PROGRAMS AS PROOFS

Robert L. Constable

November 1982

TR 82-532

Department of Computer Science
Cornell University
Ithaca, New York 14853
PROGRAMS AS PROOFS

Robert L. Constable
Cornell University

ABSTRACT

Programs are like constructive proofs of their specifications. This analogy is a precise equivalence for certain classes of programs. The connection between formal logic and programs is a foundation for programming methodology superior to that usually adopted. Moreover this equivalence suggests programming languages which are far richer than all others currently in use. These claims are established in this paper introducing parts of the PL/CV programming logics as a source of precision and examples.

KEY WORDS AND PHRASES

Algorithmic logic, automated logic, axiomatic semantics, constructive mathematics, program correctness, PL/CV, programming logic, programming methodology, realizability, while rule.

This work was supported in part by National Science Foundation Grant MCS-81-04018 and NSF-80-03349.
OUTLINE

1. Constructive Proofs
   1.1 what logicians have known
   1.2 compiling formal mathematics
   1.3 algorithms in mathematics

2. Algorithmic Proofs
   2.1 constructive logic
   2.2 proofs with conditionals
   2.3 proofs with assignments
   2.4 iteration and induction
   2.5 executing algorithmic proofs

3. Type Theoretic Proofs
   3.1 types
   3.2 propositions as types
   3.3 new programming languages
1. CONSTRUCTIVE PROOFS

1.1 what logicians have known

Ever since the 1940's work of Kleene [K45], logicians have known that various kinds of constructive proof could be "compiled" into executable code. For instance a proof of an assertion of the form "for all integers n there is an integer y such that R(n,y)" can be compiled into a computable function f such that R(n,f(n)). Results of this variety are known for many constructive formal systems [T73]. The essential ideas behind them can be grasped from an example. We choose a very simple benchmark common in the literature, the Euclidean division algorithm.

The theorem we want to prove is "for all non-negative integers a and b we can find non-negative integers q and r such that if b is positive, then a=b*q+r and 0≤r<b." To render this and other theorems more succintly we will adopt a conventional symbolic notation using &,
\( \lor, \neg, \Rightarrow \) for "and", "or", "not", and "implies" respectively and using \( \forall(x_1,\ldots,x_n):A \) and \( \exists(x_1,\ldots,x_n):A \) as the typed quantifiers "for all \( x_1,\ldots,x_n \) of type A" and "we can find \( x_1,\ldots,x_n \) of type A" respectively. As types we take \( N \) for the non-negative integers and \( Z \) for the integers. We use * to denote multiplication. The theorem now takes the following form.

Theorem (Euclidean Algorithm): \( \forall(a,b):N. \exists(q,r):N. (b \ge 0 \Rightarrow a=b*q+r \& 0 \le r < b) \)

This theorem is typically proved as a special case of the following lemma.

Lemma: \( \forall n:N. \forall(a,b):Z. (a-b < n \Rightarrow \exists(q,r):N. (a \ge 0 \& b > 0 \Rightarrow a=b*q+r \& 0 \le r < b)) \)

The lemma can be proved by induction on n, and the theorem follows by taking n to be a non-negative integer larger than a-b, say \( max(0, a-b)+1 \). The induction principle we have in mind can be stated as:
proof
for arbitrary n:N
assume P(n)

\[ \ldots \]

P(n + 1)

\[ \frac{P(0)}{\forall n:N. P(n)} \quad \text{qed} \]

This means that if we can prove P(0), and if we can prove P(n + 1) from the assumption that P(n), then P is true for all non-negative integers.

Here is an informal proof of the lemma.

Proof by induction:

Base case:

\[ a < b \text{ so take } q = 0, r = a, \text{ then} \]
\[ \text{clearly by arithmetic } a = b \cdot 0 + a \text{ and } 0 \leq a < b. \]

Induction case for arbitrary n:N:

1. Assume the result for any x - y < n, in order to show that for all (a, b):Z where a - b < n + 1 there exist appropriate q and r.

