the teacher with the child. The wide variety of behaviors recorded in the observations allowed us to abstract the following classes of significant school events.

Academic-intellectual behavior. (a) Signs of unusual strengths or weaknesses, including oral and written work, response to questions, teacher comments, etc. (b) Change from usual level of performance. (c) Response to new methods, teachers, etc. (d) Performance in various situations, including small group, individual and class instruction, and other instructional settings.

Social-emotional behavior. (a) Usual method of relating to peers. (b) Changes in peer relations. (c) Aggressive, attention-getting, immature or inappropriate behavior. (d) Successful coping with a difficult or potentially difficult situation. (e) Relationships with familiar adults--teachers, aides, etc. (f) Relationships with other adults--principal, nurse, counselor, substitute teachers, etc.

Physical appearance. (if out of ordinary, very good, or very poor). (a) General grooming. (b) Health. (c) Signs of fatigue. (d) Nutrition. (e) Stature and build.

General behavior. (a) Attention level. (b) Restlessness. (c) Alertness.

Other. (a) Contacts with other school officials. (b) Teacher reports. (c) Anything else considered important.

The "Lexicon of School Observations" is open-ended, allowing for the addition of categories as they are needed. In our work we did not focus on isolated teacher behavior since the objectives of the Stress of School

Project centered on teacher interaction with pupils. In the record of an event, antecedent teacher behaviors, teacher response to behavior, and the sequelae to the event are noted and processed as part of the event description. The APPLE programs retrieve teacher responses matched to pupil behavior. This made possible the addition of teacher behavior categories to the Lexicon from an analysis of observations of teacher-pupil interactions in a natural setting. Thus we were able to study teacher behavior in the classroom as a process of interacting with pupils and to maintain the approach of developing lexicon entries on observed rather than pre-determined categories of teacher behavior.

The logic of formatting a school event into fields which describe the school context in which the event occurred and parsing the annotation of the event into phrases which describe the behavior, giving each of these phrases a "modifier," which describes the nature of the annotation, closely follows the PsyCHES conceptualization of reducing natural language of the event into data processing form. Although APPLE required a different set of event names and modifiers, as many were retained from the PsyCHES list as seemed feasible. Abbreviations of event names in the APPLE and PsyCHES event systems are also usually the same to retain compatibility of the APPLE and PsyCHES lexicons. The lexicon in the APPLE system is open-ended. Investigators who wish to use a specific set of observations for testing hypotheses can insert their observation categories into the system.

Anecdotal Data other than Observations

The processing of other types of anecdotal data is under study at the present time. We are analyzing reports of consultations between teacher and psychologist, psychologist and principal, and psychologist and parent to develop formats for processing this important information.

Also under study are methods for processing teacher and parent conference reports. Already a part of the anecdotal information in the APPLE system are referrals, including presenting problems and follow-up placement, and annotations relative to quantitative data such as grades, test scores, and attendance follow-up which are carried along with the report information.

Since a finite list of school events does not have to be specified for the APPLE programs to be applicable, a school district user of the system may employ any set of event categories he wishes so long as he complies with the minimal requirements for formatting anecdotal data for input to the system.

APPLE File Organization, Storage, and Processing

Data Structure

Both the APPLE and the PsyCHES systems rely on storage in magnetic tapes manipulated by COBOL programs in large scale computers. The development of the programs has differed greatly. Much of the early PsyCHES programming was devoted to facilitating a transcription system that permitted keypunching a great many mnemonic codes. These codes were used by the keypunch operators in transcribing audio recordings prepared by the editors of the case history sources (usually research assistants reading and translating psychiatric protocols into event formats). Both APPLE and PsyCHES programs retain natural English language in the annotation of an event. However, PsyCHES inputs were almost completely free form and storage on magnetic tape was largely free form within large fields.

APPLE data, on the other hand, are largely structured. The high percentage of school data available in organized descriptive fields (names,

addresses, dates, etc.) and in arrays (attendance, test scores, report cards) dictated APPLE's high reliance on many special-purpose input card formats with fixed fields. Annotation of anecdotal information and observations is carried in 48-character fields. Magnetic tape storage of both anecdotal and quantitative data is formatted in card images to facilitate the use of outputs intermixed with card decks as inputs for further processing.

The APPLE data structure and the way data are organized into files give the system its potential as a research tool and its generalizability to a variety of administrative needs.

File Organization

APPLE data are maintained in four kinds of files: (a) rosters, (b) master files, (c) transaction files, and (d) interrogation files (queries, analysis, retrieval, abstraction, summary).

