To many, the computer seems both amazing and frightening, and as its use in educational settings is considered, attention is often focused on the possibilities for rigid standardization of the curriculum and for impersonalized teaching. We feel that the technology can be used to achieve just the opposite result, that is, to offer a vastly increased attention to the individual needs of students. A large portion of the research work of the Institute for Mathematical Studies in the Social Sciences at Stanford University has been devoted to the development of computer-assisted instruction (CAI) programs at the elementary school level. We have experimented with programs in drill and practice designed to provide practice in skills and concepts taught in the classroom and with tutorial programs designed to be fully instructional. We are already able to prescribe curriculum for each student at his own level of achievement and to allow him to work at his own rate. Although the technological innovations necessary to implement CAI have occurred only within the last decade, they have now reached a stage where it is possible to operate fully individualized instructional programs. We are now focusing our attention on the problems of designing an individualized curriculum.
A curriculum designed to be implemented by a computer must be more than a list of topics. It must contain a set of problems organized according to some explicit criteria. To select the content of any part of the curriculum, we need to answer such questions as: (a) Which problem types should be included? (b) How many of each type? (c) Should problems of different levels of difficulty be included; if so, how should the levels be determined? (d) What should the order of problem presentation be? In addition to examining some problems of curriculum development, we shall describe the way students interact with the computer in our present instructional programs and shall present evaluation data resulting from our attempts to assess the usefulness of the programs.

The instructional system in use at Stanford depends for its operation on a computer with time-sharing capabilities. The computer is able to interact simultaneously with each of several hundred students on an individual basis; both the lesson content and the timing of problem presentation may be different for each student.

A student taking a drill-and-practice lesson uses a teletype, a variety of electric typewriter, which is alternately controlled by the computer and the student. The computer directs the typing out of a problem, positions the teletype to receive a response, and turns the control over to the student. The student responds, using the modified typewriter keyboard of the teletype. For most problems, the computer regains control as soon as an answer has been typed, examines the answer, and either moves on to the next problem, if the response is correct, or responds with an error message, for example, "No, try again." For other problems, such as those involving word meanings, the student may execute a series of steps,
using the computer as a calculator. Control is returned to the computer only when the student is ready, for example, when he thinks he has a solution or when he needs a hint.

Block Program

The first mathematics curriculum developed, which we call the "block program," is designed for use by the teacher to supplement daily classroom work. The appropriate topics for each grade level were selected by conferring with teachers and examining texts and curriculum guides to determine which topics are customarily taught. Then lessons were written to provide drill work on these topics.

The basic structural unit of the curriculum is the block, which covers one or a selected mixture of two or three concepts. A block consists of seven lessons, each with sixteen to twenty problems. The first lesson serves as a preliminary or pretest, the next five lessons contain drill work, and the last lesson consists of a posttest on the material covered. For each of the five days of drill, five drills were written, one at each of five levels of difficulty. Thus, there is a total of twenty-five drills per block.

A block covering simple addition problems presented in horizontal format might contain problems like \(2 + 1 = \_\). Two other ways of presenting the same number fact, still in horizontal format, are \(2 + \_ = 3\), and \(\_ + 1 = 3\). We found that the last format is the most difficult for students and the first the easiest.

Further, certain number combinations are more difficult for students than others. For example, adding 0 or 1 to a number or adding doubles (1 + 1, 2 + 2, etc.) is among the easiest of the simple addition problems, but excluding these, problem difficulty increases with the size of the numbers involved. Thus, in constructing a block covering horizontal addition, even with such simple problems as those with one-digit addends, we have two variables to manipulate—position of the blank and size of the numbers—in designing drills of five difficulty levels.

Most of the major topics customarily presented to students in grades 1 through 6 are included in the block program. Between twenty and twenty-seven blocks were prepared for each grade; some focus on a single topic, others, referred to as "mixed drills,"

*February 1971 215*
include problems of more than one format (horizontal and vertical) or more than one operation (addition and subtraction, or addition, multiplication, and division). Although blocks are written for each grade, the teacher is free to select blocks from any grade level to meet the individual needs of individual students.