2. In order to apply the induction hypothesis we need inputs x, y such that x \geq 0, y > 0 and x - y < n. If we take (a - b) for x and b for y, then because b > 0 and (a - b) < n + 1 we know (a - b) - b < n by arithmetic.

3. In order to apply the induction hypothesis we need to know also that (a - b) \geq 0. We only have this if b \leq a, but when a < b we can proceed as in the base case. Here are the details.
(a-b)<0 or (a-b)\geq0 by arithmetic.
We prove the theorem in each case.
case (a-b)<0;
then a<b so take q=0, r=a.
case (a-b)\geq0;
then the induction hypothesis holds
with (a-b) for x and b for y,
so there are q', r' such that
(a-b) = b*q' + r' & 0 \leq r' < b.
a = b*(q' + 1) + r' by arithmetic.
Take q = q' + 1 and r = r' to finish
the proof in this case.

Qed

We can see the computational content of the proof rather clearly in terms of recursive procedures. When a-b<0, then the computation terminates by producing q=0 and r=a. If a-b\geq0, then it is possible to "call the lemma recursively" using (a-b) and b. These are smaller parameters, so we know that the sequence of "calls to the lemma" is finite. After the lemma is called, the result is q', r'. From these values line 5 of the proof shows how to compute the final values. The proof has a structure similar to the following Algol-like recursive procedure.

procedure lem(n, a, b, q, r)
declare n:N /* size of the inputs */;
(a, b):Z /* dividend and divisor inputs */;
(q, r):N /* quotient and remainder outputs */;
(q', r'):N /* auxiliary variables for q and r */;
assume a-b<n, a\geq0, b>0;
attain a=b*q + r & 0 \leq r < b;

if n=0 then (q:=0, r:=a)
else if (a-b)<0 then (q:=0, r:=a)
else lem(n-1, a-b, b, q', r')

/* (a-b) = b*q' + r' & 0 \leq r' < b */;
(q:=q' + 1, r:=r')
fi
end lem

Notice that the procedure will execute in cases such as n=3, a=25, b=3 for which the input assumptions cannot be satisfied at each call of lem. In such cases we simply cannot guarantee the program's behavior. The program will be correct for the division algorithm whenever n is greater than or equal to the quotient. (In many programming logics, including PL/CV, n would be part of the termination predicate and as such not an independent parameter, but the point of this procedure
is to illustrate the computational structure of the proof, not to stand on its own.)

We can at this point simply regard the assume and attain statements as comments about the behavior of the procedure.

One can see from this example the correspondence between an appeal to the induction hypothesis and a recursive call of the procedure. Likewise the correspondence between assignment, q:=0, and the choice of values for existentially quantified variables, "let q be 0", is clear.

1.2 compiling formal mathematics

The correspondences suggested by the above example can be made general, say to cover all proofs in a specific constructive theory as Kleene proved for Intuitionistic number theory [K45]. There are a number of subtle points that complicate the general result, for example translating nested implications of the form \( A \Rightarrow (B \Rightarrow (C \Rightarrow D)) \) involves higher order functions.

This correspondence suggests that the language of formal constructive proofs might be useful as a programming language. This idea was proposed by Constable [C71] and has been studied extensively by [MM73]. Recently J. Bates [B79] and C. Goad [G80] have taken further steps in this direction.

There are obvious advantages to using proofs as programs, namely one can be quite sure that they are "correct programs". A constructive proof by its nature must prove its conclusion. In programming terminology the conclusion is the problem specification. In this context the task of teaching certain aspects of programming is very similar to teaching methods of problem solving and proof. It is obvious that the methodological investigations of G. Polya [P48] and others are applicable as written. It is also obvious that certain techniques of formal logic are applicable, though many computer scientists have only recently realized this [D76, G81].

There are also obvious obstacles to using proofs as programs. In the first place for a large class of concurrent programs the correspondence is unclear. In the same vein for a large class of computational problems there is no adequate formalism, e.g., various real time control problems, various A.I. programs from chemical synthesis to expert systems. Moreover, even when we do possess a formal-
ism and a precise correspondence between proof and program, we still do not know how to compile these proofs into efficient code. This problem area is the subject of investigation at Cornell [B79, BC81] and Stanford [G80].

