Polynomial Constants are Decidable *†

Markus Müller-Olm
Universität Dortmund, FB 4, LS V
44221 Dortmund, Germany
e-mail: `mmo@ls5.cs.uni-dortmund.de`

Helmut Seidl
Universität Trier, FB 4—Informatik
54286 Trier, Germany
e-mail: `seidl@uni-trier.de`

March 28, 2002

Abstract

Constant propagation aims at identifying expressions that always yield a unique constant value at run-time. It is well-known that constant propagation is undecidable for programs working on integers even if guards are ignored as in non-deterministic flow graphs. Here we show that *polynomial constants* are decidable in non-deterministic flow graphs. In polynomial constant propagation assignment statements that use the operators $+, -, *$ are interpreted exactly but all assignments that use other operators are conservatively interpreted as non-deterministic assignments.

We present a generic algorithm for constant propagation via a symbolic weakest precondition computation and show how this generic algorithm can be instantiated for polynomial constant propagation by exploiting techniques from computable ring theory.

1 Introduction

Constant propagation is one of the most widely used optimizations in practical optimizing compilers (cf. [1, 8, 14]). Its goal is to replace expressions that always yield a unique constant value at run-time by this value. This both speeds up execution and reduces code size. Even more importantly, it can enable powerful further transformations like elimination of dynamically unreachable branches.

In order to come to grips with fundamental computability problems one often abstracts guarded branching to non-deterministic branching in program analysis. But even this abstraction leaves constant propagation undecidable for programs working on integer variables. This has already been observed in the seventies independently by Hecht [8] and by Reif and Lewis [16]. We briefly

*This paper is a pre-print of material from the first author's habilitation thesis.

†The work was supported by the RTD project IST-1999-20527 "DAEDALUS" of the European FP5 programme.

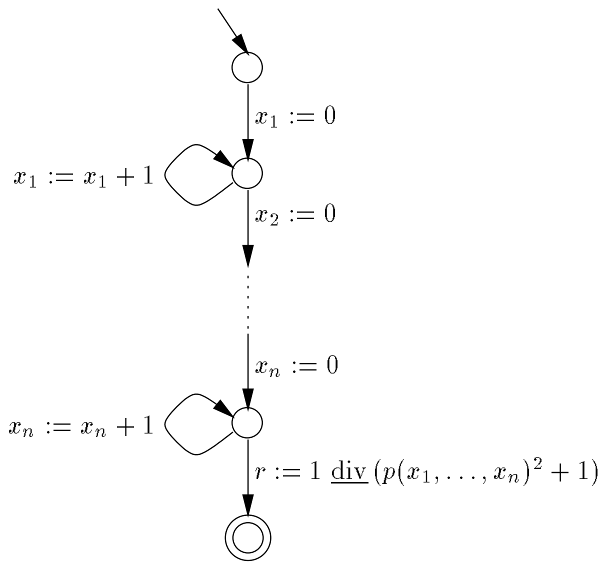


Figure 1: Undecidability of constant detection; the reduction of Reif and Lewis [16].

recall the construction of Reif and Lewis. It is based on a reduction of Hilbert's tenth problem, whether a (multi-variant) polynomial has a zero in the natural numbers, a very famous undecidable problem [11].

Assume given a (non-zero) polynomial $p(x_1, \ldots, x_n)$ in n variables $x_1, \ldots, x_n$ with natural coefficients and consider the (non-deterministic) program in Figure 1. The initializations and the loops choose arbitrary natural values for the variables x_i. If the chosen values constitute a zero of $p(x_1, \ldots, x_n)$, then $p(x_1, \ldots, x_n)^2 + 1 = 1$ and r is set to 1. Otherwise, $p(x_1, \ldots, x_n)^2 + 1 \geq 2$ such that r is set to 0. Therefore, r is a constant of value 0 at the end of the program if and only if $p(x_1, \ldots, x_n)$ does not have a natural zero.

On the other hand there are well-known and well-defined classes of constants that can be detected, even efficiently. In *copy constant detection* [6] only assignments of the form $x := c$, where c is either an (integer) constant or a program variable are interpreted; assignments with composite expressions on the right hand side are conservatively assumed to make a variable non-constant. In *linear constants* [17] also assignments of the form $x := a \cdot y + b$, where a and b are integer constants and y is a program variable, are interpreted. Another decidable class of constants are *finite constants* [18]. This motivated Müller-Olm and Rüthing [15] to examine the borderline of intractability and undecidability more closely.

They studied the complexity of constant propagation for classes that derive from restricting the signature of interpreted integer operators. An interesting question that has been left open in their study concerns the class of constants obtained from restricting the interpreted signature to the operators $+, -, *$, i.e., all standard integer operators except of the division operators. While they called the corresponding class of constants $+, -, *$-constants, we prefer the apter term *polynomial constants* here, as these signature allows just to write (multi-variant) polynomials. A deeper reason for this name will become clear soon. The detection problem for polynomial constants is PSPACE-hard [15] but no upper bound is known. In the current paper we show that polynomial constants are decidable by applying results from computable ideal theory. This decidability result identifies the division operator as the real cause for undecidability of general constant detection.

This paper is organized as follows. In the next section we define flow graphs and introduce constant propagation and weakened versions thereof. This enables us to define polynomial constants. In Section 3 we present a generic algorithmic idea for constant propagation via a symbolic weakest precondition computation and discuss the requirements for making it effective. In Section 4 we recall some results from computable ring theory that are needed in our algorithm for detection of polynomial constants and proceed with further results and observations on $\mathbb{Z}[x_1, \ldots, x_n]$, the set of multi-variant polynomial with integer coefficients. These results are exploited in Section 6, where we show how the generic algorithm from Section 3 can be instantiated for polynomial constant propagation. We finish the paper with concluding remarks and directions for future research.

2 The Framework

2.1 Flow Graphs

Suppose given a finite set of variables $X = \{x_1, \ldots, x_n\}$. Let Expr be a set of expressions over X; the precise nature of expressions is immaterial at the moment. A (deterministic) assignment is a pair consisting of a variable and an expressions written in the form $x := t$; the set of assignment statements is denoted by Asg. A non-deterministic assignment statement consists of a variable and is written $x :=?$; the set of nondeterministic assignment statements is denoted by NAsg.