With the availability of drill lessons at five levels of difficulty, it is possible to give each student a drill that corresponds to his current level of achievement. A set of branching criteria were defined that place a child at a drill level on the basis of his performance on the block pretest and that provide for a change in the level of drill after each lesson. For example, a child who works less than 20 percent of the pretest problems correctly is assigned to level 1 (the easiest), while a child whose score is between 60 and 79 percent is assigned to level 4. Whenever his drill score for a day is below 60 percent, the student is given the next easier drill the next day; he is moved to the next harder drill level when the drill score is 80 percent or higher. The results of the posttest are used to select review material which is presented to each student at the end of his regular drill lesson.

The blocks presented to the students are selected by the classroom teacher, who is free to choose blocks from any grade level to meet the needs of his students.

The block program has been in operation since 1967. In the 1968–69 school year, approximately 6,000 students from California, Kentucky, Mississippi, Iowa, Ohio, Tennessee, and Washington, D.C. participated in the program. Each student worked at a teletype, located in or near his classroom, for from five to ten minutes a day. The teletype was connected by telephone lines to the central computer at Stanford University. Although the technical difficulties encountered in operating such a far-flung system are formidable and the patience of the participants was taxed more than once, in general the performance of the system has demonstrated the feasibility of such a centrally run computer-assisted instruction program. Some of the difficulties were certainly unanticipated. For example, in order to place teletypes in some of the Eastern Kentucky schools with which we were working, we had to deal with as many as seven local telephone companies.

The response of the children has been very favorable. Many adults have expressed concern about the impersonality of inter-
Patrick Suppes and Barbara Searle

acting with a computer—or rather, its representative, the teletype. The children, though, with their great ability to personify the inanimate, talk to, even yell at, hit, and kick the teletype and advise others to avoid the machines "that give hard problems." Working at a teletype is more engrossing for the student than the usual classroom situation. Each action brings a quick, relevant, individual response, and a demand for further action. Engaging the attention of the student is a crucial prerequisite for learning arithmetic skills and concepts, and a teletype in action certainly is an attention-getting device.

We have engaged in some formal evaluation of the effects of the drill-and-practice block program and will briefly describe some of the results. Although the program contains some material of a more conceptual nature, the major emphasis has been on the acquisition of computational skills. The Stanford Achievement Test was administered to test these skills and was given to classes of children in the program and to control classes of children of comparable age, grade level, and background not in the program. Achievement is measured for both groups as a gain in scores over a school year. The most extensive evaluation was conducted for students in California and Mississippi for the academic year 1967–68. The gain in computation scores was significantly greater for those in the experimental classes compared with those in the control classes for three grades in California and for all six grades in Mississippi. The larger number of significant gain scores in Mississippi reflects the poorer performance of the control classes compared with the control classes in California. The gain in grade level, averaged over all six grades, was the same for experimental classes in both California and Mississippi. Thus, for each region the average gain was 1.5, which means that on the average students placed one-and-a-half grade levels higher on the posttest than on the pretest. The average gain scores for the control classes was 1.2 for California and 0.8 for Mississippi. Thus, with the addition of the drill-and-practice program to their curriculum, the Mississippi students were able to keep pace with the California students (who, for the most part, came from relatively affluent, middle-class neighborhoods). The effect of the accumulated falling behind of the Mississippi students clearly shows in the comparison of the control-class scores at the end of the sixth grade. California sixth graders
The Computer Teaches Arithmetic

had an average grade placement of 7.6, the Mississippi students, 5.5. Evaluation results suggest that five to ten minutes' work each day at the teletype throughout the six elementary grades would provide enough practice on computational skills, so that the performance of the Mississipians would be comparable to that of the Californians.

We should make it clear, however, that more than just five to ten minutes of practice was involved in producing the gain scores noted above. We feel that work at the teletype is beneficial to the student because it provides graded drills, with immediate knowledge of results on each problem worked, expects the same of each student regardless of color or personality characteristics, and offers other advantages of computer-assisted instruction to be discussed in greater detail later.