1.3 algorithms in mathematics

In some ways it may seem strange to treat proofs as algorithms especially since algorithms clearly have a place in mathematics which seems at first quite different from that of proofs. Algorithms are used to define functions. We would expect the "Euclidean division algorithm" to be a function \( d \) taking \( a \) and \( b \) to \( q \) and \( r \), i.e., \( d : N \times N^+ \to N \times N \) where \( N^+ = \{1,2,3,\ldots\} \). It would be customary to prove that the algorithm performed correctly, e.g., if \( d(a,b) = <q,r> \) and we write \( <q,r>.1 = q, <q,r>.2 = r \), then correctness means

\[
a = b \ast d(a,b).1 + d(a,b).2 \land 0 \leq d(a,b).2 < r.
\]

This is the viewpoint usually adopted in studies of program correctness [M74, dB80] or programming methodology [D76, AA79, G81]. In the case of procedural programs, people have in mind a function mapping states to states in such a way that if the precondition (a relation on states) is true, then the resulting function values satisfy the post condition.

We will see shortly that it makes sense to treat constructive proofs computationally because the meaning of a constructive proof is computational. We will develop these ideas in the next section.

2. ALGORITHMIC PROOFS

2.1 constructive logic

To constructively prove an assertion of the form \( A \Rightarrow B \) we must exhibit a method of transforming a proof of \( A \) into a proof of \( B \). The direct way to do this is to assume that we are given a proof of \( A \), say \( a \), and try to build a proof of \( B \) from it. The pattern of such a construction in ordinary logic is
proof
  assume A
  .
  .
  .
  B
qed

To constructively prove \( A \lor B \) we must prove either \( A \) or \( B \), and to prove \( A \land B \) we must prove both \( A \) and \( B \).

The canonical way to use \( A \lor B \) in a proof of \( C \) is to conduct a case analysis, showing that if \( A \) holds then \( C \) does and if \( B \) holds then \( C \) does. The usual pattern for this is

\[
C \text{ by cases, } A \lor B
proof
  case A
  .
  .
  .
  C
case B
  .
  .
  .
  C
qed
\]

Here is a proof using these rules

\[
(A \Rightarrow C \& B) \Rightarrow (A \lor B \Rightarrow B)
proof
  assume A \Rightarrow C \& B;
  A \lor B \Rightarrow B;
  proof
    assume A \lor B;
    B \text{ by cases, } A \lor B,
    proof
      case A;
      C \& B \text{ by applying } A \Rightarrow C \& B \text{ to } A;
      B
case B
    qed
  qed
\]

2.2 proofs with conditionals

This conventional proof style can easily be augmented by more familiar algorithmic constructs such as the conditional, if bexp then C else D fi. This is essentially a case analysis on the assertion
bexp = true ∨ bexp = false. It is used only when we know that the Boolean expression, bexp, evaluates to either true or false.

Boolean expressions are not the same as assertions. They are expressions which denote one of the two Boolean values, true and false. We use the following operations on Booleans.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>a•b</th>
<th>a + b</th>
<th>¬a</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>True</td>
<td>True</td>
<td>True</td>
<td>False</td>
<td></td>
</tr>
<tr>
<td>True</td>
<td>False</td>
<td>False</td>
<td>True</td>
<td>False</td>
<td></td>
</tr>
<tr>
<td>False</td>
<td>True</td>
<td>False</td>
<td>True</td>
<td>True</td>
<td></td>
</tr>
<tr>
<td>False</td>
<td>False</td>
<td>False</td>
<td>False</td>
<td>True</td>
<td></td>
</tr>
</tbody>
</table>