A (non-deterministic) flow graph is a structure $G = (N, E, A, \mathsf{s}, \mathsf{e})$ with finite node set N, edge set $E \subseteq N \times N$, a unique start node $\mathsf{s} \in N$, and a unique end node $\mathsf{e} \in N$. We assume that each program point $u \in N$ lies on a path from s to e. The mapping $A : E \to \mathsf{Asg} \cup \mathsf{NAsg} \cup \{\mathsf{skip}\}$ associates each edge with a deterministic or non-deterministic assignment statement or the statement skip. Edges represent the branching structure and the statements of a program, while nodes represent program points. The set of successors of program point $u \in N$ is $Succ[u] = \{v \mid (u, v) \in E\}$.

A *path* reaching a given program point $u \in N$ is a sequence of edges $p = \langle e_1, \ldots, e_k \rangle$ with $e_i = (u_i, v_i) \in E$ such that $u_1 = \mathsf{s}$, $v_k = u$, and $v_i = u_{i+1}$ for $1 \leq i < k$. In addition $p = \varepsilon$, the empty sequence, is a path reaching the start node s. We write $\mathsf{R}[u]$ for the set of paths reaching u.

Let Val be a set of values. A mapping $\sigma : X \to \mathsf{Val}$ that assigns a value to each variable is called a *state*; we write $\Sigma = \{\sigma \mid \sigma : X \to \mathsf{Val}\}$ for the set of states. For $x \in X$, $d \in \mathsf{Val}$, and $\sigma \in \Sigma$ we write $\sigma[x \mapsto d]$ for the state that maps x to d and coincides for the other variables with σ. We assume a fixed interpretation for the operators used in terms and we assume that the value of term t in state σ, which we denote by t^σ, is defined in the standard way.

In order to accommodate non-deterministic assignments we interpret statements by relations on Σ rather than functions. The relation associated with assignment statement $x := t$ is $[\![x := t]\!] \stackrel{\text{def}}{=} \{(\sigma, \sigma') \mid \sigma' = \sigma[x \mapsto t^\sigma]\}$; the relation associated with non-deterministic assignment $x :=?$ is $[\![x :=?]\!] \stackrel{\text{def}}{=} \{(\sigma, \sigma') \mid \exists d \in \mathsf{Val} : \sigma' = \sigma[x \mapsto d]\}$; and the relation associated with skip is the identity: $[\![\mathsf{skip}]\!] \stackrel{\text{def}}{=} \{(\sigma, \sigma') \mid \sigma = \sigma'\}$. This local interpretation of statements is straightforwardly extended to paths $p = \langle e_1, \ldots, e_k \rangle \in E^*$ as follows: $[\![p]\!] = [\![A(e_1)]\!] ; \ldots ; [\![A(e_k)]\!]$, where ; denotes relational composition.

2.2 Constant Propagation

A variable $x \in X$ is a *constant* at program point $u \in N$ if

$$\exists d \in \mathsf{Val}\ \forall p \in \mathsf{R}[u]\ \forall \sigma_0, \sigma \in \Sigma : \ (\sigma_0, \sigma) \in [\![p]\!] \ \Rightarrow\ \sigma(x) = d\,.$$

The universal quantification over initial states σ_0 expresses that we do not know the initial state—the state in which the program is started—which represents the input to the program.

We can weaken the demands for a constant detection algorithm by selecting a certain subset of expressions $S \subseteq \mathsf{Expr}$ that are interpreted precisely and assuming conservatively that assignments whose right hand does not belong to S assign an arbitrary values to their respective target variable. This idea can be made formal as follows.

For a given flow graph $G = (N, E, A, \mathsf{s}, \mathbf{e})$ and subset of expressions $S \subseteq \mathsf{Expr}$, let $G_S = (N, E, A_S, \mathsf{s}, \mathbf{e})$ be the flow graph with the same underlying graph but with the following weakened edge annotation:

$$A_S(e) = \begin{cases} x :=?\,, & \text{if } A(e) = (x := t) \text{ and } t \notin \mathsf{Expr} \\ A(e)\,, & \text{otherwise} \end{cases}$$

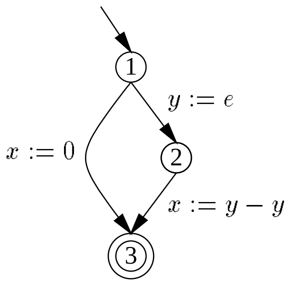
A variable $x \in X$ is then called an *S-constant* at program point $u \in N$ in flow graph G if it is a constant at u in the weakened flow graph G_S. Clearly, if x is an S-constant at u it is also a constant at u but not vice versa. The *detection problem for S-constants* is the problem of deciding for a given set of variables X, flow graph G, variable x, and program point u whether x is an S-constant at u.

To study weakened versions of constant detection problems is particularly interesting for programs computing on the integers, i.e., if Expr is the set of integer expressions formed from integer constants and variables with the standard operators $+, -, *, \underline{\text{div}}, \underline{\text{mod}}$: we have seen in the introduction that the general constant detection problem is undecidable in this case.

Let us discuss some examples for illustration. S-constants with respect to the set $S = X \cup \mathbb{Z}$, i.e., the set of non-composite expressions, are known as *copy constants* [6]. S-constants with respect to the set $S = \{a{*}x{+}b \mid a, b \in \mathbb{Z}, x \in X\}$ are known as *linear constants* [17]. In this paper we tackle constants with respect to the set $S = \mathbb{Z}[x_1, \ldots, x_n]$, the set of multi-variant polynomials in the variables $x_1, \ldots, x_n$ with coefficients in $\mathbb{Z}$, which we call *polynomial constants*. Polynomial constants correspond to $+, -, *$-must-constants in the classification of Müller-Olm and Rüthing [15] because this signature allows just to write multi-variant polynomials.

We should emphasize two points about the above framework that make the construction of S-constant detection algorithms more challenging. Firstly, in contrast to the setup in [15], we allow assignment statements, whose right hand side does not belong to S. They are interpreted as non-deterministic assignments. Forbidding them is adequate for studying lower complexity bounds for analysis questions, which is the main concern of [15]. It is less adequate when we are concerned with detection algorithms because in practice we want to detect S-constants in the context of other code.

