

More on Nonregular PDL: Finite Models and Fibonacci-like Programs

David Harel and Eli Singerman

Department of Applied Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot, Israel

We continue research on enriching propositional dynamic logic (PDL) with nonregular programs. Previous work indicates that the general problem of characterizing those extensions for which PDL becomes undecidable is probably very hard. After observing that any nonregular extension increases the expressive power of PDL, we concentrate on one-letter extensions. First, we address the issue of finite models: A general condition is formulated, and is proven to be sufficient for a one-letter extension to violate the finite model property. We show the condition to hold in several cases, including all polynomials, sums of primes, and linear recurrences. We then build on a technique of Paterson and Harel, and show that the validity problem for PDL enriched with any Fibonacci-like sequence is Π_1^1 -complete. © 1996 Academic Press, Inc.

1. INTRODUCTION

Propositional dynamic logic (PDL) was introduced by Fischer and Ladner [FL], based upon the first-order version of Pratt [P1]. It is a direct extension of the propositional calculus, in which programs can appear in the formulas. Thus, for example, $P \rightarrow \langle \alpha \rangle Q$ asserts that whenever P holds it is possible to carry out some computation of α , leading to a state in which Q holds, and $\langle \alpha \rangle P \equiv \langle \beta \rangle P$ asserts a certain kind of equivalence of programs α and β . Formulas in PDL can involve many programs, and are able to express a wide variety of properties pertaining to their input/output behavior. See [H1, KT] for detailed surveys.

In most versions of PDL, the programs are taken to be regular sets of sequences of basic programs and tests. The validity/satisfiability problem for formulas in the regular PDL has been shown to be decidable, and is actually logspace-complete for exponential time [FL, P2]. One of the fundamental facts at the base of these results is the *small model property*, to the effect that every satisfiable formula of PDL has a finite model whose size is bounded by an exponential in the length of the formula. This makes possible a decidability proof based on filtration. (When there is no restriction on the size of a satisfying model, we call this the *finite model property*, or *fmp* for short.)

In the early 1980s, the problem of extending PDL with nonregular programs was raised. In terms of programming languages, moving from regular programs to, say, context-free ones, is tantamount to moving from iterative programs to (parameterless) recursive procedures. Ladner observed in

1977 that PDL with context-free programs is undecidable, but it remained to investigate (1) the level of undecidability of context-free PDL and (2) the point at which the extensions start becoming undecidable.

The first result was strikingly negative. Denote by PDL + L the logic obtained by allowing the language L as a single new program, in addition to the regular ones. In [HPS], it was shown that the validity problem for PDL + $a^d b a^d$ is highly undecidable, *viz.*, Π_1^1 -complete, where $a^d b a^d = \{a^i b a^i \mid i \geq 0\}$. This, together with the fact that context-free PDL can be easily shown to be in Π_1^1 , settled issue (1) above.¹

In contrast to this, it was proved in [KP] that PDL + $a^d b^d$ (despite not having the fmp [HPS]) is decidable, where $a^d b^d = \{a^i b^i \mid i \geq 0\}$. This was very puzzling, due to the similarity of $a^d b a^d$ and $a^d b^d$. In [HPS] it was shown that PDL with the addition of both $a^d b^d$ and $b^d a^d$ is also Π_1^1 -complete, which made things doubly strange. Things were clarified somewhat by the results of [HR], which are the only known results to date that apply to broad classes of programs, rather than to isolated examples. One of the main results of [HR] is that PDL extended by any *simple-minded* context-free program is decidable. Simple-mindedness is the property of being accepted by a pushdown automaton whose behavior is uniquely determined by the input symbol alone, with the internal state and stack symbol only helping determine whether the machine aborts or carries out its (unique) next step. Many context-free languages are simple-minded, including ones like $a^d b^d$ and all manner of parenthesis languages (semi-Dyck sets). A large class of non-context-free programs were also shown in [HR] to retain the decidability of PDL, including $a^d b^d c^d = \{a^i b^i c^i \mid i \geq 0\}$.

This summarizes what is known for programs with two or more atomic program letters. As far as *one-letter* programs are concerned, very little is known. It was shown in [HPS] that there is a primitive recursive nonregular one-letter language (which must be non-context-free, by Parikh's theorem [P]) whose addition as a program to PDL yields

¹ The a and b here are basic (atomic) programs, and the program is added as a single new program, not as a new formation rule. The same is true of all other additions we mention.

a Π_1^1 -complete validity problem. Later, in [HP], a rather involved technique was used to prove that validity for the specific extension $\text{PDL} + 2^*$ is also Π_1^1 -complete, where $2^* = \{a^{(2^i)} \mid i \geq 0\}$. The same is true, in fact, for k^* for any fixed k . That is really all. It is not known whether there is any nonregular one-letter program whose addition to PDL retains decidability, or whether there is any one-letter program that ruins the decidability of PDL, but whose words grow sub-exponentially.

Particularly intriguing are the cases of squares and cubes: are $\text{PDL} + *^2$ and $\text{PDL} + *^3$ decidable, where $*^k = \{a^{(i^k)} \mid i \geq 0\}$? There are indications that these questions are nontrivial. For example, J. Stavi pointed out several years ago that one can easily write down a formula in $\text{PDL} + *^3$ that is valid if and only if the answer to some open problem in number theory is yes.²

All of this indicates that the general quest of characterizing those extensions for which PDL becomes undecidable is probably very hard.

In Section 2, we observe that the addition of *any* nonregular program increases the expressive power of PDL. An obvious upshot is that no nonregular extension can be shown decidable by simply reducing it to regular PDL. Indeed, the techniques of [KP, HR] are considerably more involved. With this situation in mind, there are two kinds of things that can and should be attempted. The first is to keep working hard to establish the decidability status of *specific* cases; maybe the interesting ones (like squares and cubes) will eventually yield. The second is to seek *general* results, that apply to whole classes of languages, but establishing properties weaker than decidability or undecidability. This paper reports on work we have carried out along both these lines.