The atomic Boolean expressions that we consider are x=y for x and y of the same type and for = a decidable operation. Over the integers we can prove assertions as follows:

\[
x \neq 0 \Rightarrow x\times x > 0
\]
proof

\[
\begin{align*}
\text{assume } x \neq 0; \\
\text{if } x < 0 \text{ then } x\times x > 0 \text{ by arithmetic;} \\
\text{else } x > 0 \text{ by arithmetic, } x \geq 0, x \neq 0; \\
x\times x > 0 \text{ by arithmetic.}
\end{align*}
\]

fi

qed

An equivalent proof without the conditional rule is

\[
x \neq 0 \Rightarrow x\times x > 0
\]
proof

\[
\begin{align*}
\text{assume } x \neq 0; \\
x < 0 \lor x \geq 0 \text{ by arithmetic;} \\
x\times x > 0 \text{ by cases } x < 0 \lor x \geq 0,
\end{align*}
\]

proof

\[
\begin{align*}
\text{case } x < 0; \\
x\times x > 0 \text{ by arithmetic}
\end{align*}
\]

\[
\begin{align*}
\text{case } x \geq 0; \\
x > 0 \text{ by arithmetic, } x \geq 0, x \neq 0; \\
x\times x > 0 \text{ by arithmetic}
\end{align*}
\]

qed

qed

A conditional is not only a special form of a "by cases" proof block, but it can also be used to write new assertions such as
if bexp then A else B fi

which is equivalent to the assertion

\[(bexp \Rightarrow A) \& (\neg bexp \Rightarrow B)\]

In the case of programming languages like Algol 60 or 68, the conditional is used in both roles, but in a language like PL/I without conditional expressions, it is used only to augment the proof constructs.

2.3 proofs with assignments

A common feature of informal proofs are definitions of the form "let q:=0, r:=a". Assignment statements such as q:=0, r:=a clearly resemble these definitions. Indeed if assignments are used so that a variable is never redefined, then they can be treated exactly as definitions and their resulting equalities, i.e., from "let q:=0" we conclude q=0.

When variables are redefined, as in x:=x+1, then clearly we cannot conclude the equality statement x=x+1, but moreover the whole idea of a variable and its use in a proof must be re-examined. For example even after the assignment x:=0 there are new issues to face if x is being redefined because statements involving x may change their meaning; consider this proof segment:

\[a>0; \ b>0;\]
\[b^2>0 \ by \ arithmetic, \ b>0;\]
\[b:=0;\]
\[a+b^2>0 \ by \ arithmetic, \ a>0, \ b^2>0.\]

The assignment b:=0 clearly invalidates the claim that \(b^2>0\).

Although the introduction of assignment statements which redefine variables complicates the concept of a proof somewhat, it is possible to write a reasonably elegant set of proof rules that includes them as was done by Constable and O'Donnell [CO78]. The concept of assignment has clearly proved useful in the early development of efficient high level languages such as FORTRAN and Algol; it has even crept into languages with a basically applicative structure such as Interlisp, Franzlisp, ML, etc. But whether or not the assignment concept will remain in high level languages, it is now known from extensive work on programming methodology how to treat it rigorously, and it is known from [CO78] how to build it into a formal logical system without increasing the conceptual overhead for those parts of the system which do not use it.
2.4 iteration and induction

Typically the proof rules for iterative programming constructs, such as the while loop, are shown to be sound using mathematical induction. For example the rule of Hoare [H69]

\[
\{P \& b\} \ S \ {P}, \ P \& \neg b \Rightarrow Q \\
\{P\} \ \text{while} \ b \ \text{do} \ S \ \text{od} \ \{Q\}
\]

is justified by induction on the number of executions of the loop body \(S\).

If the while rule is written in the PL/CV style [CO78, CJE82] as

\[
P; \\
\text{while} \ b \ \text{do} \\
\quad \text{assume} \ P \\
\quad S \\
\quad P \\
\text{od}
\]

\[
P \ & \neg b
\]

then it can be seen as another form of induction principle. In fact comparing side by side the induction rule of 1.1 and this while rule reveals the similarity graphically.