Secondly, a variable can be an S-constant although its value statically depends on an expression that is not in S. As a simple example consider the following flow graph and assume that the expressions 0 and $y - y$ belong to S but e does not:



Because $y - y$ is 0 for any value $y \in \mathbb{Z}$, an S-constant detection algorithm must identify x as a constant (of value 0) at program point 3, although the value of x at program point 3 statically depends on the uninterpreted expression e.

3 A Generic Algorithm

Suppose given a variable $x \in X$ and a program point $w \in N$. In this chapter we describe a generic algorithm for deciding whether x is an S-constant at w. While standard constant propagation works by forward propagation of variable assignments, we use a three phase algorithm that employs a backwards propagation of assertions:

Phase 1: In the first phase we follow an arbitrary cycle-free path leading to w, for instance using depth-first search, and compute the value d, referred to as the *candidate value*, of x on this path. This implies that, if x is a constant at w, it must be a constant of value d.

Phase 2: In the second phase we compute the weakest precondition for the assertion $x = d$ at program point w in G_S by means of a backwards data flow analysis.

Phase 3: Finally, we check whether the computed weakest precondition for $x = d$ at w is true, i.e., is valid for all states.

It is obvious that this algorithm is correct. The problem is that Phase 2 and 3 are in general not effective. However, as only assignments of a restricted form appear in G_S, the algorithm becomes effective for certain sets S, if assertions are represented appropriately. In this way Müller-Olm and Rüthing [15] constructed a polynomial-time algorithm that detects constants in flow graphs in which right-hand side expressions of assignments are constructed only with the integer operators $+$ and $-$. In their algorithm assertions are represented by affine vector spaces over $\mathbb{Q}$, which suffices for such flow graphs. In the remainder of this section we analyze the requirements for an adequate space representing assertions. For this purpose, we first characterize weakest preconditions in flow graphs.

Semantically, an *assertion* is just a subset of states $\phi \subseteq \Sigma$. Given an assertion ϕ and a statement s, the *weakest precondition* of s for ϕ, $\mathsf{wp}(s)(\phi)$, is the largest assertion ϕ' such that execution of s from all states in ϕ' is guaranteed to end in a state in ϕ. The following identities for the weakest precondition of assignment and skip statements are well-known:

$$\begin{aligned} \mathsf{wp}(x := e)(\phi) &\stackrel{\text{def}}{=} \phi[e/x] \stackrel{\text{def}}{=} \{\sigma \mid \sigma[x \mapsto e^\sigma] \in \phi\} \\ \mathsf{wp}(x :=?)(\phi) &\stackrel{\text{def}}{=} \forall x(\phi) \stackrel{\text{def}}{=} \{\sigma \mid \forall d \in \mathbb{Z} : \sigma[x \mapsto d] \in \phi\} \\ \mathsf{wp}(\mathsf{skip})(\phi) &\stackrel{\text{def}}{=} \phi \end{aligned}$$

These identities characterize weakest preconditions of basic statements. Let us now consider the following more general situation in a given flow graph $G = (N, E, A, \mathbf{s}, \mathbf{e})$: we are given an assertion $\phi \subseteq \Sigma$ as well as a program point $w \in N$. Let $W_0[w] = \phi$ and $W_0[u] = \Sigma$ and consider the following equation system consisting of one equation for each program point $u \in N$:

$$\mathbf{W}[u] \;=\; W_0[u] \cap \bigcap_{v \in Succ[u]} \mathsf{wp}(A(u,v))(\mathbf{W}[v])\,. \tag{1}$$

By the Knaster-Tarski fixpoint theorem, this equation system has a largest solution (w.r.t. subset inclusion) because $\mathsf{wp}(s)$ is well-known to be monotonic. By abuse of notation, we denote the weakest solution by the same letter $\mathbf{W}[u]$. For each program point $u \in N$, $\mathbf{W}[u]$ is the weakest assertion such that execution starting from u with any state in $\mathbf{W}[u]$ guarantees that ϕ holds whenever execution reaches w. In particular, $\mathbf{W}[\mathsf{s}]$ is the weakest precondition for validity of ϕ at w. The intuition underlying equation (1) is the following: firstly, $W_0[u]$ must be implied by $W[u]$ and, secondly, for all successors v, we must guarantee that their associated condition $\mathbf{W}[u]$ is valid after execution of the statement $A(u,v)$ associated with the edge (u,v); hence $\mathsf{wp}(A(u,v))(\mathbf{W}[v])$ must be valid at u.

For two reasons, the above equation system cannot be solved directly in general: firstly, because assertions may be infinite sets of states they cannot immediately be represented; secondly, 2^Σ has infinitely long descending chains such that standard fixpoint iteration does not terminate in general.

In order to construct an algorithm that detects S-constants we work with representations of assertions. Let us assume that we have chosen a lattice $(R, \sqsubseteq)$

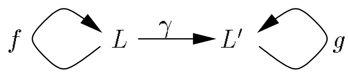


Figure 2: The situation in the transfer lemma

of representations for assertions and that $\gamma : R \to 2^{\Sigma}$ describes how the lattice element represent assertions. First of all, we require

(a) R has no infinite decreasing chains, i.e., there is no infinite chain $r_1 \sqsupseteq r_2 \sqsupseteq r_3 \sqsupseteq \ldots$.

This guarantees that maximal fixpoints of monotonic functions can effectively be computed by standard fixpoint iteration. Secondly, we suppose

(b) γ is universally conjunctive, i.e., $\gamma(\sqcap X) = \bigcap\{\gamma(r) \mid r \in X\}$ for all $X \subseteq R$.

The most important reason for making this assumption is that it ensures that we can validly compute on representations without loosing precision: if we precisely mirror the equations characterizing weakest preconditions on representations, the largest solution of the resulting equation system on representations characterizes the representation of the weakest precondition by the following well-known lemma. It appears (for the dual situation of least fixpoints) under the name 'Transfer Lemma' in [2] and under the name 'μ-Fusion Rule' in [10].