In Section 3, we consider finite models. The motivation, of course, is that if we can show that a PDL extension does not have the small model property, then it cannot be proved decidable by a standard filtration technique. (If it does not have the *finite* model property, then it cannot be proved to be even r.e. by a standard enumeration technique.) We formulate a general condition that is shown to be sufficient (but not necessary) for a one-letter nonregular extension of PDL to violate the finite model property. We then show that this condition holds true for a wide variety of one-letter extensions of PDL, including all polynomials (of which our friends, the squares and cubes are special cases), sums of primes, factorial numbers and sequences defined by linear recurrences.

These latter sequences include Fibonacci-like ones, and for these we have a much stronger result, proved in Section 4:

² For example, it is not known whether every natural number greater than 1000 can be written as the sum of seven cubes. The formula for this is simply $[(*)^3]^7 P \rightarrow [a^{1001}][a^*] P$. Thus, if $\text{PDL} + *^3$ were decidable we could—at least in principal—settle this problem.

Strengthening the technique of [HP], we prove that the validity problem for PDL enriched with any sequence of the form: f_0, f_1, f_2, \dots , where $f_0 < f_1$ and $f_{i+1} = f_i + f_{i-1}$, for $i \geq 1$, is Π_1^1 -complete.

2. PRELIMINARIES AND BASIC RESULTS

We first define PDL. Let *Prop* be an infinite set of atomic propositions and *Prog* be an infinite set of atomic programs. The set of formulas and the set of programs of PDL are defined by mutual induction as follows:

- Every proposition $P \in \text{Prop}$ is a formula.
- If φ and ψ are formulas, then so are $\neg\varphi$ and $(\varphi \vee \psi)$.
- If α is a program and φ is a formula, then $\langle \alpha \rangle \varphi$ is a formula.
- Every atomic program $a \in \text{Prog}$ is a program.
- If α and β are programs, then so are $(\alpha; \beta)$, $(\alpha \cup \beta)$ and $(\alpha)^*$.
- If φ is a formula, then $\varphi?$ is a program (called a *test*).

We sometimes omit the “;” in programs and the parentheses in formulas and programs. Also, we use standard abbreviations, such as **true**, **false**, \wedge , \equiv , and \rightarrow (for “implies”). We write $[\alpha] \varphi$ for $\neg \langle \alpha \rangle \neg \varphi$.

Intuitively, $(\alpha; \beta)$ means “do α followed by β ”, $(\alpha \cup \beta)$ means “do α or β nondeterministically, and $(\alpha)^*$ means “do α any number of times, 0 or more, nondeterministically.” The test program $\varphi?$ means “proceed, with no side-effects, only if φ is true”.

PDL formulas are interpreted over models $M = (W, \tau, \rho)$, where W is the set of *states*, $\tau: \text{Prop} \rightarrow 2^W$ provides meaning for the atomic propositions, and $\rho: \text{Prog} \rightarrow 2^{W \times W}$ provides meaning for the atomic programs. Now, by mutual induction, ρ is extended to all programs and we define the satisfaction of a formula φ in a state s of a model M , denoted $M, s \models \varphi$:

- For a proposition $P \in \text{Prop}$, we have $M, s \models P$ iff $s \in \tau(P)$.
- $M, s \models \neg\varphi$ iff it is not the case that $M, s \models \varphi$.
- $M, s \models \varphi \vee \psi$ iff $M, s \models \varphi$ or $M, s \models \psi$.
- $M, s \models \langle \alpha \rangle \varphi$ iff there exists a state t such that $(s, t) \in \rho(\alpha)$ and $M, t \models \varphi$.
- $\rho(\alpha; \beta) = \rho(\alpha) \circ \rho(\beta) = \{(s, t) \mid \text{there is } u \text{ such that } (s, u) \in \rho(\alpha) \text{ and } (u, t) \in \rho(\beta)\}$.
- $\rho(\alpha \cup \beta) = \rho(\alpha) \cup \rho(\beta)$.
- $\rho(\alpha^*) = (\rho(\alpha))^* = \{(s, t) \mid \text{there are states } s_0, s_1, \dots, s_n, \text{ for some } n \geq 0, \text{ such that } s = s_0, t = s_n, \text{ and for all } 1 \leq i \leq n, (s_{i-1}, s_i) \in \rho(\alpha)\}$.
- $\rho(\varphi?) = \{(s, s) \mid M, s \models \varphi\}$.

A formula φ is *satisfiable* if there is a model M and a state s such that $M, s \models \varphi$. The formula φ is *valid*, written $\models \varphi$, if for every model M and state s , we have $M, s \models \varphi$. Clearly, φ is valid iff $\neg\varphi$ is not satisfiable. The *satisfiability* (respectively, *validity*) *problem* is to determine, given a formula φ , whether φ is satisfiable (respectively, valid).

Together, the upper bound of Pratt [P2] and the lower bound of [FL] yield:

THEOREM 2.1. *The validity/satisfiability problem for PDL is logspace-complete for exponential time.*

Nonregular PDL is obtained by extending the programs of PDL with nonregular programs. Let L be nonregular set over the alphabet $Prog$. Denote by $PDL + L$ the logic obtained by allowing the language L as a single new program, in addition to the regular ones. Syntactically, L is treated as an atomic program (i.e., as an element of $Prog$). Semantically, $PDL + L$ is interpreted over the same models as PDL, with $\rho(L) = \bigcup_{w \in L} \rho(w)$.

The known results on nonregular PDL were mentioned in the introduction. It is interesting to observe that enriching PDL with *any* nonregular program increases expressive power:

THEOREM 2.2. *If L is any nonregular language over $Prog$, then $PDL + L > PDL$.³*

Proof. The result can be proved by embedding PDL into the second order arithmetic of k successors (SkS). It is not hard to see that any set of nodes definable in SkS is regular, so that the addition of a nonregular predicate increases its expressive power.

