1. Introduction

These lectures are addressed to mathematics students, but a certain amount of immaturity with computer science will be essential for a complete understanding.

Computer programs are algorithms written in a very stylized format and meant to be understood by a machine. The style in which we will write algorithms below is called SETL (SET Language) and has been devised by Jack Schwartz of the Courant Institute of Mathematical Sciences (N.Y.U.). It is a style that is very comfortable to a mathematician since it uses standard mathematical notations and modes of expression. The few new notions that will be required will be explained as we go along.

For our purposes a machine may be thought of as a black box, inside of which is a math graduate student slave. The student has a finite amount of paper, does not work infinitely fast, and understands the algorithms. She may respond by typing out answers on a piece of paper or printing them out on a TV-like screen.

A dialect of SETL, called SETLB, is currently running on the AEC CDC-6600 at N.Y.U. Future implementations will be orders of magnitude more efficient than the present one.
2. The Data Bank

A possibly multivalued function

\[ g: A \rightarrow B \]

may be defined by giving its graph explicitly as a set of ordered pairs. The following notations will be used.

\[
g(x) = \begin{cases} \text{unique } y \in B \text{ with } <x,y> \in g & \text{if} \\ \omega \text{ (or undefined) otherwise,} \end{cases}
\]

\[
g\{x\} = \{y \in B \mid <x,y> \in g\},
\]

and for a set \( C \subseteq A \)

\[
g[C] = \{y \in B \mid (\exists x \in C \mid <x,y> \in g)\}
\]

A data base is a tuple

\[ <A,f_1,\ldots,f_m,R_1,\ldots,R_n> \]

where \( A \) is a finite set, each \( f_i \) is a function on \( A \), and each \( R_j \) is a binary relation on \( A \).

Information that may be accessed includes all sets in the weak topology on \( A \) generated by the functions and relations defined on \( A \). A typical such set may have the form

\[
\{x \in A \mid f(x) \geq 5 \text{ and } (g\{x\} \in K \text{ or } Q(x) \neq 6)\}.
\]

This information may be augmented by using the operators:

\[ \text{arb}(S) \text{ and } \text{random}(S) \]

which chose any old element from \( S \), and a carefully chosen random element respectively.

Suppose

\[ g = \{<2,8>,<3,6>,<4,7>,<2,6>,<8,1>\}. \]

Then

\[ g(3) \text{ eq 6 , g(4) eq 7,} \]
\[ g(2) \text{ eq } \omega, \ g(15) \text{ eq } \omega, \]
\[ g\{3\} \text{ eq }\{6\}, \ g\{2\} \text{ eq }\{8,6\}, \]
\[ g\{15\} \text{ eq }\{\}\text{ (null set),} \]
\[ g\{[2,3,4]\} \text{ eq }\{8,6,7\}, \ g\{[8,9,10]\} \text{ eq }\{1\}, \]
\[ g\{[-2,1]\} \text{ eq }\{\}\]
A great deal of information can be stored in such a data base, the 1970 census for example, and this information can be used in a variety of ways.

Information may be added to the data base by augmenting the base A and updating the functions and relations. For example, if b is given along with data about b, we merely need write

\[ b \in A; f_1(b) = c; f_2(b) = g; \ldots \]
\[ R_1(b) = \{q\}; R_2(b) = \{r, s, t\}; \ldots . \]

To delete a member from the data base we proceed as follows:

\[ c \text{ out } A; f_1(c) = \omega; f_2(c) = \omega; \ldots \]
\[ R_1(c) = \omega; \ldots . \]

New functions may be defined using the operations of intersection, union, composition, transitive closure, etc. For example, if R denotes the relation "friendsof" then

\[ Q = R \cup (R \circ R) \]

is a new relation. For each \( x \in A \)
\( Q(x) \) consists of all elements of A that are either friends of \( x \) or friends of friends of \( x \).