Lemma 1 *Suppose L, L' are complete lattices, $f : L \to L$ and $g : L' \to L'$ are monotonic functions and $\gamma : L \to L'$ (Fig. 2).*

If γ is universally conjunctive and $\gamma \circ f = g \circ \gamma$ then $\gamma(\nu f) = \nu g$, where νf and νg are the largest fixpoints of f and g, respectively.

Other important consequences of (b) will become clear soon.

We must mirror the elements comprising the equation system characterizing weakest preconditions on representations precisely. Firstly, we need to represent the start value, W_0. Universal conjunctivity of γ implies that $\gamma(\top) = \Sigma$, i.e., the top value of R is a precise representation of Σ. In addition, we require:

(c) Assertion $x = d$ is representable: for each $x \in X$, $d \in \mathsf{Val}$ there is $r_{x=d} \in R$ with $\gamma(r_{x=d}) = \{\sigma \in \Sigma \mid \sigma(x) = d\}$.

Secondly, we need effective representations for the operators appearing in equations. Requirement (b) implies that the meet operation of R precisely abstracts intersection of assertions. In order to enable effective computation of intersections, we need in addition:

(d) for given $r, r' \in R$, we can effectively compute $r \sqcap r'$.

By induction this implies that we can compute finite meets $r_1 \sqcap \ldots \sqcap r_k$ effectively.

The only remaining operations on assertions are the weakest precondition transformers of basic statements. We must represent $\mathsf{wp}(x := t)$, which is the substitution operator $(\cdot)[t/x]$ on assertions. However, it must be represented only for expressions $t \in S$, as the S-constant detection algorithm computes the weakest precondition in weakened flow graph G_S, where assignments with other right hand side are replaced by non-deterministic assignments.

(e) There is a computable substitution operation $(\cdot)[t/x] : R \to R$ for each $x \in X$, $t \in S$, which satisfies $\gamma(r[t/x]) = \gamma(r)[t/x]$ for all $r \in R$.

Obviously, $\mathsf{wp}(\mathsf{skip})$, the identity, is precisely represented by the identity on R. It remains to represent $\mathsf{wp}(x :=?)$:

(f) There is a computable projection operation $proj_i : R \to R$ for each variable $x_i \in X$ such that $\gamma(proj_i(r)) = \forall x_i(\gamma(r))$ for all $r \in R$.

Finally, we need the following, in order to make Phase 3 of the algorithm effective.

(g) Assertion true is decidable, i.e., there is a decision procedure that decides for a given $r \in R$, whether $\gamma(r) = \Sigma$ or not.

If, for a given set $S \subseteq \mathsf{Expr}$, we can find a lattice of representations satisfying requirements (a)–(g), we can effectively execute the three phase algorithm from the beginning of this section by representing assertions by elements from this lattice. This results in a detection algorithm for S-constants.

In this paper we are interested in detection of polynomial constants. Thus, from now on, let $\mathsf{Val} = \mathbb{Z}$ and $S = \mathbb{Z}[x_1, \dots, x_n]$. They key idea for the detection of polynomial constants is to represent assertions by the zeros of ideals in the polynomial ring $\mathbb{Z}[x_1, \dots, x_n]$ and to apply techniques from computable ideal theory. A full introduction to this area is beyond the scope of this paper but we recall the facts needed in the next section and make some additional observations in Section 5. Accessible introductions can be found in standard textbooks on computer algebra. The case of polynomial rings over fields is covered, e.g., by [5, 7, 19], while [13] treats the more general case of polynomial rings over rings, that is of relevance here, as $\mathbb{Z}$ is an integral domain but not a field.

4 A Primer on Computable Ideal Theory

Let us first recall some notions from ring theory. Recall that $\mathbb{Z}$ together with addition and multiplication forms a commutative ring, i.e., a structure $(R, +, \cdot)$ with a non-empty set R and two inner operations $+$ and $\cdot$ such that $(R, +)$ is an Abelian group, $\cdot$ is associative and commutative, and the distributive law $a \cdot (b + c) = a \cdot b + a \cdot c$ is valid for all $a, b, c \in R$. On the set of polynomials, $\mathbb{Z}[x_1, \dots, x_n]$, we can define addition and multiplication operations in the standard way; this makes $\mathbb{Z}[x_1, \dots, x_n]$ a commutative ring as well.

A non-empty subset $I \subseteq R$ of a ring R is called an *ideal* if $a + b \in I$ and $r \cdot a \in I$ for all $a, b \in I$, $r \in R$. The ideal *generated* by a subset $B \subseteq R$ is

$$(B) = \{r_1 \cdot b_1 + \ldots + r_k \cdot b_k \mid r_1, \dots, r_k \in R, b_1, \dots, b_k \in B\},$$

and B is called a *basis* or *generating system* of I if $I = (B)$. An ideal is called *finitely generated* if it has a finite basis $B = \{b_1, \dots, b_m\}$. Hilbert's famous basis theorem tells us that $\mathbb{Z}[x_1, \ldots, x_n]$ is *Noetherian*, since $\mathbb{Z}$ is Noetherian, i.e., that there are no infinitely long strictly increasing chains $I_1 \subset I_2 \subset I_3 \subset \ldots$ of ideals in $\mathbb{Z}[x_1, \ldots, x_n]$. This implies that every ideal of $\mathbb{Z}[x_1, \ldots, x_n]$ is finitely generated.

It is important for effectivity of our algorithm that we can compute effectively with ideals. While Hilbert's basis theorem ensures that we can represent every ideal of $\mathbb{Z}[x_1, \ldots, x_n]$ by a finite basis, in itself it does not give effective procedures for basic questions like membership tests or equality tests of ideals represented in this way. Indeed, Hilbert's proof of the basis theorem was famous (and controversial) at its time for its non-constructive nature.

Fortunately, the theory of Gröbner bases and the Buchberger algorithm provides a solution for some of these problems. While a complete presentation of this theory is way beyond the scope of this paper—the interested reader is pointed to the books mentioned above—a few sentences are in order here. A Gröbner basis is a basis for an ideal that has particularly nice properties. It can effectively be computed from any given finite basis by the Buchberger algorithm. There is a natural notion of reduction of a polynomial with respect to a set of polynomials. Reduction of a polynomial p with respect to a Gröbner basis always terminates and yields a unique result. This result is the zero polynomial if and only if p belongs to the ideal represented by the Gröbner basis. Hence reduction with respect to a Gröbner basis yields an effective membership test, that in turn can be used to check equality and inclusion of ideals.