**PEANO INDUCTION**

\[
\text{proof for arbitrary } n \\
\text{assume } Q(n) \\
\ldots \\
Q(n+1) \\
\text{qed}
\]

\[
\forall n \in \mathbb{N}. Q(n)
\]

**WHILE INDUCTION**

\[
Q; \\
\text{while} \ b \ \text{do} \\
\quad \text{assume} \ Q \\
\quad \ldots \\
\text{od}
\]

\[
Q \ & \neg b
\]

The base case for Peano induction is the proof of \(Q(0)\) whereas for while induction it is the proof of \(Q\) prior to the loop statement.

*While induction* is useful only in the presence of assignment statements which are used to redefine at least one of the variables occurring in the Boolean expression. But regardless of this dynamic interpretation of the loop, it can also be seen as another induction principle offering another way to prove an assertion such as \(Q \ & \neg b\). Given the computational interpretation however, one can see

\[+\text{ In PL/CV one can actually prove arbitrary assertions this way. See [CJE82] for the general rule.}\]
while induction as a proof method which is in some cases more efficient than Peano induction. Here is a proof of the Euclidean algorithm using while induction and the assignment rule.

```
divide: procedure (a, b, q, r)
declare (a, b):Z;
declare (q, r):N;

assume a \geq 0, b > 0;
attain a = b*q + r, 0 \leq r < b;

a = b*0 + a by arith;
r := a, q := 0;
\exists i:N. (i \geq 0 \land r \leq i) by intro, r;
while ((r-b) \geq 0) do
  assume a = b*q + r, r \geq 0;
  arbitrary i:N where r \leq i,
  \neg (r \leq 0) by arith, r-b \geq 0, +, b > 0;
  a = b*(q+1) + (r-b) by arith, a = b*q + r;
  r-b \leq i-1
  r := r-b, q := q+1
  \end
d
  r \leq b by arith, \neg(r-b \geq 0);
a = b*q + r, 0 \leq r < b
return
```

2.5 executing algorithmic proofs

One of the principal advantages of a proof system such as PL/CV which includes algorithmic constructs such as conditionals, assignments and iterative induction is that we already know how to efficiently execute a large class of the proofs. For example we recognize the proof at the end of (2.4) as the Euclidean division algorithm. Thus PL/CV is an example of a system of formal constructive mathematics whose proofs can be very efficiently executed, as efficiently as those in any high level programming language.

3. TYPE THEORETIC PROOFS

3.1 types

In addition to the base types we have mentioned, N, Z, we have been implicitly interested in all functions from one type to another. Given types A_1, ..., A_n and B, let us write the type of all functions from A_1, ..., A_n into B as prod(A_1, ..., A_n, B). We are also interested in n-tuples (or records) and we write the type of n tuples \langle a_1, ..., a_n \rangle with a_i from A_i as prod(A_1, ..., A_n). Both prod
and \( \text{prod} \) are similar to the Algol-68 types \( \text{proc} \) and \( \text{struct} \) respectively. We will also allow an Algol-like union, denoted \( \text{union}(A_1, \ldots, A_n) \). In this notation the divide example of 1. has type \( \text{prod}(N,Z,Z)\text{prod}(N,N) \).

When we examine the divide procedure as it is used in a constructive programming logic such as PL/CV, it makes sense to say that in addition to its numerical parameters \( n, a \) and \( b \) it has a parameter which is a proof of the input assumptions \( a \geq 0, b > 0 \). Such a proof is needed in a programming logic because without satisfying the input assumptions, the procedure cannot be invoked. The only way to satisfy these assumptions in the logic is to provide a formal proof. This notion of proof is very concrete. So it makes sense to think of the required formal proof as an input.

We could unify the notation and conception of the logic considerably if we allowed assertions to be types whose members were their proofs. Then we could say that the inputs to the divide procedure are of types \( N, Z, Z \) and \( a \geq 0 \& b > 0 \) respectively. The output could also include a proof of the output specification (variously called the post condition or the attain statement). Again in a programming logic this information must be present; indeed the particular asserted program builds a particular proof of the output specification. In the case of the divide procedure, the output would be two integers, \( q \) and \( r \), and a proof of the assertion \( a = b \times q + r \& 0 \leq r < b \).