Rosters. These files are essentially lists of names. If the list has been gathered for some special purpose, some small amount of supplemental information may accompany the name. Schools generally have to compile many lists: class lists, honor rolls, graduation lists, athletic eligibility, visually impaired, high risk of failure, gifted, teacher, alumni, etc. It is assumed that APPLE rosters would serve many such needs. Rosters of pupils who are in special categories can be used to match against the files in order to abstract specified kinds of data from the dossiers of the pupils listed on the roster.

In a system of several thousand pupils we would expect a file to control the assignment of pupil dossier codes, showing the present status of all codes previously assigned (active, withdrawn, name change). New code

assignments would be checked against this roster and it would be searched for new students' names in the possibility that they are reentrants with already existing file materials. The code used in the APPLE system consists of the first three letters of the pupil's last name, followed by a three-digit number unique for that letter combination that places the name in alphabetic order within the alphabetic portion of the code. For example: Charles Johnson--JOH200; Linda Johnson--JOH230; Phillip Johnson--JOH245; Eric Johnston--JOH300. These codes can be assigned by a facility of the APPLE main program.

Master file. The master file is the data base, the repository for all data entering the system, and is continually updated with new material from the transaction files. It is stored on magnetic tape for economy and transferred to magnetic disc for high speed analysis and retrieval. The APPLE system organizes pupil data in the master files in the following hierarchy, which is described in detail below:

Cohort

Pupil dossier

Topic

Data species

Date

Entry

Line

1. Cohort. A cohort is a group of pupils who "travel" together in the school setting--essentially, those who are at the same grade level. On the way some old pupils may drop out and new pupils may enter, but the group as a whole can always be identified as "the class of 19__." In a

long-range record system the advantage of the cohort designation versus a changing grade level designation is obvious. Each cohort in the file has as its starting boundary a cohort entry, which contains school and cohort codes in a fixed field as well as a free-form description of the cohort. Cohorts are sequenced in the master file in ascending order of cohort code.

2. Pupil dossier. A pupil dossier is the collection of all entries belonging to one pupil. It always starts with the name entry, which contains the pupil dossier code and serves as the boundary point for the dossier. Pupil dossiers are sequenced within cohort in ascending order of dossier code.

3. Topic. Data in the pupil dossier are grouped into topics that cover different areas of the pupil's life. Topics presently implemented for APPLE are identification (i.e. name, address, etc.), family and socio-economic data, and school records (i.e. class history, absence records, report cards, test scores, and observations of pupil behavior). The high-order character of the two-character format code is used to identify the topic, thus permitting future expansion of the system to include other areas of data.

4. Data species. All entries containing the same kind of data in the same format are considered to belong to a data species. Data species are identified by the format code, and are sequenced within the dossier in ascending order of format code. Examples are name, test scores, observations. Some data species have unique entries, that is, they occur only once in a pupil dossier (e.g. name). Others by their nature have repeating entries; they may occure more than once (e.g. observations, test scores, referrals).

5. Date. Every repeating entry contains the entry-date in a uniform field. This is the date associated with the data (e.g. the date an observation was recorded, the date a test battery was administered, or the closing date of an absence report) and is not to be confused with the posting date, which indicates when an entry was incorporated into the master file. Repeating entries are sequenced within species in ascending date order. If several entries of the same date are input at the same time, they will be stored in the order of input. If subsequently another entry of that date is merged with the file, it will be sequenced after all previous entries of the same date.

6. Entry. An entry is a set of one or more lines. The first line is in a fixed-field format belonging to a particular data species and any subsequent lines are annotation. An observation of pupil behavior and a test score report are both entries. Any entry may include annotation. An entry is treated as a unit for purposes of storage and may be treated so in other processing such as sorting or abstraction.

7. Line. The line is the basic unit of input in the APPLE system. It is an 80-column punched card image. As stored on tape, it is identical in format to the punched card or teletype line on which it was originally input, except that the sequence number, which appeared in columns 75-80 of the punched card, appears on tape in columns 3-8, and the pupil dossier code, which appeared in columns 3-8 on the punched card, is moved to columns 75-80. In the master file the pupil dossier code is replaced by a posting date which indicates when the line was written in the file, and the sequence number is replaced by one of a new series that puts all the lines in the file in one continous sequence.

The format of each line, that is, the location and type of each item of data it contains, is indicated by a two character format code which appears in column 1-2. The format code has a number of functions: (a) It defines a keypunch card layout for input. (b) It defines the particular species of data that is punched in that layout (e.g. name, address, test scores, observations of pupil behavior) except for the general purpose formats noted in (e) below. (c) It defines a set of data items to be used in processing, especially in retrieval. (d) It determines an output format which permits the line to be printed in attractive, legible form and may signal the computer to output headings or captions and/or control vertical spacing on the page. (e) It sequences the line within the pupil dossier for purposes of storage and retrieval. Format codes appear in the dossier in ascending sequence, with the exception of certain formats whose codes use characters other than letters and numbers. Of special interest are the format codes with a "+" or a "-" in the low-order character (column 2). They are used for annotation, and a line with such a code has no intrinsic place in the sequence, but is treated, at the time of storage or retrieval as an appendage of the line it follows. Any characteristics of its contents, likewise, will depend on the species of data to which it is appended.