A more direct proof can be obtained as follows: For $Prog = \{a_1, \dots, a_k\}$, define the model $T = (W, \tau, \rho)$ to be the complete k -ary tree, in which P holds at the root only, each internal node has k offspring, one for each program a_i , but with all edges pointing upward. Thus, from the node $u = a_{i_1} \dots a_{i_j}$, the only program possible that leads to a state satisfying P is u^R , i.e., $a_{i_j} \dots a_{i_1}$, which is u in reverse. Now, let $Lan(\varphi) = \tau(\varphi)^R$, i.e., the language over $Prog$ whose words are precisely the paths in T that lead from states that satisfy φ to the root. One now shows that $Lan(\varphi)$ is a regular set over the alphabet $Prog$, by induction on the structure of φ . Hence, $\langle L \rangle P$ cannot be equivalent to any PDL formula, since $Lan(\langle L \rangle P) = L$ is nonregular. (Details of this proof appear in the preliminary version of this paper [HS].) ■

3. THE FINITE MODEL PROPERTY

The technique used in [HPS] to show that $PDL + a^A b^A$ violates the fmp uses a comb-like model built up from two

³ Here, $>$ denotes "strictly greater than in expressive power": There is a formula in $PDL + L$ that has no equivalent formula in PDL, where equivalent means having the same truth value in every state of every model.

atomic programs. It thus does not work for one-letter alphabets. In fact, the only information we have about the fmp for one-letter extensions of PDL is the trivial fact that—PDL + k^* for a fixed k —violates it. (These extensions are known to be *highly* undecidable [HP] and therefore cannot have the fmp.) Since we are particularly interested in extensions such as squares and cubes ($PDL + *^2$ and $PDL + *^3$), for which the decidability status is unknown and seems to be hard to establish, we tackle the more humble issue of the fmp here. We prove a general result in Theorem 3.3 that constitutes a tool for showing that one-letter extensions of PDL do not have the fmp. We then use this result to obtain specific results, such as:

PROPOSITION 3.1 (Squares and Cubes). *$PDL + *^2$ and $PDL + *^3$ do not have the fmp.*

Let us now prepare for the theorem.

DEFINITION 3.2. For a language L over the alphabet Σ with $a \in \Sigma$, let

$$\#L = \{i \mid a^i \in L\}.$$

If α is a program, we use $\#\alpha$ as an abbreviation of $\#L(\alpha)$, where $L(\alpha)$ is the language depicted by the program α . (Thus, for $S \subseteq \mathcal{N}$, we can use a^S to denote the language $\{a^i \mid i \in S\}$, so that we have $\#a^S = S$.)

Note. In the rest of this section all arrows in models denote a -transitions.

THEOREM 3.3. *Let $S \subseteq \mathcal{N}$. Suppose that for some program α in $PDL + a^S$ with $L(\alpha) \subseteq a^*$, the following conditions are satisfied:*

1. *There exists some n_0 such that for each $x \geq n_0$ and for all $i \in \#\alpha$,*

$$x \in S \Rightarrow x + i \notin S.$$

2. *For every $l, m > 0$ there exists $x, y \in S$, with $x > y \geq l$ and $d \in \#\alpha$, such that $(x - y) \equiv d \pmod{m}$.*

Then $PDL + a^S$ does not have the fmp.

Proof. The proof is based on the following intuition: every infinite path in a finite model must "close up" in a circular fashion. This forces some periodic property along such a path on every formula over one-letter programs. Let S and α satisfy the conditions in the statement of the theorem. We will use the nonperiodic nature of the set S to construct a satisfiable formula φ in $PDL + a^S$ that has no finite model.

Let $\varphi = \varphi_1 \wedge \varphi_2 \wedge \varphi_3$, where

$$\varphi_1 = [a^*] \langle a \rangle \text{ true}$$

$$\varphi_2 = [a^S] P$$

$$\varphi_3 = [a^{n_0}] [a^*] (P \rightarrow [\alpha] \neg P).$$

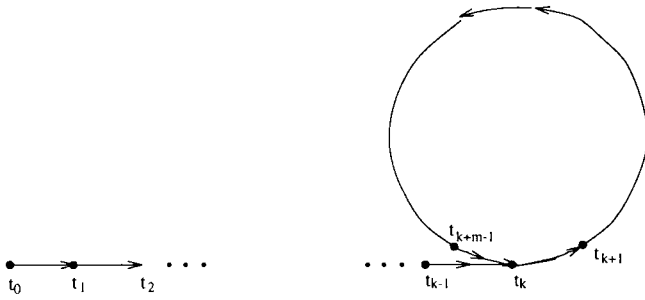
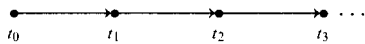


FIGURE 1

Here, n_0 is the constant from property (1), and a^{n_0} is written out in full.

In the infinite model



in which $\tau(P) = \{t_x \mid x \in S\}$, we obviously have $t_0 \models \varphi$. (That φ_3 holds in t_0 is due to property (1).) We will show that φ does not have a finite model. Suppose that $M, t_0 \models \varphi$ for some finite model M . By φ_1 and the finiteness of M it follows that there exists in M some path of the form shown in Fig. 1.

Let m denote the size of the cycle in Fig. 1. For every $z \in \mathcal{N}$, denote by z' the remainder of $(z - k)$ when divided by m . Note that for $z \geq k$, the state $t_{k+z'}$ can be reached from t_0 by executing the program a^z .

By property (2), we can find $x, y \in S$ and $d \in \# \alpha$ such that

$$x > y > \max(n_0, k) \quad \text{and} \quad (x - y) \equiv d \pmod{m}.$$

The truth of φ_2 at t_0 implies that

$$t_{k+y'} \models P \quad \text{and} \quad t_{k+x'} \models P.$$

Since $y > n_0$, it follows from φ_3 that $t_{k+(y+d)'} \models \neg P$. However,

$$(x - y) \equiv d \pmod{m} \Rightarrow (y + d)' = x'.$$

Hence, $t_{k+x'} \models \neg P$, which is a contradiction. ■

Remark. It is sometimes useful to replace condition (2) of the theorem by the weaker condition, call it (2'), in which the consequent does not have to hold for every modulus m , but only for every $m \geq m_0$, for some fixed m_0 . To prove this, we have to add the following conjunct to φ :

$$\varphi_{1.5} = Q_0 \wedge [a^*] \left(\bigwedge_{i=0}^{m_0-1} (Q_i \rightarrow [a] Q_{(i+1) \bmod m_0}) \wedge \bigwedge_{0 \leq i < j \leq m_0-1} \neg(Q_i \wedge Q_j) \right).$$

It forces the size of the cycle in Fig. 1 to be a positive multiple of m_0 .