In the terminology of [13], $\mathbb{Z}[x_1, \ldots, x_n]$ is a *strongly computable ring*. This implies that the following operations are computable for ideals $I, I' \subseteq \mathbb{Z}[x_1, \ldots, x_n]$ given by finite bases B, B' and polynomials $p \in \mathbb{Z}[x_1, \ldots, x_n]$, cf. [13]:

Ideal membership: Given an ideal I and a polynomial p. Is $p \in I$?

Ideal inclusion: Given two ideals I, I'. Is $I \subseteq I'$?

Ideal equality: Given two ideals I, I'. Is $I = I'$?

Sum of ideals: Given two ideals I, I'. Compute a basis for $I + I' \stackrel{\text{def}}{=} \{p + p' \mid p \in I, p' \in I'\}$. As a matter of fact, $I + I' = (B \cup B')$.

Intersection of ideals: Given two ideals I, I'. Compute a basis for $I \cap I'$.

It is straightforward (and well-known) that $I + I'$ and $I \cap I'$ are again ideals if I and I' are. We can use the above operations as basic operations in our algorithms.

5 More About $\mathbb{Z}[x_1, \ldots, x_n]$

5.1 $\mathbb{Z}[x_1, \ldots, x_n]$ as a Complete Lattice

Interestingly, the ideals in $\mathbb{Z}[x_1, \ldots, x_n]$ form also a complete lattice under subset inclusion $\subseteq$. Suppose given a set $\mathcal{I}$ of ideals in $\mathbb{Z}[x_1, \ldots, x_n]$. Then the largest ideal contained in all ideals in $\mathcal{I}$ obviously is $\bigcap \mathcal{I}$, and the smallest ideal that contains all ideals in $\mathcal{I}$ is $\sum\mathcal{I} := \{r_1 \cdot a_1 + \ldots + r_k \cdot a_k \mid r_1, \ldots, r_k \in \mathbb{Z}[x_1, \ldots, x_n], a_1, \ldots, a_k \in \bigcup\mathcal{I}\}$. The least element of the lattice is the zero ideal $\{0\}$ that consists only of the zero polynomial and the largest element is $\mathbb{Z}[x_1, \ldots, x_n]$. While this lattice does not have finite height it is Noetherian by Hilbert's basis theorem such that we can effectively compute least fixpoints of monotonic functions on ideals of $\mathbb{Z}[x_1, \ldots, x_n]$ by standard fixpoint iteration.

5.2 Zeros

As mentioned, we represent assertions by the zeros of ideals in our algorithm. A state σ is called a *zero* of polynomial p if $p^\sigma = 0$; we denote the set of zeros of polynomial p by $\mathcal{Z}(p)$. More generally, for a subsets $B \subseteq \mathbb{Z}[x_1, \ldots, x_n]$, $\mathcal{Z}(B) = \{\sigma \mid \forall p \in B : p^\sigma = 0\}$. Let us for later use state some facts concerning zeros, in particular of the relationship of ideal operations with operations on their zeros.

Lemma 2 *Suppose B, B' are sets of polynomials, q is a polynomial, I, I' are ideals, and $\mathcal{I}$ is a set of ideals in $\mathbb{Z}[x_1, \ldots, x_n]$.*

1. *If $B \subseteq B'$ then $\mathcal{Z}(B) \supseteq \mathcal{Z}(B')$.*
2. *$\mathcal{Z}(B) = \mathcal{Z}((B)) = \bigcap_{p \in B} \mathcal{Z}(p)$. In particular, $\mathcal{Z}(q) = \mathcal{Z}((q))$.*
3. *$\mathcal{Z}(\sum\mathcal{I}) = \bigcap\{\mathcal{Z}(I) \mid I \in \mathcal{I}\}$. In particular, $\mathcal{Z}(I + I') = \mathcal{Z}(I) \cap \mathcal{Z}(I')$.*
4. *$\mathcal{Z}(\bigcap\mathcal{I}) = \bigcup\{\mathcal{Z}(I) \mid I \in \mathcal{I}\}$. In particular, $\mathcal{Z}(I \cap I') = \mathcal{Z}(I) \cup \mathcal{Z}(I')$.*
5. *$\mathcal{Z}(\{0\}) = \Sigma$ and $\mathcal{Z}(\mathbb{Z}[x_1, \ldots, x_n]) = \emptyset$.*
6. *$\mathcal{Z}(I) = \Sigma$ if and only if $I = \{0\} = (0)$.*

5.3 Substitution

Suppose given a polynomial $p \in \mathbb{Z}[x_1, \ldots, x_n]$ and a variable $x \in X$. We can define a substitution operation on ideals I as follows: $I[p/x] = (\{q[p/x] \mid q \in I\})$, where the substitution of polynomial p for x in q, $q[p/x]$, is defined as usual. By definition, $I[p/x]$ is the smallest ideal that contains all polynomials $q[p/x]$ with $q \in I$. From a basis for I, a basis for $I[p/x]$ is obtained in the expected way: if $I = (B)$, then $I[p/x] = (\{b[p/x] \mid b \in B\})$. Thus, we can easily obtain a finite basis for $I[p/x]$ from a finite basis for I: if $I = (b_1, \ldots, b_k)$ then $I[p/x] = (b_1[p/x], \ldots, b_k[p/x])$. Hence we can add substitution to our list of computable operations.

The substitution operation on ideals defined in the previous paragraph mirrors precisely semantic substitution in assertions which has been defined in connection with $\mathsf{wp}(x := e)$.

Lemma 3 $\mathcal{Z}(I)[p/x] = \mathcal{Z}(I[p/x])$.

We leave the proof of this equation that involves the substitution lemma known from logic to the reader.