Transaction files. Information enters the system as "transactions" which are transcribed in batches into machine-sensible form, stored in temporary files on magnetic tape for further processing, and finally merged with the master file. These temporary files are called transaction files. Every line of data entering a transaction file contains a pupil dossier code in a file control field; this code remains in the line for positive

identification until it is incorporated into the master file. The line is then positively associated with a name entry that contains the pupil dossier code in a special field, and the code in the file control field is replaced by the posting date to maintain a history of the master file.

Before they are ready for merging with the master file, transaction files must undergo one or more of a series of processing steps which are performed by APPLE storage and maintenance facilities.

1. Reformatting. Where transactions are received as machine-sensible records in non-APPLE formats, the REFORMAT facility of the APPLE main program can abstract specified items from such records and incorporate them into an APPLE format to produce a file acceptable as APPLE input.

2. Sorting. Because APPLE records are accessed sequentially, a transaction file must be in correct sequence before it can be merged with the master file. The SORT PACKAGE, an APPLE supporting program, will sort files by cohort, pupil dossier code, format code and date, perserving the input order of entries within date and lines within entry.

3. Proof-listing and sequence numbering. At the time a batch of transactions is stored on magnetic tape it can be listed in report format, which spaces out items in each line for readability and adds captions and headings to make the data meaningful (see Fig. 1). The PROOF facility of the APPLE main program produces both the report and the file copy on magnetic tape. As part of the operation, the RESEQUENCE option will cause a 6-digit sequence number to be inserted in each line of the file. The lines are numbered in increments of ten to permit insertion of new lines.

4. Correction of data. The ARD (add-replace-delete) facility of the APPLE main program permits the user to delete any line or lines of a file and to insert new lines. A line may be replaced by deleting the old line and inserting a new one. Sequence numbers are used to find lines to be deleted and to determine where new lines are to be inserted.

5. File-editing. To assure that it will be possible to merge additional data with the master file and to retrieve data from it, the master file and all update material must meet certain criteria. Every batch of data must be headed by a cohort entry, since the cohort code is not carried in individual pupil entries. Every entry must be associated with a name entry, either in the file being edited or in the file with which it is being merged. No pupil dossier code may contain any illegal characters that would cause it to be sequenced out of order. And of course all data must be in proper sequence. The MERGE facility of the APPLE main program edits a transaction file by rejecting any entries that do not meet these criteria, either at the time of creating a master file, at the time of merging transactions with the master file, or at a previous time. It also creates, maintains, and checks for each file it processes a file status entry that indicates whether the file is a master file, whether its sequence has been disrupted by an ARD operation, and the data of the last update.

Interrogation files. Once the master file is established as the data base, it can be used to produce many kinds of files that contain subset and/or summary information.

1. File analysis. The retrieval capability of the main APPLE program permits the user to identify those pupils who do or do not satisfy specified requirements. These requirements can be expressed as logical products of up to fifty criteria. Counts are produced showing how the

pupils as individuals and groups satisfy each criterion.

2. Abstraction. Subset files may be produced by selection on two dimensions: (a) Pupil. Information may be abstracted for all pupils, or only for those on a selected list, including the kind of list produced by the file analysis described above. (b) Data. Complete dossiers may be abstracted, or subsets may consist of selected date ranges within selected data species.

On another level, selected data items (for all or for selected pupils) and/or summary information may be incorporated into user-specified formats to produce files for input to statistical and other programs.

When interrogation files are to be used for research purposes by parties not entitled to have access to confidential information, the DISGUISE option can be used to provide a copy of any file with disguised pupil dossier codes and with names and addresses obliterated.

Files produced by interrogation may need to be sorted in order to be useful. Disguised files, for instance, should be sorted on the disguised pupil dossier number. Subset files may need to be sorted on relevant data items. The SORT PACKAGE is a generalized program that can be used to sort files in a variety of ways on user-specified items, with specified entries treated as "masters" and the input order of other specified entries preserved.

Retrieval Power

The APPLE data structure facilitates abstraction of a wide variety of subsets of both anecdotal and fixed-field data from the pupil dossiers. It also facilitates application of a powerful system of logical analysis to the entire file for the purpose of identifying those dossiers of special

interest. Lastly, but of great importance for research applications, the structure facilitates the abstraction of numerical matrices in suitable form for input to statistical computer programs.