The following propositions provide examples of one-letter extensions, for which the absence of the fmp can be proved using Theorem 3.3. (In the proofs we refer to the conditions of the theorem by (1) and (2) (or (2')).)

First, we prove the “squares” part of Prop. 3.1. Let $S_{\text{squares}} = \{i^2 \mid i \in \mathcal{N}\}$. To satisfy (1), take $n_0 = 1$ and $\alpha = a$. (Hence, $\# \alpha = \{1\}$.) As for (2), given $l, m > 0$, let $d = 1$, choose $y = (q \cdot m)^2 > l$ and $x = (q \cdot m + 1)^2$. Then indeed, $x, y \in S_{\text{squares}}$, $x > y \geq l$ and $x - y = (q \cdot m + 1)^2 - (q \cdot m)^2 \equiv d \pmod{m}$.

This can be generalized as follows:

PROPOSITION 3.4 (Polynomials). *For every polynomial of the form*

$$p(n) = c_i n^i + c_{i-1} n^{i-1} + \dots + c_0 \in \mathcal{P}[n]$$

with $i \geq 2$ and positive leading coefficient $c_i > 0$, let $S_p = p(\mathcal{N}) \cap \mathcal{N}$. Then, $\text{PDL} + a^{S_p}$ does not have the fmp.

Proof. To satisfy the conditions of Theorem 3.3, choose j_0 such that

$$p(j_0) - c_0 > 0.$$

Take α such that $\# \alpha = \{p(j_0) - c_0\}$. Find some n_0 such that for every $x \geq n_0$, we have

$$p(x + 1) - p(x) > p(j_0) - c_0.$$

This takes care of (1). Next, given $l, m > 0$, for $d = p(j_0) - c_0$, $y = p(q \cdot m) > l$, and $x = p(q' \cdot m + j_0) > y$, we have

$$x - y = p(q' \cdot m + j_0) - p(q \cdot m) \equiv p(j_0) - c_0 \pmod{m}. \quad \blacksquare$$

PROPOSITION 3.5 (Sums of Primes). *Let p_i be the i th prime (with $p_1 = 2$), and define*

$$S_{\text{sop}} = \left\{ \sum_{i=1}^n p_i \mid n \geq 1 \right\}.$$

Then, $\text{PDL} + a^{S_{\text{sop}}}$ does not have the fmp.

Proof. Clearly, (1) is satisfied with $n_0 = 3$ and $\alpha = a$. To see that (2) holds, we use a well known theorem of Dirichlet: A necessary and sufficient condition for the existence of infinitely many primes in the arithmetic progression $s + j \cdot t$, $j \geq 0$, is that $\text{gcd}(s, t) = 1$.

Now, given $l, m > 0$, find some i_0 such that

$$p_{i_0-1} > l \quad \text{and} \quad p_{i_0} \equiv 1 \pmod{m}.$$

(The existence of such p_{i_0} follows from Dirichlet's theorem applied to the arithmetic progression $1 + j \cdot m, j \geq 0$.)

Let

$$d = 1, \quad y = \sum_{i=1}^{i_0-1} p_i, \quad \text{and} \quad x = \sum_{i=1}^{i_0} p_i.$$

Then $x, y \in S_{\text{top}}, x > y \geq l$ and $x - y = p_{i_0} \equiv d \pmod{m}$. ■

PROPOSITION 3.6 (Factorials). For $S_1 = \{n! \mid n \in \mathcal{N}\}$, $\text{PDL} + a^{S_1}$ does not have the fmp.

Proof. Here, we take $\alpha = a^{S_1}$ and $n_0 = 2$. In order to satisfy (1), we have to show that if $2 \leq x$ and $y \in S_1$, we have $x + y \notin S_1$. Let $x = s!$ and $y = t!$, and assume for $t \geq s \geq 2$ that

$$(i) \quad s! + t! = u!.$$

Obviously $u > t$. If $t \neq s$, then by dividing both sides of (i) by $s!$ we obtain $(s + 1) \mid 1$. If $t = s$, we obtain $(s + 1) \mid 2$. Both of these contradict the fact that $s \geq 2$. Turning now to (2), given $l, m > 0$, let $i_0 > \max(m, l)$. Take $d = y = i_0!$ and $x = (2i_0)!$ (since $\# \alpha = S_1$, we indeed have $d \in \# \alpha$ and $x > y \in S_1$). We now have $x, y > l$ and $x - y \equiv d \pmod{m}$. ■

PROPOSITION 3.7 (Fibonacci Numbers). Let $S_f = \{f_n\}_{n \geq 0}$, where $\{f_n\}_{n \geq 0}$ is the Fibonacci sequence: $f_0 = 1, f_1 = 2, f_{n+1} = f_n + f_{n-1}$, for $n \geq 1$. Then, $\text{PDL} + a^{S_f}$ does not have the fmp.

Proof. For $\alpha = a$ and $n_0 = 2$, (1) is satisfied. For (2) we shall use the following:

CLAIM. For every $l > 0$ and $m \geq 3$, there exists i such that $f_i \geq l$ and

$$f_i \equiv 1 \pmod{m}.$$

Proof of Claim. Denote by r_n the remainder of f_n when divided by m . Consider the first $m^2 + 1$ pairs of the sequence

$$\langle r_0, r_1 \rangle, \langle r_1, r_2 \rangle, \langle r_2, r_3 \rangle, \dots, \langle r_{m^2}, r_{m^2+1} \rangle, \dots$$

Since there are m different remainders modulo m , there must be some s and t such that

$$(ii) \quad m^2 \geq s > t \quad \text{and} \quad \langle r_s, r_{s+1} \rangle = \langle r_t, r_{t+1} \rangle.$$

Let t be the least number for which there exists an s as above. We show that $t = 0$. Otherwise, if $t > 0$, then by (ii) and the facts that $r_{s-1} = r_{s+1} - r_s$ and $r_{t-1} = r_{t+1} - r_t$, we have $r_{s-1} = r_{t-1}$. Hence $\langle r_{s-1}, r_s \rangle = \langle r_{t-1}, r_t \rangle$, which is a contradiction to the minimality of t .