The block curriculum was, at the time of its development, one of the most elaborate and highly individualized arithmetic programs available. However, as we gained experience with its implementation, we became more aware of its drawbacks. It was quite possible for a student to spend time working and performing well on one type of problem, say, addition, when what he really needed was more practice on multiplication. And it was possible for a student to be confronted with a problem type in which even the easiest examples were too difficult for him because he had not mastered the necessary subordinate skills. It became clear to us that in designing a program oriented toward supplementing the work in the classroom we were focusing attention on such questions as, "What grade is the child in?" and, "What is usually taught at that grade level?" rather than, "What concepts has the child mastered?" and, "What should this child learn next?"

These considerations suggested a major reorganization of the curriculum. First, each child should be given the opportunity to master subordinate skills and concepts before moving on to those of greater complexity or difficulty, no matter what his grade placement in school. This implies that the set of problems covering a single topic must be ordered according to difficulty and complexity so that the curriculum will contain the problem types providing practice in the appropriate skills and so that the student will encounter them in the proper order.

Second, it is a common observation that the difficulty experi-
enced with topics differs from one individual to the next, and these differences need to be accommodated in a curriculum designed to meet the needs of the individual. Thus, progress through each part of the curriculum should be independent of performance on other parts. Finally, it would be useful to provide extra practice for each student on the skills in which he is weakest, thereby attempting to raise the average level of his performance.

The Strand Program

We developed a new elementary mathematics curriculum, designed to meet the objectives mentioned above, which we called the “strand program.” This curriculum is divided into fifteen topics which include number concepts, horizontal addition, vertical addition, horizontal subtraction, vertical subtraction, measurement, equations, horizontal multiplication, vertical multiplication, division, decimals, laws of arithmetic, fractions, negative numbers, and problem solving. All the problems of a single topic are called a “strand,” and the problems of a single type within a strand constitute an “equivalence class.” Each equivalence class has an explicit definition in terms of the format of the problem, the size of the numbers involved, the operation used, and so forth. The equivalence classes, which number 725 for grades 1 through 6, cover all (or almost all) of the topics of elementary arithmetic. Any problem in the elementary curriculum that is covered by the strand program belongs to one and only one equivalence class.

Two questions enter into the definition of a strand: Which problem types should be included, and how should the classes that result from these choices be ordered? As our search for problem types led us to examine contemporary texts, we were astonished to find how widely texts differed with regard to the problem types chosen for each concept as well as in the order of presentation of these topics. We shall illustrate some of these differences because they emphasize the need for systematic research in the area of curriculum development. To date, no body of data exists that can serve as a guide for the kinds of curriculum decisions we have been discussing.

The topic of vertical addition lends itself well to a precise and orderly definition of problem types. Each problem can be charac-
The Computer Teaches Arithmetic

terized by the number of rows, the number of columns, and whether or not each row has the same number of digits. Thus, a problem of the form $\frac{ab}{+e}$ is said to have two rows, two columns and to be nonrectangular because one row has fewer than two digits. The problem type $\frac{ab}{+ed}$ is called rectangular. In addition, we can characterize problems by the size of the sum of each column. If the sum of any column except the left-hand one is greater than nine, the problem is said to have a “carry” in that column. If the sum of the left-hand column is greater than nine, we say the problem has a “basic addition fact.” A problem with a basic addition fact has an answer that is one digit longer than the longest row. Using these definitions, we constructed a set of equivalence classes in such a way that every possible vertical addition problem of a given size is included in a separate class. For the smaller problems (for example, two rows and two columns) it seems reasonable to assume that each student should have the opportunity to practice on each type, at least until such time as we have performance data to show that practice on one type is equivalent to practice on some other type.

However, in examining several second-grade texts we found only one that contained two-column, two-row problems with a basic addition fact and no carry, and only one (a different one) that had two-row, two-column problems with both a basic addition fact and a carry. We think it likely that such omissions occur because of the failure of textbook writers (ourselves included) to analyze the content of the curriculum in a really systematic way.