Flexibility of Content

The APPLE system permits an almost unlimited number of compatible formats, all of which can be manipulated by the processing facilities. The Stress of School Project has employed only a portion of the possible kinds of information which one can conceive of as being available for processing into the APPLE system. Further flexibility in the use of formats is made possible by the way the system handles array data. Formats that hold array information such as test, absence, and report card data, are designed to be user-defined. That is, while information must be stored in fields of fixed size, the user may utilize as many or as few of these fields as he chooses, and may define the information stored in each field by imposing his own report headings as part of the input and by accessing each field by its location rather than by predetermined names for items or groups of items. Thus the logic of the file system and the main and supporting programs permit the user to specify the information which is to be maintained, the categories of information to be processed, and the specific entries which are to be made.

References

Amidon, E. Interaction analysis applied to teaching. Bulletin of the National Association of Secondary School Principals, 1966, 50, 93-97.

Barker, R. G. & Wright, H. F. Midwest and its children: The psychological ecology of an American town. New York: Harper & Row, 1955.

Caldwell, B. M., Honig, A. S., & Wynn, R. L. Coding manual for APPROACH. Technical Report, 1968, State University of New York and Syracuse University, Grant No. D-156(R), Children's Bureau, Welfare Administration, and Grant No. MH-07649, National Institutes of Health, Department of Health, Education, and Welfare.

Eiduson, B. T., Brooks, S. H., & Motto, R. L. A generalized psychiatric information-processing system. Behavioral Science, 1966, 11(2), 133-143.

Flory, D. L. Intellectual development, social class, and the child's activities in the home. Unpublished doctoral dissertation, University of California, Berkeley, 1968.

Goodlad, J. I., Caffrey, J. G., O'Toole, J. F., & Tyler, L. L. Application of electronic data processing methods in education. Report of the Cooperative Research Project, January, 1965, University of California, Los Angeles, Project No. F-026, Office of Education, United States Department of Health, Education, and Welfare.

Medley, D. M., Schluck, C. G., & Ames, N. P. Assessing the learning environment in the classroom: A manual for users of OScAR 5V. Research memorandum #69-9. Princeton, N. J.: Educational Testing Service, 1968.

Reckless, W. C. & Smith, M. The agreement of three observers after practice in simultaneous recording of behavior. Journal of Applied Psychology, 1934, 18, 635-644.

Ryan, D. G. Some correlates of teacher behavior. Educational and Psychological Measurement, 1959, 19, 3-12.

Seton, D. A., Collins, D., & Koo, G. Factor analysis of the Winnetka Scale for rating school behavior. Journal of Experimental Education, 1965, 33, 373-378.

Wrightstone, J. W. Constructing an observational technique. Teacher College Record, 1935, 37, 1-9.

Footnote

[1]This research was supported by USPHS, NIH Grants MH 14,605-01 and 02.

Note

An APPLE system user's manual being developed will enable the user to select appropriate options, prepare input, and interpret output.

Two report-prefacing facilities are provided in the program in addition to the file-manipulation modules and processing options described previously. The user may select any of these facilities by means of a set control cards that precede any input data cards. The APPLE program control cards, referred to as "star-statements," are distinguished by an asterisk in column 1 and are used to select options and to provide parameters to the program modules. Figure 2 diagrams the way in which the report-prefacing facilities and the file-manipulations modules are selected.

Report-prefacing facilities. These options permit the user to introduce his own identifying material or documentation at the beginning of a report, following the title page.

The *.LIST statement causes text cards following it to be printed in 80-column format centered on the page. (The text cards must not have an asterisk in column 1, and the process continues until another star-statement is encountered).

The *.DISPLAY statement assembles a 132-character line from each pair of text cards following it, using a 66-column field from each card of the pair. This facility is particularly useful for developing new report formats to be incorporated into the program. Dummy reports can be produced by keypunching and easily modified until satisfactory formats are arrived at. Then the necessary programming can be implemented.

Any number of *.LIST and *.DISPLAY statements may be included in the program control set. Each one will cause the printout of the text cards following it to start on a new page.

File-manipulation modules. Of the six modules, whose functions are described briefly on pp. 24-26, five are selected by star-statements, i.e. REFORMAT, MERGE, ARD, ANALYZE, and ABSTRACT. The PROOF module is selected by default, if no other module is selected. Its only function is to produce a copy of an input file in report format and, optionally, to store it on tape. As indicated in the diagram, MERGE may be selected to follow REFORMAT or ABSTRACT may be selected to follow ANALYZE; in each case both operations will take place in the same job. Similarly, a series of files may be merged together by executing MERGE in several passes on successive files in one job. No other multiple operations are permissible. If an illegal combination of modules is selected, the first one shown in the diagram is executed, and any other module selection is ignored.

Concurrent optional processing. RESEQUENCE and DISGUISE may be selected with appropriate modules. These options are described respectively on p. 24 and p. 26. An option called PROOF may also be used when a non-default module is selected. When used as a concurrent process (rather than as a module by default), PROOF causes an output file to be printed in report format in addition to being stored as usual on tape.

