Polynomial fingerprinting for formulas

Mihai Prunescu

FMI & IMAR

FROM 2025
Iași, România

Outline

Motivation

- Zero Knowledge Proofs for the correctness of mathematical proofs.
- Every formula must be associated with a number.
- As the proof steps must be easily performed...
- ... the encryption has to be homomorphic for modus ponens and for substitution.

Lemma Schwartz-Zippel

Theorem

Let $\mathbb{F}$ be a finite field and let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a non-zero polynomial of degree $d \geq 0$. If $r_1, r_2, \ldots, r_n$ are selected randomly and if the choices are independent in $\mathbb{F}$, then:

$$Pr[f(r_1, r_2, \ldots, r_n) = 0] \leq \frac{d}{|\mathbb{F}|}.$$

Idea

The family of polynomial matrices

$$A(k) = \begin{pmatrix} x_{4k+1} & x_{4k+2} \\ x_{4k+3} & x_{4k+4} \end{pmatrix}$$

are non-commutative to such extent that if two products are equal:

$$A(i_1)\dots A(i_n) = A(j_1)\dots A(j_m)$$

then $n = m$ and $i_1 = j_1, \dots, i_n = j_n$.

Realization

Definition

Let x_1 be a polynomial variable. Let:

$$A(x_1) = \begin{pmatrix} x_1 & 1 \\ 0 & 1 \end{pmatrix}$$

We consider that $A(x_1) \in M_{2\times 2}(\mathbb{Z}[x_1, x_2, \ldots])$.

Main property

Lemma

Consider a set of different variables $V = \{x_1, x_2, \dots, x_k\}$. *Suppose that* $0 \leq i_1, \dots, i_n, j_1, \dots, j_m \leq k$. *If:*

$$A(x_{i_1})A(x_{i_2})\dots A(x_{i_n}) = A(x_{j_1})A(x_{j_2})\dots A(x_{j_m}),$$

then the following equalities take place: $n = m$, $i_1 = j_1$, $\dots$, $i_n = j_n$.

Definition

For every specific symbol c of arity $d = d(c)$ a number of $d + 1$ different fixed edge variables $C, C_1, \dots, C_d \in \{x_1, x_2, \dots\}$ are associated.
Suppose that a tree T has root c and the sub-trees connected with c are $T_1, \dots, T_d$. Suppose that one already associated matrices

$$[T_1], \dots, [T_d] \in M_{2\times 2}(\mathbb{Z}[x_1, x_2, \dots])$$

with these sub-trees. Then we associate with T the pair:

$$[T] = A(C) + A(C_1)[T_1] + \cdots + A(C_d)[T_d],$$

where $C, C_1, \dots, C_d$ are the associated edge variables.

Unicity

Definition

If φ is a formula or a term, let $[\varphi]$ denote the polynomial matrix associated with its tree.

Theorem

A matrix represents at most one formula.

Example

1. The letters x, y, z are atomic propositional formulas.
2. If φ and ψ are formulas, then:

$$\neg\,\varphi, \;\; \varphi \to \psi,$$

   are formulas.

The alphabet is $A = \{x, y, z, \neg, \to\}$.

Example

The variables x, y, z are symbols of arity 0 and will always be final nodes. We associate them with the matrices:

$$[x] = A(X) = \begin{pmatrix} X & 1 \\ 0 & 1 \end{pmatrix},$$

$$[y] = A(Y) = \begin{pmatrix} Y & 1 \\ 0 & 1 \end{pmatrix},$$

$$[z] = A(Z) = \begin{pmatrix} Z & 1 \\ 0 & 1 \end{pmatrix}.$$

Example

The symbols with positive arity are $\{\neg, \to\}$. We associate with $\neg$ the matrices:

$$A(N) = \begin{pmatrix} N & 1 \\ 0 & 1 \end{pmatrix}, \quad A(N_1) = \begin{pmatrix} N_1 & 1 \\ 0 & 1 \end{pmatrix}.$$

We associate with $\to$ the matrices:

$$A(I) = \begin{pmatrix} I & 1 \\ 0 & 1 \end{pmatrix}, \quad A(I_1) = \begin{pmatrix} I_1 & 1 \\ 0 & 1 \end{pmatrix}, \quad A(I_2) = \begin{pmatrix} I_2 & 1 \\ 0 & 1 \end{pmatrix}.$$

The 7 variables $X, Y, Z, N, N_1, I, I_1, I_2$ are pairwise different.

Example

The inductive steps are given by:

$$[\neg\,\alpha] = A(N) + A(N_1)[\alpha],$$

$$[\alpha \to \beta] = A(I) + A(I_1)[\alpha] + A(I_2)[\beta],$$

Inverse matrix

$$A(x) = \begin{pmatrix} x & 1 \\ 0 & 1 \end{pmatrix}$$

$$A(x)^{-1} = \begin{pmatrix} x^{-1} & -x^{-1} \\ 0 & 1 \end{pmatrix}$$

$$A(x)^{-1} \notin M_{2\times 2}(\mathbb{Z}[x_1, x_2, \ldots])$$

$$x \in \mathbb{F} \setminus \{0\}, \;\; A(x)^{-1} \in M_{2\times 2}(\mathbb{F})$$

Modus ponens

$$\varphi$$
$$\varphi \to \psi$$
$$- - -$$
$$\psi$$

$$[\varphi \to \psi] = A(I) + A(I_1)[\varphi] + A(I_2)[\psi].$$

$$[\psi] = A(I_2)^{-1} \left([\varphi \to \psi] - A(I) - A(I_1)[\varphi] \right).$$