In order to treat these assertions as types, we must allow the type to depend on values of the other parameters. For example the output type \( a = b \times q + r \& 0 \leq r < b \) is parameterized by its inputs, \( a, b \). In general if \( B(x) \) is a type parameterized by \( x \) of type \( A \), then we write \( \text{proc}(x:A)B(x) \) to denote those functions which on input \( a \) of type \( A \) produce a result of type \( B(a) \). This is the function space constructor of AUTOMATH [dBr70].

To describe the dependency between the input type \( a \geq 0 \& b > 0 \) and its parameters, \( a, b \), we can use a dependent product construct of the form \( \text{prod}(x:A, y:B(x)) \) which consists of those pairs \( <a,b> \) such that \( a \) is of type \( A \) and \( b \) is of type \( B(a) \). In general we allow \( \text{prod}(x:A, y:B(x), z:C(x,y)) \), etc. This dependent product construct is essentially the infinite union of Scott [S70] and Martin-Löf [ML75]. The input to the divide procedure can be expressed with this constructor as:

\[
\text{prod}(n:N, a:Z, b:Z, p:a \geq 0 \land b > 0).
\]
The kind of logic described so far is essentially PL/CV2, but the syntax is somewhat altered so that assertions do not appear explicitly as types.

### 3.2 propositions as types

The programming logic outlined above is more complex than is necessary. Building on the propositions-as-types principle [H80, dBr70] Per Martin-Lof described a comprehensive theory of types in which propositions can be identified with types. The identification is suggested by this correspondence and has been extensively explained elsewhere [ML75, 79, C81, C82, BC81].

<table>
<thead>
<tr>
<th>proposition</th>
<th>type</th>
</tr>
</thead>
<tbody>
<tr>
<td>A &amp; B</td>
<td>$\text{prod}(A,B)$</td>
</tr>
<tr>
<td>A ∨ B</td>
<td>$\text{union}(A,B)$</td>
</tr>
<tr>
<td>A $\Rightarrow$ B</td>
<td>$\text{proc}(A)B$</td>
</tr>
<tr>
<td>$\forall x:A. B(x)$</td>
<td>$\text{proc}(x:A)B(x)$</td>
</tr>
<tr>
<td>$\exists x:A. B(x)$</td>
<td>$\text{prod}(x:A,B(x))$</td>
</tr>
</tbody>
</table>

The union construction is as in Algol 68 [McG68].

### 3.3 new programming languages

The types $\text{proc}(x:A)B(x)$, and $\text{prod}(x:A,y:B(x))$, $\text{union}(A,B)$ clearly extend the usual Algol-like types $\text{proc}$, $\text{struct}$ and $\text{union}$. They in fact provide types not available in any other programming language. If a notion of recursive data types is also provided as in [GMW79, CZ81] then these types subsume those of nearly all other programming languages.

A programming logic for a theory of this kind [CZ81, C82, BC81] is so rich that it appears to be adequate to formalize all of constructive mathematics (see [ML79]). If one considers the problems of reasoning about numerical algorithms and other sequential computational problems arising in mathematics, it is difficult to see how a language which falls short of providing these types will ever be completely adequate. This understanding will not be apparent from a viewpoint which starts with a language like FORTRAN or Algol and proceeds to extend it on an *ad hoc* basis. It is revealed by considering the nature of programming problems and the complete context in which they are posed.
and solved. In this context, I believe that the view of a sequential program as an (efficiently) executable proof is more fertile than any other.

ACKNOWLEDGEMENTS

The programming logic PL/CV2 used to illustrate these points was designed and built by a team of my colleagues, Michael J. O'Donnell, Scott D. Johnson, Dean B. Krafft, Tat-Kung Chan, Daniel Zlatin, Carl D. Eichenlaub and Ryan Stansifer. To them and my colleagues on the PRL project, Joseph L. Bates and Alan Demers, I owe a great deal. I also want to thank Donette Isenbarger for carefully preparing the manuscript.

REFERENCES


[C82]: Constable, Robert L., "Intensional Analysis of Functions and Types", Computer Science Department, University of Edinburgh, June 1982.