5.4 Projection

In this section we define projection operators $proj_i$, $i = 1, \ldots, n$, such that for each ideal I, $\mathcal{Z}(proj_i(I)) = \forall x_i(\mathcal{Z}(I))$. Semantic universal quantification over assertions has been defined in connection with $\mathsf{wp}(x :=?)$.

A polynomial $p \in \mathbb{Z}[x_1, \ldots, x_n]$ can uniquely be written as a polynomial in x_i with coefficients in $\mathbb{Z}[x_1, \ldots, x_{i-1}, x_{i+1}, x_n]$, i.e., in the form $p = c_k x_i^k + \ldots + c_0 x_i^0$, where $c_0, \ldots, c_k \in \mathbb{Z}[x_1, \ldots, x_{i-1}, x_{i+1}, x_n]$, and $c_k \neq 0$ if $k > 0$. We call $c_0, \ldots, c_k$ the coefficients of p with respect to x_i and let $\mathcal{C}_i(p) = \{c_0, \ldots, c_k\}$.

Lemma 4 $\forall x_i(\mathcal{Z}(p)) = \mathcal{Z}(\mathcal{C}_i(p))$.

Proof. Let $p = c_k x_i^k + \ldots + c_0 x_i^0$ with $\mathcal{C}_i(p) = \{c_0, \ldots, c_k\}$.
'$\supseteq$': Let $\sigma \in \mathcal{Z}(\mathcal{C}_i(p))$. We have $c_k^{\sigma[x_i \mapsto d]} = c_k^\sigma = 0$ for all $d \in \mathbb{Z}$ because c_k is independent of x_i. Hence, $p^{\sigma[x_i \mapsto d]} = c_k^{\sigma[x_i \mapsto d]} d^k + \ldots + c_0^{\sigma[x_i \mapsto d]} d^0 = 0d^k + \ldots + 0d^0 = 0$ for all $d \in \mathbb{Z}$, i.e. $\sigma \in \forall x_i(\mathcal{Z}(p))$.
'$\subseteq$': Let $\sigma \in \forall x_i(\mathcal{Z}(p))$. We have $c_k^{\sigma[x_i \mapsto d]} = c_k^\sigma$ for all $d \in \mathbb{Z}$ because c_k is independent of x_i. Therefore, $c_k^\sigma d^k + \ldots + c_0^\sigma d^0 = c_k^{\sigma[x_i \mapsto d]} d^k + \ldots + c_0^{\sigma[x_i \mapsto d]} d^0 = p^{\sigma[x \mapsto d]} = 0$ for all $d \in \mathbb{Z}$ because of $\sigma \in \forall x_i(\mathcal{Z}(p))$. This means that the polynomial $c_k^\sigma x_i^k + \ldots + c_0^\sigma x_i^0$ vanishes for all values of x_i. Hence, it has more than k zeros which implies that it is the zero polynomial. Consequently, $c_j^\sigma = 0$ for all $j = 0, \ldots, k$, i.e., $\sigma \in \mathcal{Z}(\mathcal{C}_i(p))$. $\square$

Suppose $I \subseteq \mathbb{Z}[x_1, \ldots, x_n]$ is an ideal with basis B.

Lemma 5 $\forall x_i(\mathcal{Z}(I)) = \mathcal{Z}(\bigcup_{f \in B} \mathcal{C}_i(f))$.

Proof. $\forall x_i(\mathcal{Z}(I)) = \forall x_i(\mathcal{Z}(B)) = \forall x_i(\bigcap_{p \in B} \mathcal{Z}(p)) = \bigcap_{p \in B} \forall x_i(\mathcal{Z}(p)) = \bigcap_{p \in B} \mathcal{Z}(\mathcal{C}_i(p)) = \mathcal{Z}(\bigcup_{p \in B} \mathcal{C}_i(p))$. $\square$

In view of this formula, it is natural to define $proj_i(I) = (\bigcup_{p \in B} \mathcal{C}_i(p))$ where B is a basis of I. It is not hard to show that this definition is independent of the basis; we leave this proof to the reader. Obviously, $proj_i$ is effective: if I is given by a finite basis $\{b_1, \ldots, b_k\}$ then $proj_i(I)$ is given by the finite basis $\bigcup_{j=1}^k \mathcal{C}_i(b_j)$.

Corollary 1 $\forall x_i(\mathcal{Z}(I)) = \mathcal{Z}(proj_i(I))$.

Proof. $\forall x_i(\mathcal{Z}(I)) = \mathcal{Z}(\bigcup_{p \in B} \mathcal{C}_i(p)) = \mathcal{Z}((\bigcup_{p \in B} \mathcal{C}_i(p))) = \mathcal{Z}(proj_i(I))$. $\square$

6 Detection of Polynomial Constants

We represent assertions by ideals of the polynomial ring $\mathbb{Z}[x_1, \ldots, x_n]$ in the detection algorithm for polynomial constants. Thus, let R be the set of ideals of $Z[x_1, \ldots, x_n]$) and $\sqsubseteq$ be $\supseteq$. The representation mapping is $\gamma = \mathcal{Z}$. Note that the order is *reverse* inclusion of ideals. This is because larger ideals have smaller set of zeros. Thus, the *meet* operation is the *sum* operation of ideals and the top element is the ideal $\{0\} = (0)$.

In a practical algorithm, ideals are themselves represented by finite bases. For transparency, we suppress this further representation step in the description that follows but ensure that all used operations on ideals can effectively be computed on bases.

The lattice $(R, \supseteq)$ satisfies requirements (a)–(g) of Section 3:

(a) This requirement is satisfied because $\mathbb{Z}[x_1, \ldots, x_n]$ is Noetherian.

(b) The identity $\mathcal{Z}(\sum \mathcal{I}) = \bigcap\{\mathcal{Z}(I) \mid I \in \mathcal{I}\}$ means that $\mathcal{Z}$ is universally conjunctive.

(c) Suppose $x \in X$ and $d \in \mathbb{Z}$. Certainly, a state is a zero of the ideal generated by the polynomial $x - d$ if and only if it maps x to d. Hence, we choose $r_{x=d}$ as the ideal $(x - d)$ generated by $x - d$.

(d) In Section 4 we have seen that the sum of two ideals can effectively be computed on bases.