Of interest to a sports car salesman might be the set

\[ \text{potentialcustomers} = \{ x \in A \mid \text{age}(x) \geq 18 \text{ and } \text{age}(x) \leq 30 \text{ and } \text{income}(x) \geq 20000 \text{ and } \text{sex}(x) \equiv \text{male} \text{ and } \text{trafficpenaltypoints}(x) \leq 4 \}; \]

The salesman may get a list to give his secretary as follows:

\[ (\forall x \in \text{potentialcustomers}) \]
\[ \text{print name}(x), \text{phoneno}(x); \text{end} \forall x; \]

or he may print out the following type of form letter:

\[ (\forall x \in \text{potentialcustomers}) \]
\[ \text{print name}(x); \]
\[ \text{print address}(x); \text{print}; \]
\[ \text{print "Dear", name}(x); \]
\[ \text{print "As a man with an income of", income}(x), \]
\[ \text{"I am sure you will be interested in our new", favorite_color}(x), \text{"dinkycar"}.} \]
\[ \text{print "Some weeks ago I ran into your friend", random friendsof}(x), \text{"and he mentioned that you were dissatisfied with your", presentcar}(x); \]
\[ \text{common = clubs}(x) \cap \text{clubs(number("HalBean")));} \]
\[ \text{if common ne nl then print "Perhaps I'll see you at the", random(common); else print "Please drop by my office"; \}
\[ \text{print "Personally yours";} \]
\[ \text{print "Hal Bean";} \]
3. Simulation

There are many types of situations that can be modeled on a machine. When complete information is known, the model is decisive and empirical evidence can often be used in place of a more difficult analysis of the situation. This will be illustrated with the game of "craps."

A player rolls two dice to establish his point \( \in \{2,3,\ldots,12\} \). If his point is 2, 3, or 12 he loses. If it is 7 or 11 he wins. If neither of the above is the case, he continues rolling until he rolls a 7 in which case he loses, or his point again in which case he wins. A probabilistic analysis shows that the odds are slightly in favor of the house. An empirical estimate of this is given on the right where the game is played one million times and the number of wins is recorded. The program illustrates the use of labels, while and go to statements, as well as the use of a procedure.

There are two types of procedures in SETL. An example of a functional type of procedure is

```plaintext
define g(x); return x^2-3x + 4; •
```

A procedure of subroutine type does not return a value, it does something and then merely returns. An example might be

```plaintext
define message(x); print x; return; •
```

```plaintext
craps = \{2,3,12\}; wins = \{7,11\};
nowins = 0; nogames = 0;
die = \{1 \leq n \leq 6\};

define roll; return random(die) + random(die);

playgame: if nogames \geq 1000000 then
  go to finished;;
  no games = nogames + 1; point = roll;
  if point \in craps then go to playgame; else
  if point \in wins then nowins = nowins + 1;
  go to playgame;;
  newroll = roll;
  (while newroll \notin \{7, point\}) newroll = roll;;
  if newroll \equiv point then nowins = nowins + 1; go to playgame;;

finished: print nowins,"games won out of",10000;
```

The following conversations actually occurred.

M=mathematician C=computer scientist MC=mathematician and computer scientist

M to MC: What is a procedure?
MC to M: It's like a lemma!
C to MC: What is a lemma?
MC to C: It's like a procedure!
4. Comparison of Notations

It is quite easy to translate a mathematical algorithm into a SETL program. Translation of some typically mathematical phrases or constructions are shown below:

1. The $\Gamma$-function
   \[ \Gamma(1) = 1 \text{ and for } n > 1, \Gamma(n) = n\Gamma(n-1) \]

2. Let \( \{S_1, S_2, \ldots, S_n\} \) denote the elements of the set of integers $A$ arranged in increasing order.

3. The elements of the above set arranged in decreasing order.

4. Is $n$ a prime number?

5. The smallest prime greater than a given $n$

6. A list of the first 500 primes.

7. The set of primes smaller than 500

8. The above set in increasing order

9. The greatest common denominator of two positive integers $m$ and $n$.

10. Single valued function

11. Domain of function

12. Inverse of function

The programs below are not written from the point of view of efficiency which is a serious issue and will be considered later.

```setl
define f(n); return if n eq 1 then 1 else n*f(n-1); ·

define incorder(A); S=nl; (\forall x \in A) k=1+\#\{y \in A | y \lt x\}; S(k)=x; end \forall x; ·

define decoder(A); S=incorder(A); (1 \leq n \leq \#S) S(n)=S(n-\#S+1); end \forall n; return S; ·

define prime(n); return if(1 \leq k < n | n \div k eq 0) then false else true; ·

define primeafter(n); k=n+1; (while not prime(k)) k=k+1; end while; return k; ·

list(1)=1; (while \#list lt 500) list(\#list+1)=primeafter(list(\#list)); end while

primes={1 < p < 500 | prime(k)};

incorder(primes);

gcd(m,n); return if m eq n then n else gcd(|m-n|,\min{m,n}); ·

function(f); return type(f) eq set and (\forall x \in f | \pair(x)) and (\forall x \in f, y \in f | x(1) eq y(1) implies x(2) eq y(2)); ·

domain(f); return \{x(1), x \in f\}; ·

inverse(f); return \{<x(2), x(1)>, x \in f\}; ·
```
13. Image of function

14. Injectivity

15. Composition of functions (or relations)

16. The $n^{th}$ power of a function

17. The transitive closure $R^*$ of a relation $R$

18. The set of all functions from $A$ to $B$

19. Permutations of a set $A$

20. Given an integer $n>1$, if it is even divide it by 2, and if it is odd multiply by 3 and add 1. It is not known if iteration of this always leads to 1. For $n=7$ we have 7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1

21. The pair of prime twins greater than $n$.

22. The center of a group $g$ whose multiplication table is $m$.

define image(f); return domain(inverse(f));

define injective(f); return(\forall x \in domain(f), y \in domain(f) | x \neq y \implies f(x) \neq f(y));

define $R \circ S$; return \{<x(1), y(2)>, x \in R, y \in S | x(2) eq y(1)};

define power(R, n); return if $n \leq 1$ then $R$ else if $n \leq 2$ then $R \circ R$ else $R \circ power(R, n-1)$;

define star(R); $T=R; n=2;$
  (while power(R, n) ne n) $n=n+1;$
  $T=T \cup \text{power}(R, n); ; \text{return} \ T;$

define fncts(A, B); return \{f \in \text{powset}(A \times B) | \text{domain}(f) eq A and function (A)};

define permutations(A); return \{f \in \text{fncts}(A, A) | \text{injective}(f)};

define getstone(n); return if $n \leq 1$ then true else if even($n$) then getstone($n/2$) else getstone($3*n+1$);

define even(n); return $2*(n/2) \ eq n$;

define nextprimepair(n); $k=n+1;$
  (while not (prime($k$) and prime($k+2$)))
  $k=k+2;$ end while; print <k, k+2>;

define center = \{y \in g | (\forall x \in g | m(x, y) eq m(y, x))$;
Most programming languages, e.g., FORTRAN, BASIC, APL, ALGOL, LISP, PL/1, COBOL, have evolved up from the machine and programming manuals often read like old fashioned text books on tensor analysis. SETL descends into computer science from mathematics, and is essentially a coordinate-free approach to programming. Problems of efficiency, which often amount to something like choosing a suitable coordinate system, can be handled automatically. The essential contribution of SETL is the introduction of abstract objects in place of artificial coordinate representations of them.

There seems to be a difference of opinion concerning the SETL style of programming as compared to the current one. The humorous excerpt below is taken from a 1971 review of a 300-page SETL manuscript written by an "expert" on programming languages.

"In summary then, SETL is at worst just a collection of strange notations and devices, and at best it is "just another programming language." Compared with the elegance and clean design of APL, SETL fails to attract the mathematician's mind. It does not use the mathematician's symbols, his notation, his precedences, or his identities. Its mass of petty detail is no smaller than that of other languages. The algorithms presented are little more than transliterations of what would be written in ALGOL or APL. Yet the idea of a set as a datatype (or data structure) and the partially-fulfilled idea of specifying operations on all the elements of a set are very powerful notions and are good candidates for incorporation in some existing programming languages."
5. The Labyrinth

The following is from Algorithms and Automatic Computing Machines by B.A. Trakhtenbrot.
The setting is a symmetric relation $R$ over a set of nodes. Two nodes are given and the problem is to construct a path between them (if one exists).

In order to construct such an algorithm, we prescribe a special method of search. At each step of the search we can separate the corridors into three classes: those through which Theseus has never passed (we shall call these green corridors), those through which he has passed once (yellow), and those through which he has passed twice (red). Furthermore, from any junction Theseus may move to an adjacent junction in one of two ways:

1. Unwinding the thread. Theseus moves along any green corridor to an adjacent junction, unwinding Ariadne's thread as he goes; this corridor is then considered yellow.