STRESS OF SCHOOL PROJECT--UNIVERSITY OF CALIFORNIA

NADINE M LAMBERT--PRINCIPAL INVESTIGATOR

C S HARTSOUGH--PROJECT MANAGEMENT

F L CONVERSE--SYSTEM DESIGN

--

--

--

Figure 1. Sample Proof Report

INQUIRIES MAY BE ADDRESSED TO ANY OF THE ABOVE AS APPROPRIATE TO YOUR INTEREST

SCHOOL OF EDUCATION UNIVERSITY OF CALIFORNIA BERKELEY CA 94720 415 642-7581

SPONSORED UNDER GRANT MH14605-03 U S DEPARTMENT OF HEALTH, EDUCATION AND WELFARE

[PUPIL] CONVERSE ALEXANDER MILOVAN [BIRTH DT] 62 11 24 [SEX]M [ETH]O [FILE NO] CON777 [PREV FILE NO]
[NOTE] ALEXANDER MILOVAN KAROLY CHAIM KAPLAN CONVERSE

560 SANTA CLARA AV BERKELEY CA[ZIP]94707 [PHONE AREA CD] [DAY] 642-7581 [NIGHT] 524-7727 PRES PERM LEGAL

[SCHL HIST]	[ACTION]	[DISTRICT]	[ZIP]	[SCHOOL]	[STAFF]	[ACTIVITY]	[GD]	[GRP]
69:06:00	SOURCE	LOS ANGELES	90066	MARVISTA			01	
69:09:02	ENROLLED	BERKELEY	94707	KENNEDY	W E PARRIS	HMRDOM	02	B10
69:09:15	SPEC-TCH	BERKELEY	94707	KENNEDY	J K LADISH	SCIENCE	02	B10
69:10:00	SPEC-TCH	BERKELEY	94707	KENNEDY	K L MILLER	MUSIC	02	B10
69:12:18	THERAPY	BERKELEY	94707	KENNEDY	G O SCOTT	SPEECH	02	A04
70:06:14	PROMOTED	BERKELEY	94707	KENNEDY	K D DID	HMROOM	03	H22
70:12:11	WITHDREW	PALO ALTO	94303	HOOVER			03	

[ATTENDANCE] [N-MO]	[1/6 MONTH] EN AB OC TY	[2/7 MONTH] EN AB OC TY	[3/8 MONTH] EN AB OC TY	[4/9 MONTH] EN AB OC TY	[5/10 MONTH] EN AB OC TY	[SUMMER 1] EN AB OC TY	[SUMMER 2] EN AB OC TY
70:01:30	16 00 00 00	20 01 01 00	17 00 00 01	18 03 01 00	19 00 00 00		
70:06:12	17 01 01 00	15 00 00 01	19 00 01 00	20 00 00 00	15 00 00 00		
71:01:15	17 00 00 01	20 01 01 00	17 00 00 00	18 00 00 00	10 00 00 00		

[REFERRALS]	[SERVICE]	[RESOURCE]	[RQSTED BY]	[CAUSE]	[FOLLOW UP]
69:10:08	RECOMMENDATN	SPEECH	W E PARRIS	PERFORMANCE	69:11:
[NOTE]	TONGUE TRAINING NEEDED				
69:10:15	TEST-HEARING	AUDIOMETRIST	G O SCOTT	INSUFNT-INFO	69:10:

[TEST RECORDS]	[BATTERY]	[TEST]	[RAW]	[CVT]	[PTL]	[TEST]	[RAW]	[CVT]	[PTL]	[TEST]	[RAW]	[CVT]	[PTL]
69:11:00	COOP-PRI	MATH	087	573	620	LSTNG	033	482	470				
70:05:00	LRG-THRNDK	VERBAL	072	507	510	N-VRBL	082	550	580	TOTAL	154	524	540
[NOTE]	TEST INTERRUPTED	BOMB SCARE	TIME ADJUSTED										

[DATE/EVENT]	[ACTVTY]	[CONTEXT]	[LOCATN]	[SUBJ]	[MODIFIER]	[ANNOTATION]	[RPTR/CD]
69:09:17							
SITTING BEH	SCIENCE	T-CLASS	CLASSRM	SELF			
					NATURE	THE SUBJECT FALLS OFF HIS CHAIR WHILE SEEMINGLY JUST SITTING THERE.	
69:09:23							
ATTNTVNESS	READING	T-CLASS	CLASSRM	SELF			
					NATURE	SUBJECT SEEMS VERY INTERESTED AND EVEN INTENT UPON ANSWERING TEACHER-S QUESTIONS ABOUT WORDS.	
69:10:04							
PUPIL HELP	MATH	I-CLASS	CLASSRM	SELF			
					ANTECDNT	THE SUBJECT FOUND A DOLLAR ON THE FLOOR.	
					NATURE	THE SUBJECT GAVE IT TO THE TEACHER.	
					ADD INFO	ANOTHER PUPIL HAD LOST IT, THE TEACHER DISCOVERED.	