(e) We have seen in Section 5.3 that $(\cdot)[p/x]$ is an adequate and computable substitution operator.

(f) From Section 5.4 we know that $proj_i$ is an adequate and computable projection operator.

(g) We know that $\mathcal{Z}(I) = \Sigma$ if and only if $I = \{0\}$. Moreover, the only basis of the ideal $\{0\}$ is $\{0\}$ itself. Hence, in order to decide whether an ideal I given by a basis B represents Σ, we need only check whether $B = \{0\}$.

We can thus apply the generic algorithm from Section 3 for the detection of polynomial constants. In order to make this more specific, we put the pieces together, and describe the resulting algorithm in more detail.

Suppose given a variable $x \in X$ and a program point $u \in N$ in a flow graph $G = (N, E, A, \mathbf{s}, \mathbf{e})$. Then the following algorithm decides whether x is a polynomial constant at u or not:

Phase 1: Determine a candidate value $d \in \mathbb{Z}$ for x at u by executing an arbitrary (cycle-free) path from $\mathbf{s}$ to u.

Phase 2: Associate transfer functions $f_e : R \to R$ with each edge $e \in E$ that represent $\mathsf{wp}(A_S(e))$:

$$f_e(I) = \begin{cases} I & \text{if } A(e) = \mathsf{skip} \\ I[p/x] & \text{if } A(e) = (x := p) \text{ with } p \in \mathbb{Z}[x_1, \ldots, x_n] \\ proj_x(I) & \text{if } A(e) = (x := t) \text{ with } t \notin \mathbb{Z}[x_1, \ldots, x_n] \\ proj_x(I) & \text{if } A(e) = (x := ?) \end{cases}$$

Set $A_0[w] = (x - d)$ and $A_0[u] = (0)$ for all $u \neq w$ and compute the largest solution of the equation system

$$\mathbf{A}[u] = A_0[u] + \sum_{v \in Succ[u]} f_e(\mathbf{A}[v]) \quad \text{for each } u \in N\,.$$

We can do this, for instance, as follows. Starting from $A_0[0]$ we iteratively compute, simultaneously for all program points u, the following sequences of ideals

$$A_{i+1}[u] = A_i[u] + \sum_{v \in Succ[u]} f_e(\mathbf{A}_i[v]) \, .$$

We stop upon stabilization, i.e., when we encounter an index i_s with $A_{i_s+1}[u] = A_{i_s}[u]$ for all $u \in N$. Obviously, $A_0[u] \subseteq A_1[u] \subseteq A_2[u] \subseteq \ldots$, such that computation must terminate eventually because $\mathbb{Z}[x_1, \ldots, x_n]$ is Noetherian. In this computation we represent ideals by finite bases and perform Gröbner basis computations in order to check whether $A_{i+1}[u] = A_i[u]$.[1]

Phase 3: Check, if the ideal computed for the start node, $A_{i_s}[\mathbf{s}]$, is (0). If so, x is a polynomial constant of value d at u; otherwise, x is not a polynomial constant at u.

Phase 2 can be seen as a backwards data flow analysis in a framework in which ideals of $\mathbb{Z}[x_1, \ldots, x_n]$ constitute data flow facts, the transfer functions are the functions f_e specified above, and the start value is A_0. Of course, we can use any evaluation strategy instead of the naive iterative approach described above.

7 Conclusion

In this paper we have shown that polynomial constant are decidable. In polynomial constant detection the integer operators $+, -, *$ are interpreted but all expressions using other operators are conservatively assumed to yield non-constant values. The detection algorithm for polynomial constants uses an indirect three phase approach. In the first phase a candidate value is computed that is verified in the second and third phase by means of a symbolic weakest precondition computation. The same algorithmic idea has been used by Müller-Olm and Rüthing in an efficient detection algorithm for Presburger constants [15]. In Presburger constant detection only the integer operators $+$ and $-$ are interpreted. While they represent assertions by affine vector spaces, we use here ideals in the polynomial ring $\mathbb{Z}[x_1, \ldots, x_n]$. More generally, we have analyzed the demands for making this general algorithmic idea effective.

Standard constant propagation uses a forward propagation while we use here a backwards propagation of assertions. Interestingly, Presburger constants can also be detected by a forward propagation of affine spaces. Karr [9] describes an algorithm that detects all Presburger constants, although he does not address completeness issues. There is a problem with forward propagation of affine spaces. In forward propagation of assertions we effectively compute strongest post conditions rather than weakest precondition and this computation involves union of assertions rather than intersections. Because affine spaces are not closed under union, Karr defines a (complicated) union operator of affine spaces that over-approximates their actual union by an affine

[1] As $A_{i+1}[u] \supseteq A_i[u]$ by construction, it suffices to check $A_{i+1}[u] \subseteq A_i[u]$.

space. One might be tempted to consider forward propagation of ideals of the polynomial ring $\mathbb{Z}[x_1, \ldots, x_n]$. At first glance, this idea looks promising, because ideals are closed under intersection and intersection of ideals correspond to union of their set of zeros, such that we can even precisely represent the union of assertions. There is, however, another problem: $\mathbb{Z}[x_1, \ldots, x_n]$ is not 'co-Noetherian', i.e., there are infinitely long strictly *decreasing* chains of ideals, e.g., $(x) \supset (x^2) \supset (x^3) \supset \ldots$. Therefore, strongest postcondition computation with ideals cannot be guaranteed to terminate in general.

Our approach to compute weakest preconditions symbolically with effective representations is closely related to abstract interpretation [3, 4]. Requirement (b) of the generic algorithm, universal conjunctivity of the representation mapping $\gamma : R \to 2^\Sigma$, implies that γ has a lower adjoint, i.e., that there is a monotonic mapping $\alpha : 2^\Sigma \to R$ such that (α, γ) is a Galois connection [12]. In the standard abstract interpretation framework, we are interested in computation of least fixpoints and the lower adjoint, α, is the abstraction mapping. Here, we are in the dual situation: we are interested in computation of greatest fixpoints. Thus, the role of the abstraction is played by the upper adjoint, $\gamma : R \to 2^\Sigma$. Funnily, this means that in a technical sense the members of R provide more concrete information than the members of 2^Σ and that we compute on the concrete side of the abstract interpretation. Thus, we may interpret the lattice R as an *exact partial representation* rather than an abstract interpretation. The representation via R is *partial* because it does not represent all assertions exactly; this is indispensable due to countability reasons because we cannot represent all assertions effectively. It is an *exact representation* because it allows us to infer the weakest preconditions arising in the S-constant algorithms precisely, which is achieved by ensuring that the initial value of the fixpoint computation is represented exactly and that the occurring operations on representations mirror the corresponding operations on assertions precisely.