2. Rewinding the thread. Theseus returns along a yellow corridor to an adjacent junction, rewinding Ariadne's thread as he goes; this corridor is then considered red.

Note that Theseus is not allowed to go through a red corridor. We assume that Theseus makes some mark by which he can later distinguish a green corridor from a red one. He can distinguish the yellow corridors because they have Ariadne's thread stretched along them. The choice of each move depends upon the conditions which Theseus finds at the junction where he happens to be. These conditions will be one or more of the following:

1. Minotaur. The Minotaur is discovered at the given junction.
2. Loop. Ariadne's thread already passes through the given junction; in other words, there are at least two other yellow corridors leading from the junction.
3. Green. There is at least one green corridor leading from the junction.

4. Ariadne. Ariadne is at the given junction.

A fifth case. None of the above conditions prevails. Our search method may now be described by means of the following table.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Move</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Minotaur</td>
<td>Stop</td>
</tr>
<tr>
<td>2. Loop</td>
<td>Rewind the thread</td>
</tr>
<tr>
<td>3. Green</td>
<td>Unwind the thread</td>
</tr>
<tr>
<td>4. Ariadne</td>
<td>Stop</td>
</tr>
<tr>
<td>5. Fifth case</td>
<td>Rewind the thread</td>
</tr>
</tbody>
</table>

define labyrinth(ariadne, minotaur, R);
green=R; yellow=nl; red=nl; marked=nl;
/* at each step after the first we have just traversed a path <a,b>. We have the following cases*/

**case1 = b eq minotaur**
**case2 = b marked**
**case3 = {x ε green | x(l) eq b} ne nl**
**case4 = b eq ariadne**
/* the possible actions are*/
**stop = go to done**
**rewind = <a,b> in red; <b,a> in red; <a,b> out yellow; <b,a> out yellow; <a,b> = <b,a>; go to nextmove**
**unwind = x = arb {y ε green | y(l) eq b}; b in marked; <a,b> = x; <a,b> in yellow; <b,a> in yellow; <a,b> out green; <b,a> out green; go to nextmove**
x=arb{y ε R | y(l) eq ariadne}; <a,b> = x; <a,b> in yellow; <b,a> in yellow; <a,b> out green; <b,a> out green; nextmove: if case 1 then stop; if case2 then rewind; if case3 then unwind; if case4 then stop; rewind;
done: if b eq ariadne then print "no path exists" else print "The solution is", yellow; end if;
6. Finite Groups

We will write an algorithm to check the truth of the Feit-Thompson theorem for all groups of even order up to one million.

```
define group(G); <e,g,m,i> = G; return
type g eq set and type m eq set and
type i eq set and
e ε g and i[g] eq g and m[g,g] eq g and
(∀xεg m(x,e) eq x and m(e,x) eq x) and
(∀x,y,zεg m(x,i(x)) eq e and m(i(x),x) eq e) and
(∀x,y,zεg m(x,m(y,z)) eq m(m(x,y),z));
```

```
define subgroup(h,G); <e,g,m,i> = G; return hcg and group(<e,h,m,i>);
```

```
define normalsubgroup(h,G); return subgroup(h,G) and (∀xεh, yεG(2)|m(i(y),m(x,y))εh);
```

```
define generatedsubgroup(A,G); <e,g,m,i>=G; G=Awithe; n=0; (while n lt #B) n=#B;
(∀x,yεB) m(x,i(y)) in B; end ∀x;
end while; return B;
```

```
define generatednormalsubgroup(A,G); <e,g,m,i>=G; A ⊂ B and normalsubgroup(B,G)]B;
```

```
define commutator(G); <e,g,m,i>=G;
A={m(x,m(y,m(i(x),i(y))),xεg,yεg};
return generatednormalsubgroup(A,G);
```