When we examined the order of introduction of different problem types in several standard texts, we found differences that were just as startling. For example, it is common practice, currently, to introduce the operations of multiplication and division at approximately the same time at the third-grade level. One third-grade text introduces all the basic multiplication facts first, followed by the corresponding division facts. The order in which the digits 1 through 9 are presented as factors is 1, 0, 2, 3, 4, 5, 6, 7, 8, and 9. Another text introduces products with one factor up to 5, the other up to 10, the related division facts, then the remainder of the products through $10 \times 10$, followed by the related division problems with divisors of 6 through 10. A third text takes an entirely different tack. Products with a factor of 5 are presented, and then
these multiplication facts are used to introduce a series of concepts, including dividing by 5, without and with remainder, finding multiples of a number, finding the greatest number of 5's in a number (solving inequalities), and finding areas. Then products with 2 as a factor are introduced, and the same series of concepts are discussed again. This pattern is followed, with factors presented in the following order: 0, 1, 10, 3, 4, and 6. Products with factors of 7, 8, and 9 are delayed until the fourth grade.

Each of these patterns of presentation represents the judgment of the respective authors of the texts and is based on experience and intuition. All may be equally effective, some one pattern may emerge as superior, or different patterns of ordering material will be found to be effective for different students. At any rate, the systematic examination of alternatives has scarcely begun. We are optimistic that the performance data we obtain from the strand program will begin to shed light on some of these questions.

Although we discovered that we could expect little guidance from texts on the selection and order of topics, we nevertheless used a detailed problem count to obtain approximate grade levels for problem types. We were also able to obtain performance data from the block program, and, relying on intuition as well as experience, we ordered the equivalence classes within each strand. We expect the structure established will undergo drastic revisions as we accumulate performance data for the students now in the program.

In summary, let us look at the structure of a strand. For a given strand the set of appropriate problem types has been divided into equivalence classes which have been ordered according to criteria of difficulty and complexity. Then, for each class, specific problems are written that fit the definition of that class. Within each class the problems are arranged in random order. A strand therefore consists of a set of problems arranged in a fixed, linear order, with a first problem, a second problem, and so on, to the last.

Although the strand program is designed to allow free movement of the student through the curriculum, the curriculum itself is organized according to grade levels. Each equivalence class is identified by a number called the "grade level" of the class. Since there may be up to twenty classes for each grade, the grade-level numbers are given to two decimal places. Thus, the first five classes
of the number concepts strand, which starts at the beginning of the first grade, are 1.00, 1.05, 1.10, 1.15, and 1.20. The class numbers specify the level of a strand at which a student is working and are used in computing an average grade placement for each student as a measure of progress through the strand program. Clearly, however, there is no necessary connection between a student’s real grade in school and his grade placement on the strand program. At present, the latter number cannot be said to represent an actual performance level standardized against a large group of students, but only a level of performance on the particular curriculum structure which developed.

Little has been said so far about the movement of the individual student through the strand program. We would like a movement scheme to satisfy the following criteria: Each student should progress in such a way that his movement on each strand is independent of his performance on any other; a student should spend more time doing problems on more important topics; he should have more practice on topics on which he performs poorly, less where he performs well.

To specify the appropriate mix of problems for each student at any given time, we define a “distribution function,” which may be thought of as a table that gives the proportion of problems the student should receive from each strand. We shall describe first how the distribution function depends on the curriculum, and how it depends on the performance of the student.

Decisions about the importance of topics in the curriculum were made after a careful examination of three typical elementary mathematics text series. For each text we classified each problem as one of the fifteen strand topics and counted the number of problems for each topic in each half of the book. The problem count was used to decide at what levels to begin and end each strand. Thus, for example, the horizontal addition strand was designed to extend from the beginning of first to the end of third grade to correspond with the occurrence of horizontal addition problems in the texts. The problem count was also used to define the proportion of problems at each half-grade level which was devoted to each strand topic. We found, for example, that for the first half of the second grade the distribution of topics was as follows: 16 percent number concepts, 14 percent vertical addition, 28 percent horizon-
tal addition, 12 percent vertical subtraction, 14 percent horizontal subtraction, 10 percent equations, and 6 percent measurement. We used these figures to determine what proportion of time the average second grader should spend working problems at the teletype from each appropriate strand. Using the problem count in this way, we were able to define a standard curriculum distribution for each half year for grades 1 through 6.