[DATE/EVENT]	[ACTVTY]	[CONTEXT]	[LOCATN]	[SUBJ]	[MODIFIER]	[ANNOTATION]	[RPTR/CD]
69:11:12							
ACAD PRFRMNC	N A	I-CLASS	CLASSRM	SELF			
					NATURE	TURNED IN PAPER, TEACHER GAVE IT AN A.	
					RESPONSE	EXCITEDLY TOLD SEVERAL CLASSMATES HE HAD GOTTEN AN A.	
69:11:19							
PHYS COORDN	N A	I-INDIVID	CLASSRM	SELF			
					NATURE	MUCH DIFFICULTY FOLLOWING LINES	
					STATUS	BUT COMPLETED IT REASONABLY WELL	
					CRCMSTNC	WITH TEACHER-S HELP.	
69:11:19							TEACHER
INATTENTION	N A	N A		SELF			
					NATURE	SUBJECT SELDOM PAYS ATTENTION TO WHAT IS SAID TO HIM	
					ADD INFO	ALTHOUGH HE TALKS TO OTHERS VERY WELL.	
69:12:05							
ART ACTIVITY	N A	I-INDIVID	CLASSRM	CAL250			
					OBJECT	CON777	
					ANTECDNT	PUPIL ASKED THE SUBJECT TO DRAW A WHALE FOR HIM.	
					NATURE	HE DREW IT IN THE MIDDLE OF HIS PICTURE, RATHER THAN IN THE WATER, AND <MESSED IT UP.>	
					ATTITUDE	SEEMED TO DO IT PURPOSELY, WAS DISSATISFIED WITH HIS OWN DRAWING.	
69:12:15							
VERBAL BEHAV	N A	T-GROUP	CLASSRM	MULTI			
					CO-SUBJ	HON283	
					CO-SUBJ	CON777	
					CRCMSTNC	WHILE OTHERS ARE TAKING TURNS AT CHARADES IN FRONT OF THE CLASS,	
					NATURE	TWO SUBJECTS ARE HUDDLED CLOSE TOGETHER, TALKING.	