Substitution

$$[\varphi(x)] = \sum_{\text{nodes } c} A(X_{i_1}) \dots A(X_{i_n}) \cdot A(X_c).$$

The monomial $A(X_{i_1}) \dots A(X_{i_n})$ consists of the edge-variables on the path from the root to the node c.

$$[\varphi(x/\psi)] = [\varphi(x)] - A(X_{i_1}) \dots A(X_{i_n})[x] - A(X_{j_1}) \dots A(X_{j_m})[x] +$$
$$+ A(X_{i_1}) \dots A(X_{i_n})[\psi] + A(X_{j_1}) \dots A(X_{j_m})[\psi].$$

In general, let x be a propositional or a first-order element variable, and let X be the polynomial variable associated to this symbol of arity 0. Let φ be a formula or a term. We denote by:

$$\sum_{c=x} A(X_{i_1}) \dots A(X_{i_n}) \cdot A(X_c) := [\varphi]_x \cdot A(X_c).$$

It follows that in general for every formula or term ψ,

$$[\varphi(x/\psi)] = [\varphi] - [\varphi]_x \cdot A(X) + [\varphi]_x \cdot [\psi].$$

Fingerprints

Definition

Let φ be a well-formed expression over A, i.e. a term or a formula. Suppose that $x_1, \ldots, x_k$ are the free variables in φ, which may be propositional variables or first-order element variables. We call the fingerprint of φ the tuple:

$$([\varphi], [\varphi]_{x_1}, \ldots, [\varphi]_{x_k}).$$

We denote the fingerprint of φ with $F(\varphi)$.

Fingerprints

Theorem

Suppose that formulas φ and $\varphi \to \psi$ have fingerprints:

$$F(\varphi) = ([\varphi], [\varphi]_{x_1}, \ldots, [\varphi]_{x_k}),$$

$$F(\varphi \to \psi) = ([\varphi \to \psi], [\varphi \to \psi]_{x_1}, \ldots, [\varphi \to \psi]_{x_k}).$$

Then the fingerprint of ψ is:

$$F(\psi) = ([\psi], [\psi]_{x_1}, \ldots, [\psi]_{x_k}),$$

where:

$$[\psi] = A(I_2)^{-1} \left([\varphi \to \psi] - A(I) - A(I_1)[\varphi] \right),$$

$$[\psi]_{x_i} = A(I_2)^{-1} \left([\varphi \to \psi]_{x_i} - A(I_1)[\varphi]_{x_i} \right).$$

Fingerprints

Theorem

Let φ and ψ be formulas or terms. Suppose that their fingerprints are:

$$F(\varphi) = ([\varphi], [\varphi]_{x_1}, \dots, [\varphi]_{x_k}),$$

$$F(\psi) = ([\psi], [\psi]_{x_1}, \dots, [\psi]_{x_k}).$$

$$F(\varphi(x_i/\psi)) = ([\varphi(x_i/\psi)], [\varphi(x_i/\psi)]_{x_1}, \dots, [\varphi(x_i/\psi)]_{x_k})$$
$$[\varphi(x_i/\psi)] = [\varphi] - [\varphi]_{x_i} \cdot A(X_i) + [\varphi]_{x_i} \cdot [\psi],$$

and, if $j \neq i$, then:

$$[\varphi(x_i/\psi)]_{x_j} = [\varphi]_{x_j} + [\varphi]_{x_i}[\psi]_{x_j}$$

while if $j = i$, then:

$$[\varphi(x_i/\psi)]_{x_i} = [\varphi]_{x_i}[\psi]_{x_i}$$

A formalized proof

Axioms:

$$K(\alpha, \beta): \quad \alpha \to (\beta \to \alpha),$$
$$S(\alpha, \beta, \gamma): \quad (\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma)),$$
$$N(\alpha, \beta): \quad (\neg\alpha \to \neg\beta) \to (\beta \to \alpha).$$

We consider the following theorem:

Theorem

$$A \to A.$$

A formalized proof

Consider the formula $B := A \to A$. By making corresponding substitutions, we write down:

$$S(A, B, A): \quad (A \to (B \to A)) \to ((A \to B) \to (A \to A)),$$

$$K(A, B): \quad A \to (B \to A),$$

$$K(A, A): \quad A \to (A \to A),$$

At this point we observe that $K(A, A)$ is in fact:

$$K(A, A): \quad A \to B,$$

$$MP(K(A, B), S(A, B, A)) = C: \quad (A \to B) \to (A \to A),$$

$$MP(K(A, A), C): \quad A \to A.$$

We also observe that the conclusion is the same as B. $\square$

A formalized proof

$$
\begin{aligned}
[A] &= A, \\
[B] &= I + I_1[A] + I_2[A], \\
[B \to A] &= I + I_1[B] + I_2[A], \\
[K(A, B)] &= I + I_1[A] + I_2[B \to A], \\
[K(A, A)] &= I + I_1[A] + I_2[B], \\
[C] &= I + I_1[K(A, A)] + I_2[B], \\
[S(A, B, A)] &= I + I_1[K(A, B)] + I_2[C], \\
I_2^{-1}([S(A, B, A)] - I - I_1[K(A, B)]) &= [C], \\
I_2^{-1}([C] - I - I_1[K(A, A)]) &= [B].
\end{aligned}
$$

THANK

YOU!