9.

```
define quotient(G,R); <e,g,m,i> = G;
cosets = nl; (while g ne nl) x = arb g;
coset = {yεg|m(x,i(y)|εh};
coset in cosets; coset outof g; end while;
M=nl; I=nl;
(∀aεcosets, bεcosets) M(a,b)=m[a,b];
I(a) = i[a]; end ∀a; return
<h,cosets,M,I>;
```

```
define solvable(G); (while G ne
commutator(G)) G=quotient(G,
commutator(G)); end while; return
#G(2) eq 1;
```

```
A = {1<n<100000}; g = permutations(A);
i = {<x,Inverse(x)>, xεg};
m = {<x,y,x•y>, xεg,yεg};
e = {<a,a>, aεA};
P = <e,g,m,i>;
```

```
if ∃ B ⊂ g | subgroup(B,P) and even (#B)
and not solvable (<e,B,m,i>) then
print "Feit-Thompson Theorem is wrong";
else print "OK through 1000000";
```

7. Drawing Pictures

A = <70,30>; B = <70,10>; C = <90,10>; 
D = <90,30>; E = <80,40>; 

house = brokenline(<A,B,C,D,E,A,D>); 
sun = {<x,z>£P(x-80)^2+(y-80)^2 le 100}; 
ray = {<x,y>£P y ~ 80 and 95 le x le 100}; 
rays = ray; (1<Vn<100)rotated(ray,<80,80>,n*30) into rays; end Vn; 
tower = {<x,2y>,<x,y>£translated(house,<-60,0>)}; 
picture = house u sun u rays u tower; 
draw(picture); 

The code to implement the above is given on the right. A three dimensional version could be used to draw a picture in one perspective, rotate the axes, and then view it from a different perspective.
Assuming that the draw routine prints on 35mm film one can use the routines above to make movies. Suppose the sets

tail, body, fleg, bleg

are given along with the pivots A, B, C and
dog = tail \cup body \cup fleg \cup bleg. Then
(1 \leq n \leq 50) draw \(dogs(n)\);
will produce the movie on the right if the
dogs routine is defined as follows:

define dogs(n); if odd(n) then return
trans(dog,\langle n,0 \rangle);
\theta = if divides(4,n) then -1 else 1;
newdog = body \cup
rot(fleg,C,\theta*45) \cup
rot(bleg,B,-\theta*45) \cup
rot(tail,A,\theta*30);
return trans(newdog,\langle n,0 \rangle);
A powerful aspect of SETL is that it allows procedure names to be elements of sets or components of tuples. This use will be illustrated below.

A binary tree is either an atom or a pair of trees. The atoms in the tree are called its leaves. Two trees will be called similar if they have the same structure when the leaves are ignored. A forest is defined as a set of trees. A forest may be enlarged by adding to it all trees of the form T(2) when T(1) and T already belong to the forest. Iterating this procedure will eventually lead to a fully grown forest. Suppose F is a forest. Then the program

```
(while #F ≤ # enlarged(F)) F=enlarged(F) ;;
```

expands F until it produces no new trees. The growth can be sped up considerably using features where a feature (of trees) is a function defined on all trees. Using the definition on the right of enlarging the forest, F, with respect to a set of features, feats, gives the more efficient growth

```
(while #F ≤ # enlargedwrt(feats, F))
F = enlargedwrt(feats, F) ;;
```

How efficient the growth is dependent on how quickly the features can be evaluated. If f is a feature and the test f(S) eq f(T) can be computed much more quickly than the test S sim T then the feature may be useful.
9. Real Reals

This section follows the spirit of Errett Bishop's Foundations of Constructive Analysis (pp. 18, 19, 26).

Let $\mathbb{Z}^+$ denote the positive integers.

A procedure $x$ defined $\forall n \in \mathbb{Z}^+$ is said to be real if for every $k \in \mathbb{Z}^+$ there is an $A(k) \in \mathbb{Z}^+$ satisfying

$$|x(n) - x(m)| \leq \frac{1}{k} \quad \forall m, n \geq A(k).$$

Two reals $x$ and $y$ are equal ($x = y$) if for every $k \in \mathbb{Z}^+$ there is an $B(k) \in \mathbb{Z}^+$ satisfying

$$|x(n) - y(n)| \leq \frac{1}{k} \quad \forall n \geq B(k).$$

A real $x$ is positive ($x > 0$) if for some $k \in \mathbb{Z}^+$ there is a $C(k) \in \mathbb{Z}^+$ such that

$$x(n) \geq \frac{1}{k} \quad \forall n \geq C(k).$$

A real $x$ is nonnegative ($x \geq 0$) if for each $k \in \mathbb{Z}^+$ there is a $D(k) \in \mathbb{Z}^+$ such that

$$x(n) \geq -\frac{1}{k} \quad \forall n \geq D(k).$$

It can occur that a given real number $x$ is known to satisfy $x > 0$, but that neither $x > 0$ nor $x = 0$ is known to be true. More striking is the fact that there is no general method (nor will there ever be one) for deducing that $x > 0$ or $x = 0$ given that $x \geq 0$. Examples are given on the right.