We have thus shown that the first pair that repeats itself is $\langle r_0, r_1 \rangle$. Now, since $f_0 = 1, f_1 = 2$ and $m \geq 3$, we have $r_0 = 1, r_1 = 2$. Therefore, $\langle r_s, r_{s+1} \rangle = \langle 1, 2 \rangle$, for some $s > 0$. We can now apply the previous argument to the sequence of pairs starting at $\langle r_s, r_{s+1} \rangle$, i.e., to

$$\langle 1, 2 \rangle = \langle r_s, r_{s+1} \rangle, \langle r_{s+1}, r_{s+2} \rangle, \dots$$

and find a further occurrence of $\langle 1, 2 \rangle$. Induction completes the proof of the Claim. ■

It is now easy to verify (2'). Let $m_0 = 3$. Given $m \geq m_0$ and $l > 0$, use the Claim to find some $f_{i_0} > 2l$ with $f_{i_0} \equiv 1 \pmod{m}$. For $d = 1, y = f_{i_0-1}$ and $x = f_{i_0+1}$, we have

$$x, y \geq l \quad \text{and} \quad x - y = f_{i_0+1} - f_{i_0-1} = f_{i_0} \equiv 1 \pmod{m}. \quad \blacksquare$$

Proposition 3.7 can be generalized to a large family of sequences defined by linear recurrences with constant coefficients:

PROPOSITION 3.8 (Linear Recurrence). Let $S_l = \{l_n\}_{n \geq 0}$ be the sequence defined by:

$$l_n = c_1 l_{n-1} + \dots + c_{k-1} l_{n-k+1} + l_{n-k}, \quad \text{for } n \geq k,$$

where $c_i \in \mathcal{N}$ for each $i, c_1 \neq 0$, and $l_1, \dots, l_k \in \mathcal{N}^+$ are arbitrary. Then, $\text{PDL} + a^{S_l}$ does not have the fmp.

Proof. Since the growth of the sequence is exponential, (1) is satisfied for every α for which $L(\alpha) \subseteq a^+$ is finite (the particular α we shall use is defined later.) For (2') (the version of (2) described in the Remark following Theorem 3.3) we use the following:

CLAIM. Let $i_0 \geq 0$. For every $m > l_{i_0}$, there are infinitely many l_i with $l_i \equiv l_{i_0} \pmod{m}$.

Proof of Claim. Let $i_0 \geq 0$ and $m > l_{i_0}$. Denote by r_n the remainder of l_n when divided by m . Observe that the r_n 's satisfy—modulo m —the same recurrence relation as the l_n 's, namely, that

$$(iii) \quad r_n \equiv c_1 r_{n-1} + \dots + c_{k-1} r_{n-k+1} + r_{n-k} \pmod{m},$$

$$\text{for } n \geq k.$$

Among the first $m^k + 1$ k -tuples of the sequence

$$\langle r_{i_0}, \dots, r_{i_0+k-1} \rangle, \langle r_{i_0+1}, \dots, r_{i_0+k} \rangle, \langle r_{i_0+2}, \dots, r_{i_0+k+1} \rangle, \dots$$

there is at least one k -tuple that repeats itself. Let $t \geq i_0$ be the least number for which there exists an s such that

$$(iv) \quad i_0 + m^k \geq s > t$$

and

$$\langle r_s, \dots, r_{s+k-1} \rangle = \langle r_t, \dots, r_{t+k-1} \rangle.$$

We show that $t = i_0$, i.e., that the first k -tuple that repeats itself is $\langle r_{i_0}, \dots, r_{i_0+k-1} \rangle$. Otherwise, if $t > i_0$, then by (iii) and (iv) we have $r_{s-1} \equiv r_{t-1} \pmod{m}$. However, both r_{s-1} and r_{t-1} are remainders modulo m , so that $r_{s-1} = r_{t-1}$. Therefore $\langle r_{s-1}, \dots, r_{s+k-2} \rangle = \langle r_{t-1}, \dots, r_{t+k-2} \rangle$, which is a contradiction to the minimality of t .

We have thus shown that $r_s = r_{i_0}$, for some $s > i_0$. Since $m > l_{i_0}$, we have $r_{i_0} = l_{i_0}$. Hence, $l_s \equiv l_{i_0} \pmod{m}$, for some $s > i_0$. We can now apply the previous argument to the sequence of k -tuples starting at $\langle r_s, \dots, r_{s+k-1} \rangle$, and find a further occurrence of $r_s = r_{i_0} (= l_{i_0})$. Induction completes the proof of the claim. ■

We can now verify the conditions of the theorem. Choose i_0 and j_0 such that $l_{i_0} < l_{j_0}$. Take α such that $\# \alpha = \{l_{j_0} - l_{i_0}\}$. Find some n_0 such that the difference between any two consecutive elements of $\{l_n\}_{n \geq 1}$ greater than n_0 is more than $l_{j_0} - l_{i_0}$. This takes care of (1). For (2'), let $m_0 = l_{j_0} + 1$. Given $m \geq m_0$ and $l > 0$, use the claim to find some $l_i \geq l$ with $l_i \equiv l_{i_0} \pmod{m}$, and some $l_j > l_i$ with $l_j \equiv l_{j_0} \pmod{m}$. For $d = l_{j_0} - l_{i_0}$, $y = l_i$ and $x = l_j$, we have

$$x > y \geq l \quad \text{and} \quad x - y = l_j - l_i \equiv l_{j_0} - l_{i_0} \pmod{m}. \quad \blacksquare$$

4. UNDECIDABILITY FOR FIBONACCI PROGRAMS

That the validity problem for PDL + 2* is highly undecidable (actually, Π^1_1 -complete) is proved in [HP] by a rather complex reduction from a recurring tiling (domino) problem. In this section we extend the method of [HP] to prove the Π^1_1 -completeness of different one-letter extensions, the Fibonacci-like sequences. Let $F = \{f_i\}_{i \geq 0} \subseteq \mathcal{N}$ be a sequence defined by $f_0 < f_1$ (arbitrary), and $f_{i+1} = f_i + f_{i-1}$, for $i \geq 1$.