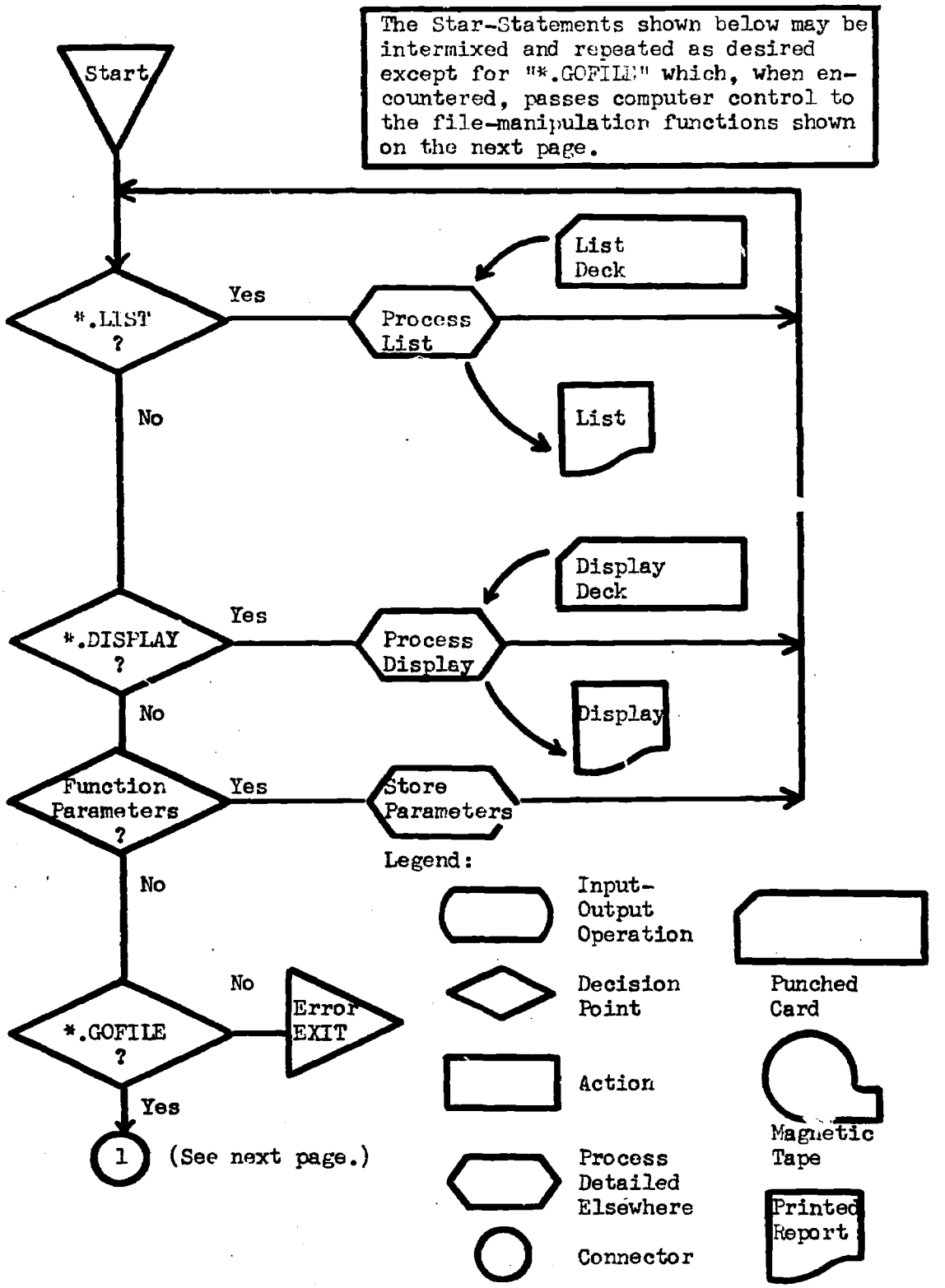

Figure 2. Overview of APPLE Main Program Options

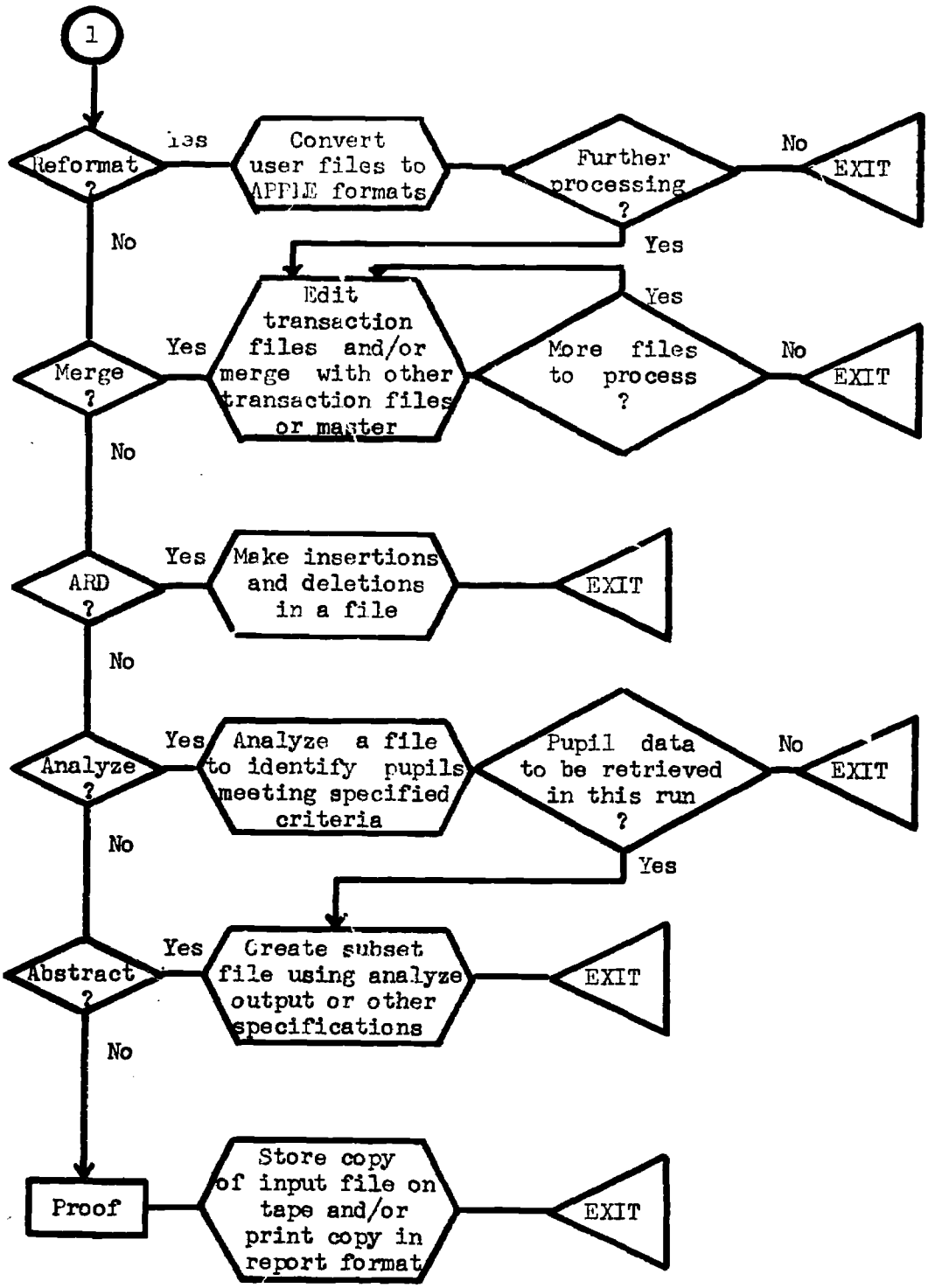