By the very nature of Galois connections, the representation mapping γ and its lower adjoint α satisfy the two inequalities $\alpha \circ \gamma \sqsubseteq \mathsf{Id}_R$ and $\mathsf{Id}_\Sigma \subseteq \gamma \circ \alpha$, where Id_R and Id_Σ are the identities on R and Σ, respectively. Interestingly, for fundamental reasons none of these inequalities degenerates to an equality when we represent assertions by ideals of $\mathbb{Z}[x_1, \ldots, x_n]$ as in our algorithm for detection of polynomial constants. On the one hand, $\gamma \circ \alpha \neq \mathsf{Id}_\Sigma$ because the representation is necessarily partial. On the other hand, $\alpha \circ \gamma \neq \mathsf{Id}_R$ because the representation of assertions is not unique. For example, if $p \in \mathbb{Z}[x_1, \ldots, x_n]$ does not have a zero in the integers, we have $\mathcal{Z}((p)) = \emptyset$ such that $\mathcal{Z}((p)) = \mathcal{Z}((1)) = \mathcal{Z}((\mathbb{Z}[x_1, \ldots, x_n]))$. But by undecidability of Hilbert's tenth problem, we cannot decide whether we are faced with such a polynomial p and thus cannot effectively identify (p) and (1). This forces us to work with a non-unique representation. While we cannot decide whether the set of zeros of an ideal I given by a basis B is empty, we can decide whether it equals Σ because this only holds for $I = (0)$. Fortunately, this is the only question that needs to be answered for the weakest precondition.

As a consequence of non-uniqueness, the weakest precondition computation on ideals does not necessarily stop once it has found a collection of ideals that represents the largest fixpoint on assertions but may proceed to larger ideals that represent the same assertions. Fortunately, we can still prove termination by arguing on ideals directly.

Let us discuss some possible directions for future work. Firstly, it is interesting to implement the detection algorithm for polynomial constants and evaluate how it performs in practice. Secondly, we can look for other classes of constants to which the generic algorithm can be applied. One obvious class are polynomial constants over $\mathbb{Q}$ rather than $\mathbb{Z}$, where we can use essentially the same algorithm because $\mathbb{Q}[x_1, \ldots, x_n]$ is also a strongly computable ring. But we may also identify other totally different classes where assertions can be symbolically represented. On the theoretical side, there is the challenge to diminish the gap between the upper and lower complexity bound for the detection problem of polynomial constants. Currently, we have decidability as an upper bound, as witnessed by the algorithm in this paper, and PSPACE-hardness as a lower bound [15].

References

[1] A. V. Aho, R. Sethi, and J. D. Ullman. *Compilers: Principles, Techniques, and Tools.* Addison-Wesley, 1986.

[2] K. R. Apt and G. D. Plotkin. Countable nondeterminism and random assignment. *Journal of the ACM*, 33(4):724–767, 1986.

[3] P. Cousot and R. Cousot. Abstract interpretation: A unified lattice model for static analysis of programs by construction or approximation of fixpoints. In *Proceedings 4th POPL*, Los Angeles, California, January 1977.

[4] P. Cousot and R. Cousot. Abstract interpretation frameworks. *J. Logic Computat.*, 4(2):511–547, 1992.

[5] J. H. Davenport, Y. Siret, and E. Tournier. *Computer Algebra: Systems and Algorithms for Algebraic Computation.* Academic Press, 1988.

[6] C. Fischer and R. LeBlanc. *Crafting a Compiler.* Benjamin/Cummings Publishing Co., Inc., Menlo Park, CA, 1988.

[7] K. O. Geddes, S. R. Czapor, and G. Labahn. *Algorithms for Computer Algebra.* Kluwer, 1992.

[8] M. S. Hecht. *Flow analysis of computer programs.* Elsevier North-Holland, 1977.

[9] M. Karr. Affine relationships among variables of a program. *Acta Informatica*, 6:133–151, 1976.

[10] Mathematics of Program Construction Group. Fixed-point calculus. *Information Processing Letters*, 53(3):131–136, 1995.

[11] Y. V. Matiyasevich. *Hilbert's Tenth Problem.* The MIT Press, 1993.

[12] A. Melton, D. A. Schmidt, and G. E. Strecker. Galois connections and computer science applications. In D. Pitt, S. Abramsky, A. Poigné, and D. Rydeheard, editors, *Category Theory and Computer Programming*, volume 240 of *LNCS*, pages 299–312. Springer-Verlag, 1985.

[13] B. Mishra. *Algorithmic Algebra*. Springer-Verlag, 1993.

[14] S. S. Muchnick. *Advanced compiler design and implementation*. Morgan Kaufmann Publishers, San Francisco, California, 1997.

[15] M. Müller-Olm and O. Rüthing. The complexity of constant propagation. In D. Sands, editor, *ESOP 2001 (10th European Symposium on Programming)*, volume 2028 of *Lecture Notes in Computer Science*, pages 190–205. Springer, 2001.

[16] J. R. Reif and H. R. Lewis. Symbolic evaluation and the global value graph. In *Conf. Rec. 4th ACM Symposium on Principles of Programming Languages POPL'77*, pages 104–118, Los Angeles, CA, January 1977.

[17] M. Sagiv, T. Reps, and S. Horwitz. Precise interprocedural dataflow analysis with applications to constant propagation. *Theoretical Computer Science*, 167(1–2):131–170, 1996.

[18] B. Steffen and J. Knoop. Finite constants: Characterizations of a new decidable set of constants. *Theoretical Computer Science*, 80(2):303–318, 1991.

[19] F. Winkler. *Polynomial Algorithms*. Springer-Verlag, 1996.