Once the curriculum distribution is determined, a straightforward application of probability theory is made to calculate the performance criterion for each equivalence class. This criterion might be that the student advances to the next class if he gets at least eight problems out of a block of eleven problems correct. The basis for the calculation is that the average student should be expected to make one year's progress in grade placement in one school year. The underlying model of student learning and performance is drawn from contemporary work in learning theory.

The basic distribution is modified for each student to give more weight to the strand on which the student is performing poorly. We shall not discuss the details of how this is accomplished; a relatively straightforward mathematical formula uses the difference between a student's performance on his best strand and the other strands and the curriculum distribution appropriate for the student's average performance level to specify an individually prescribed distribution function.

One further feature controls the movement of students on the strands. In order to provide more practice on specific problem types that are difficult for the student and less on those which are easy, a student is programmed to skip problems in a strand when he makes a series of correct responses or to move backward, and hence repeat problems, when he makes incorrect responses.

What sort of lesson can a student on the strand program expect when he sits down at the teletype? After signing in with his name and number, he is presented with a problem from a strand that has been selected at random (with a probability determined by his distribution function). The specific problem presented to him depends on the last problem he worked from that strand at any previous session and whether he got that previous problem correct or incorrect. If now he responds correctly, he will be given a new problem, probably from a different strand. If he responds incor-
rectly, he will receive the message, "No, try again," and the problem will be retyped. A second incorrect response will bring the message, "No, the answer is ———. Try again," with the correct answer provided. Thus, before moving on, the student has the opportunity to see, and type in, the correct response. As the student moves through the session, which has a fixed time, the number of problems he receives will depend on how fast he works and how many incorrect responses he makes. He continues to receive a mixture of problem types, which result from the mix of strands he is working on and his movement from one equivalence class to the next on each strand. At the end of the session, the student is told how many problems he worked and the percentage correct. Then the teletype bids him good-bye, by name, and asks that he tear off his printout on the dotted line. The student then has a printed record of the entire lesson, which he may use in any way he chooses.

The strand program described represents the current state of our thinking about how best to structure an individualized drill-and-practice mathematics curriculum. At the time of developing the block program we did not foresee those disadvantages that eventually led us to restructure the curriculum into strands. Now, as operational experience with the strand program accumulates, we anticipate further growth in our understanding of how to individualize instructional programs.

There are, however, some characteristics of both programs that are known to enhance learning and that we expect to retain in further modifications. Students learn faster when they are told immediately that their replies are right or wrong and when they are given the opportunity to correct their errors. Both these features are part of our programs.

By matching student capabilities with the difficulty level of problems, we can prevent the experience of repeated failure that is so discouraging, especially for slower students, and we can avoid the excessive repetitions that bore the brighter children.

Many schools classify students at the beginning of a year, or even upon entrance, as "slow" or "fast" and, using these categorizations, provide different instructional programs or "tracks." The individualized CAI program accomplishes a kind of "micro-tracking," selecting the appropriate level of instruction for each small topic,
thus avoiding the labeling of children and the frequent mismatch between capabilities and level of instruction.

Several advantages accrue to the teachers of students receiving CAI. The computer relieves the teacher of the dreary and time-consuming task of correcting drill exercises, keeps records, and diagnoses areas of students' strengths and weaknesses. The teacher can use the diagnoses for designing special instruction for individual students and can devote the class time previously applied to drill to more imaginative activities.

We feel encouraged that our experience with CAI will lead us to increasingly sophisticated ways of viewing the problem of providing for each child a continuing learning experience that will meet his individual needs and capabilities.