Figure 2 (continued). Overview of APPLE Main Program Options

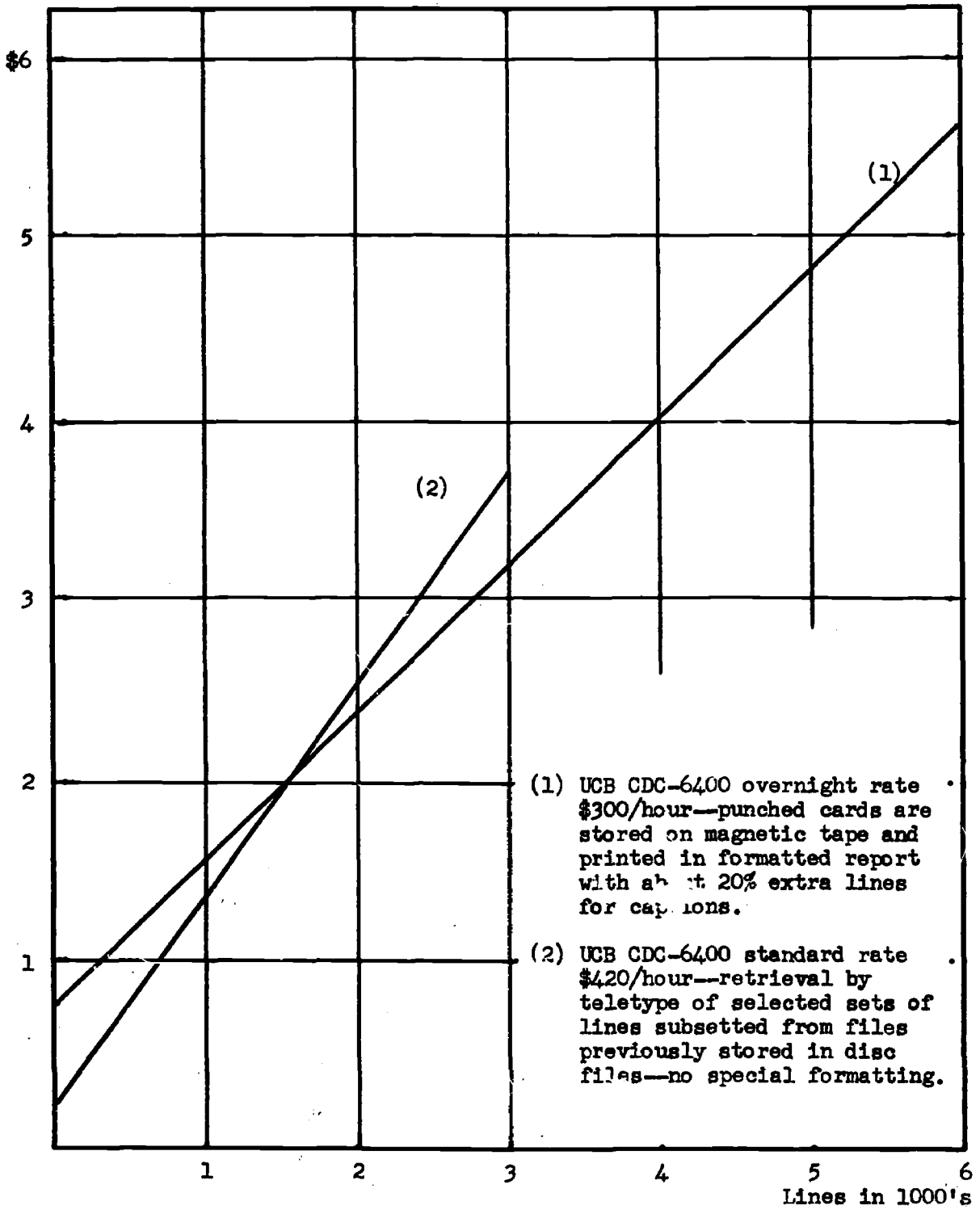

Figure 3. Computer Cost Estimates