THEOREM 4.1. *The validity problem for PDL + a^F is Π^1_1 -complete.*

The rest of this section is devoted to the proof. For convenience, we shall talk about the equivalent formulation—that satisfiability is Σ^1_1 -complete. The upper bound, i.e., that the problem is in Σ^1_1 , can be established without difficulty using standard arguments, cf. [HPS, Lemma 6.3]. We thus concentrate on the lower bound, carrying out a reduction from a recurring tiling problem, shown in [H2] to be Σ^1_1 -complete. We first describe the tiling problem and prove a grid-like property of the set F , and then present the details of the reduction.

Remark. The theorem concerns the family of all Fibonacci-like sequences. However, throughout the proof we demonstrate the central issues with the original Fibonacci sequence, which is the case of $f_0 = 1, f_1 = 2$.

4.1. Preparation

Let G' be the strict upper positive octant of the integer grid \mathcal{Z}^2 , i.e., $G' = \{(i, j) \mid 0 \leq i < j\}$. Three diagonals and three columns in G' play a special role, these are

$$D_1 = \{(i, i + 1) \mid i \geq 0\}, \quad D_2 = \{(i, i + 2) \mid i \geq 0\}$$

and $D_3 = \{(i, i + 3) \mid i \geq 0\}$,

$$C_0 = \{(0, i) \mid i \geq 1\}, \quad C_1 = \{(1, i) \mid i \geq 2\}$$

and $C_2 = \{(2, i) \mid i \geq 3\}$.

The portion to be tiled is $G = G' - D_1 - C_0$. It is known that the problem of deciding whether a given finite set of tiles can tile G is undecidable, viz., co-r.e. complete [B]. Imposing recurrence conditions on the tiling (e.g., requiring that a certain tile appears infinitely often) makes these problems highly undecidable, viz. Σ^1_1 -complete, as shown in [H2]. We shall use a specialized version of the problem. For every $i \geq 0$, let

$$G_i = \{(j, i) \mid 0 \leq j \leq i - 1\} \cup \{(i - 2, i + 1)\} \\ \cup \{(i, j) \mid j \geq i + 1\}.$$

PROPOSITION 4.2. *The problem of deciding whether a given set $T = \{d_0 \dots d_m\}$ of tile types can tile G such that all the tiles in D_2 have the same color on their left edges, and for $i \geq 1$, the tile d_0 occurs at least once in every set $G_i \cap G$, is Σ^1_1 -complete.*

The problem in Proposition 4.2 is very similar to problem R6 in [H2], and its Σ^1_1 -completeness can be established easily using the techniques given there. To get a feeling for the problem, consult Fig. 2, in which the emphasized edges are to be monochromatic and one of the sets used in the recurrence condition, in this case G_4 , is marked.

The undecidability proof for PDL + 2* in [HP] was made possible by the following grid-like property of sums of powers of 2:

$$\text{if } i \neq j \text{ and } 2^i + 2^j + 2^k \in \{2^n + 2^m \mid n, m \geq 0\}$$

$$\text{then } k = i \text{ or } k = j.$$

We will use a similar, though less obvious, grid-like property of the set F .

DEFINITION 4.3. A sum $f_{i_1} + \dots + f_{i_n}$ of elements of F is called *proper* if the difference between every two indices is at least 2 (i.e., $|i_j - i_l| \geq 2$ for $1 \leq j \neq l \leq n$). A proper sum of $n \geq 1$ elements is called a *proper n -sum*.

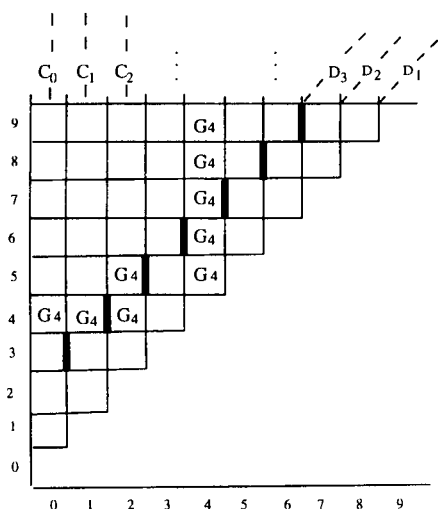


FIGURE 2

In a way similar to the proof in [HP], we mark our portion of the grid with sums of pairs of elements of F , and analyze which of the other points in this portion can be reached by adding to such sums (i.e., to an element in the grid) a *third* element of F . The situation is captured by the following lemma.

LEMMA 4.4. (Grid-like Property). *For any $i \geq 1$ and $j \geq i + 2$, consider the proper 2-sum $f_i + f_j$. The following constitute precisely all pairs (n, m) for which $f_n + f_m$ is a proper 2-sum (with $m \geq n + 2$), which satisfies $f_n + f_m = f_i + f_j + f_k$, with $f_k \in F$. Given such i and j , these pairs depend on the value of $j - i$, as follows:*

- (Δ) $j - i = 2$: $(i, j + 1), (i, j + 2), (i - 2, j + 1)$.
- (\circ) $j - i = 3$: $(i + 1, j), (i, j + 1), (i, j + 2), (i + 2, j + 1)$.
- (\square) $j - i > 3$: $(i + 1, j), (i + 2, j), (i, j + 1), (i, j + 2)$.