\begin{verbatim}
define a(n); return 0; \bullet
define b(n); return 1/n; \bullet
define c(n); return n/(n+1); \bullet
define d(n); return 13. if (1<3k<2n prime(k) and prime(2n-k))
         then 0 else 1; \bullet
define e(n); return[+:1\leq k\leq n]d(k)/10^k; \bullet
\end{verbatim}

The procedures $a, b, c$, and $e$ are real numbers. On the other hand the procedure $d$ is not known to be real. It generates a sequence of zeros and ones, but the Cauchy criterion cannot be verified since it would involve finding a proof of Goldbach's conjecture.

The number $e$ satisfies $e \geq 0$ since we have that $e(n) \geq 0$ for all $n$. To show that $e = 0$ would amount to showing that Goldbach's conjecture were true and a proof that $e > 0$ would provide a counter example to the conjecture. Hence, from $e > 0$ we cannot conclude either $e > 0$ or $e = 0$, as is done in the formal setting. In fact, many basic results in formal analysis are false in the constructive sense.
10. References and Remarks


   This is the basic reference work on SETL. It includes a discussion of programming principles, details concerning the design of the language, and a variety of algorithms written up in SETL.

2. SETL-Newsletters 1-100+. N.Y.U. Dittoed Notes.

   A miscellaneous collection of papers mostly concerned with implementation problems, but some have an independent theoretical importance. Variously authored.


   A users' manual for SETLB with all the gory details explained and illustrated.


   A comprehensive discussion of the problems involved in getting a machine to understand a language. Parsing algorithms and optimization techniques are discussed at great length.


   These two papers are concerned with a quite different approach to programming. It will be illustrated with a few examples.

   Suppose x = <2,8,6,4>.

   Then, setting y=x(3) gives y the value 6.

   By contrast, setting x(3)=45 will change the value of x which then becomes <2,8,45,4>.

   Rewriting x(3)=45 as x(3) eq 45 = true, opens the door to a new style of programming. For example, instead of writing an algorithm to solve the pair of equations

   \[
   2x + 3y = 7 \quad & \quad 4x - 5y = 3,
   \]

   one would merely write

   \[
   2x + 3y \text{ eq } 7 \quad \text{and} \quad 4x - 5y \text{ eq } 3 = \text{true};
   \]

   The compiler would then fiddle away until it established the truth of this statement by finding \( x = 2 \) and \( y = 1 \).

   The general principle here is quite simple and programs could be written in the following form:

   Whatever you wish to be true = true.
A programming language developed along these lines would lead to a descriptive, rather than algorithmic, style of programming, and this could be very powerful. A descriptive sort of a sequence \( x(m) \) has the form

\[
(1 \leq \forall k < \#x \mid x(k) \lt x(k+1)) = \text{true.}
\]

Eventually one might be able to write the ultimate program

\[
\text{Let there be light} = \text{true.}
\]


In these three papers as well as his book (cited earlier) Bishop speaks about writing a compiler for mathematics. The input to the compiler would be a mathematical proof of the existence of some object, and the output would be a program that constructs the object. This would of course only be feasible in the constructive framework where "proof" refers to a constructively valid argument.


From the book jacket: "During the past two decades, as the large-scale record systems of most government agencies and private organizations have been computerized, the American public has become concerned that something fundamentally dangerous might be happening."


A philosophical work which is pertinent to the design of software for the increasingly popular notion of the teaching machine. Chapter 2, The Purpose of Education, leads off with the following quote from H.A. Overstreet:

"...education is undergoing important transformations. From a relatively external process of pouring in facts, it is increasingly becoming a process of evoking the deeper, generative possibilities that lie within the individual."