These three cases are illustrated in Fig. 3, using the symbols attached to each (recall that the illustrations are of the standard Fibonacci sequence): Within each symbol, at location (i, j) , is the value $f_i + f_j$, and the arrows point to the corresponding $f_n + f_m$.

Proof. We shall use the fact that two proper sums with different sets of indices cannot represent the same number. To see this, one can use the easily verified inequalities $f_0 + f_2 + \dots + f_{2l} < f_{2l+1}$ and $f_1 + f_3 + \dots + f_{2l+1} < f_{2l+2}$, and prove that the sum containing the largest index (not common to both sums) must represent a larger number.

It follows, in particular, that a proper 3-sum cannot equal a proper 2-sum. This narrows down the number of possible k 's for which adding f_k to the proper 2-sum $f_i + f_j$ might result in another proper 2-sum $f_n + f_m$. We prove in detail the $j - i = 2$ case; the remaining cases are proven similarly.

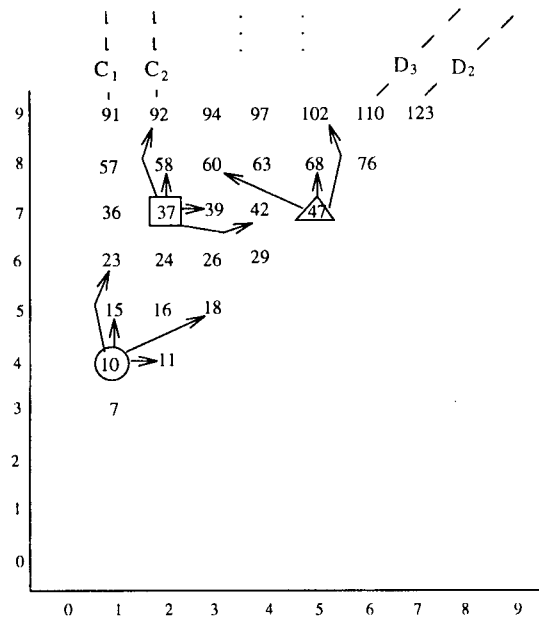


FIGURE 3

Since $f_i + f_j + f_k$ (i.e., $f_i + f_{i+2} + f_k$) is a proper 3-sum (and therefore cannot equal a proper 2-sum), for every $k < i - 1$ and for every $k > i + 3$, the only k 's we ought to consider are the following:

- $k = i + 3$ ($= j + 1$): Here, $f_i + f_j + f_k = f_i + f_{i+2} + f_{i+3} = f_i + f_{i+4}$, so that $n = i$ and $m = j + 2$.
- $k = i + 2$ ($= j$): For $i > 1$, we have $f_i + f_{i+2} + f_{i+2} = (f_{i-2} + f_{i-1}) + (f_i + f_{i+1}) + f_{i+2} = f_{i-2} + f_{i+1} + f_{i+3}$, which is not a proper 2-sum. For $i = 1$, the result is not a proper 2-sum either, as can be checked easily.
- $k = i + 1$: Here, $f_i + f_{i+2} + f_{i+1} = f_i + f_{i+3}$, so that $n = i$ and $m = j + 1$.
- $k = i$: For $i > 1$, we have $f_i + f_{i+2} + f_i = f_{i-2} + f_{i-1} + f_i + f_{i+2} = f_{i-2} + f_{i+1} + f_{i+2} = f_{i-2} + f_{i+3}$. Hence, $n = i - 2$ and $m = j + 1$. For $i = 1$, the result is not a proper 2-sum.
- $k = i - 1$: Here $f_i + f_{i+2} + f_{i-1} = f_{i+1} + f_{i+2} = f_{i+3}$, which is not a proper 2-sum. ■

4.2. The Reduction

Suppose we are given a set $T = \{d_0, \dots, d_m\}$ of tile types involving the colors c_0, \dots, c_l , where each d_i is given by a quadruple $(left_i, right_i, up_i, down_i)$. We describe an effective procedure for constructing a formula φ_T of $PDL + a^F$ which is satisfiable if and only if T can tile G as in Proposition 4.2. Thus, satisfiability for $PDL + a^F$ will be Σ_1^1 -hard.

Without loss of generality, assume that k is a power of 2. We use LEFT, RIGHT, UP, and DOWN to abbreviate four sets of log k atomic formulas that will be used to encode

entries refer to the unique R_j true at each point. The first conjunct marks the two columns C_1 and C_2 , respectively. Each of these columns is then marked with the appropriate R_j 's. This is done in two steps:

1. The second conjunct of (3) forces the four lowest elements to be marked, such that for $1 \leq i \leq 2$, $0 \leq j \leq 3$, R_j holds at $(i, i+j+2)$.

2. The third conjunct forces the R_j marking of C_1 and of C_2 to propagate upwards such that, for $1 \leq i \leq 2$, $0 \leq j \leq 3$, R_j holds at $(i, 4r+i+j+2)$, for all $r \geq 0$. This is done using the fact that, for $1 \leq i \leq 2$, the a^F -successors in column C_i of a C_i -point are precisely those at one and at two positions above it.

Finally, The remaining conjunct of (3) makes $R_{(r-s-2) \bmod 4}$ true at (s, r) for all $(s, r) \in G$. This can be easily verified by induction.

$$\begin{aligned} & \bigwedge_{i=1}^2 [a^F][a^{f_i}] C_i \wedge \bigwedge_{i=1}^2 \bigwedge_{j=0}^3 [a^{f_i}][a^{f_{i+j+2}}] R_j \\ & \wedge \bigwedge_{i=1}^2 \bigwedge_{j=0}^3 [a^F][a^{f_i}] \\ & \times \left(R_j \rightarrow [a^F] \left(C_i \rightarrow \bigvee_{1 \leq k \leq 2} R_{(j+k) \bmod 4} \right) \right) \\ & \wedge [a^F a^F] \left((R_0 \vee R_1) \rightarrow \right. \\ & \left. [a^F] \left(P \rightarrow \left(Q \vee \bigvee_{0 \leq j \leq 3 \text{ and } j \neq i} R_j \right) \right) \right) \\ & \wedge \left((R_2 \vee R_3) \rightarrow [a^F] \left(P \rightarrow \bigvee_{0 \leq j \leq 3 \text{ and } j \neq i} R_j \right) \right). \quad (3) \end{aligned}$$

The next two clauses assert, respectively, that tiles from T are placed at each point of G , and that this tiling satisfies the usual color-matching constraints required of the edges, together with the special requirement of the colors on the left-hand edges of D_3 -points (as stated in Prop. 4.2).

$$[a^F a^F] \left(P \wedge \neg Q \rightarrow \bigvee_{i=0}^m \text{LRUD} = d_i \right) \quad (4)$$

$$\begin{aligned} & [a^F a^F] \left(\bigwedge_{i=0}^3 \bigwedge_{j=0}^3 (R_i \rightarrow ((\text{RIGHT} = c_j \rightarrow \right. \\ & \left. [a^F](R_{(i-1) \bmod 4} \rightarrow \text{LEFT} = c_j)) \right. \\ & \left. \wedge (\text{UP} = c_j \rightarrow [a^F](R_{(i+1) \bmod 4} \rightarrow \text{DOWN} = c_j))) \right). \quad (5) \end{aligned}$$

To see why (5) forces tiles at D_2 -points to have the same color on their left-hand edge, note that by Lemma 4.4 (the

(\circ)-case), every D_3 -point has two D_2 -points among its a^F -successors; these two are indistinguishable, as they are both labeled with R_0 .

Finally, clause (6) forces d_0 to occur at least once on every set G_i .

$$[a^F](\neg P' \rightarrow \langle a^f \rangle (\neg Q \wedge \neg P' \wedge \text{LRUD} = d_0)). \quad (6)$$

This is true since, for any $i \geq 1$, G_i corresponds to the set $\{f_i + f_j \mid j \neq i-1 \wedge j \neq i+1\}$, and (6) thus states that, for each $i \geq 1$, there is a $j \neq i-1$ or a $j \neq i+1$ such that the point $f_i + f_j$ is associated with d_0 . (The first $\neg P'$ in (6) is used to exclude $[a^{f_0}]$, since the recurrence condition is not required to hold for G_0 . The second $\neg P'$ is used to eliminate the possibility of choosing $\langle a^{f_i} \rangle$ in the $[a^{f_i}]$ -case.)

We can now complete the proof of Theorem 4.1 by proving:

CLAIM. φ_T is satisfiable if and only if T satisfies the property described in Proposition 4.2.

Proof. (If). Given that T can tile G as in Prop. 4.2, construct the model $M = (W, \tau, \rho)$, with $W = \{s_i \mid i \geq 0\}$ and $\rho(a) = \{(s_i, s_{i+1}) \mid i \geq 0\}$. Now, regardless of the interpretations on points outside the $a^F a^F$ grid $G' = \{s_k \mid k = f_i + f_j \text{ for some } i, j \geq 0\}$, τ will interpret the R_i on G as in Fig. 5, and will interpret the LEFT, RIGHT, UP, and DOWN combinations to encode the given tiling on G . It is now easy to see that $M, s_0 \models \varphi_T$.

(Only if). If $M, s \models \varphi_T$ for some model $M = (W, \tau, \rho)$, $s \in W$, then the $[a^*] \langle a \rangle$ true part of clause (1) forces the existence of an infinite a -path $s = s_0, s_1, s_2, \dots$ (the states need not necessarily be distinct.) Upon this path the remaining parts of φ_T force the R_i to behave as in Fig. 5 and the LEFT, RIGHT, UP, and DOWN combinations to correspond to a tiling of G with the set T , as in Proposition 4.2. ■

Received August 29, 1994; final manuscript received May 10, 1996

REFERENCES

- [FL] Fischer, M. J., and Ladner, R. E. (1979), Propositional dynamic logic of regular programs, *J. Comput. System Sci.* **18**, 194–211.
- [H1] Harel, D. (1984), Dynamic Logic, in "Handbook of Philosophical Logic" (D. Gabbay and F. Guenther, Eds.), Reidel, Dordrecht, 1984, Vol. II, pp. 497–604.
- [H2] Harel, D. (1985), Recurring dominoes: Making the highly undecidable highly understandable, *Ann. Discrete Math.* **24** (1985), 51–72.
- [HP] Harel, D. and Paterson, M. S., Undecidability of PDL with $L = \{a^{2^i} \mid i \geq 0\}$, *J. Comput. System Sci.* **29**, 359–365.
- [HPS] Harel, D., Pnueli, A., and Stavi, J. (1983), Propositional dynamic logic of nonregular programs, *J. Comput. System Sci.* **26** (1983), 222–243.

- [HR] Harel, D. and Raz, D. (1993), Deciding properties of nonregular programs, *SIAM J. Comput.* **22** (1993), 857–874.
- [HS] Harel, D. and Singerman, E. (1995), More on nonregular PDL: Expressive power, finite models, fibonacci programs, in “3rd Israel Symp. on the Theory of Comp. Sys.,” IEEE Press, New York, pp. 140–149.
- [KP] Koren, T. and Pnueli, A. (1983), There exist decidable context-free propositional dynamic logics, in “Proc. Sympo. on Logics of Programs,” Lecture Notes in Computer Science, Springer-Verlag, New York, Vol. 164, pp. 290–312.
- [KT] Kozen, D. and Tiuryn, J. (1990), Logics of programs, in “Handbook of Theoretical Computer Science” (J. Van Leeuwen, Ed.), Elsevier, Amsterdam, Vol. B, pp. 789–840.
- [P] Parikh, R. J. (1966), On context-free languages, *J. Assoc. Comput. Mach.* **13**, 570–581.
- [P1] Pratt, V. R. (1976), Semantical considerations on Floyd–Hoare logic, in “17th IEEE Symp. Found. Comput. Sci.,” IEEE Press, New York, pp. 109–121.
- [P2] Pratt, V. R. (1980), A near optimal method for reasoning about action, *J. Comput. System Sci.* **20**, 231–254.