A Coq formalization of a sign determination algorithm

TYPES – Tallinn, May 20 2015

Cyril Cohen and Mathieu Kohli

Context

Fundamental step in some algorithms in real algebraic geometry is the sign determination.

A naive sign determination algorithm has already been formalized (cf Cohen, Mahboubi, LMCS 2012.)

Our goal: formalize more efficient versions, in order to perform computations.

Example of application: *Formally-Verified Decision Procedures for Univariate Polynomial Computation Based on Sturms and Tarskis Theorems*, Narkawicz, Muoz, Dutle, JAR 2015

Statement of the problem

Knowing how to compute

$$\mathrm{TaQ}(P, Q) = \sum_{x \in \mathrm{roots}(P)} \mathrm{sign}(Q(x)),$$

Given a polynomial P and a list of n polynomials $\vec{Q}$ and a list of sign conditions $\vec{\sigma} \in \{0, 1, -1\}^n$ we want to compute:

$$\mathrm{cnt}(P, \vec{Q}, \vec{\sigma}) = |\{x \in \mathrm{roots}(P) | \forall i, \mathrm{sign}(Q_i(x)) = \sigma_i\}|,$$

using multiple calls of $\mathrm{TaQ}(P, Q^{\vec{\alpha}})$, with $\vec{Q}^{\vec{\alpha}} = \prod_i Q_i^{\alpha_i}$.

Naive solution

(*Algorithms in real algebraic geometry*, Basu, Pollack, Roy)

Trivially

$$\begin{pmatrix}\mathrm{T}(1) & \mathrm{T}(Q) & \mathrm{T}(Q^2)\end{pmatrix} = \begin{pmatrix}\mathrm{C}(Q,0) & \mathrm{C}(Q,+1) & \mathrm{C}(Q,-1)\end{pmatrix} \cdot \begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 1\\ 1 & -1 & 1\end{pmatrix}.$$

More generally,

$$\left(\mathrm{TaQ}(P, \vec{Q}^{\vec{\alpha}})\right)_{\vec{\alpha}\in\{0,1,2\}^n} = \left(\mathrm{cnt}(P, \vec{Q}, \vec{\sigma})\right)_{\vec{\sigma}\in\{0,1,-1\}^n} \cdot \begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 1\\ 1 & -1 & 1\end{pmatrix}^{\otimes n}$$

by induction on n, with appropriate generalization, cf Cohen, Mahboubi, LMCS 2012.

Efficiency issues

Given a polynomial P and a list of n polynomials $\vec{Q}$ and a list of sign conditions $\vec{\sigma} \in \{0, 1, -1\}^n$ we want to compute:

$$\mathrm{cnt}(P, \vec{Q}, \vec{\sigma}) = |\{x \in \mathrm{roots}(P) | \forall i, \mathrm{sign}(Q_i(x)) = \sigma_i\}|,$$

using multiple calls of $\mathrm{TaQ}(P, Q^{\vec{\alpha}})$, with $\vec{Q}^{\vec{\alpha}} = \prod_i Q_i^{\alpha_i}$, but:

- not too many calls, i.e. using only a small subset A of $\{0, 1, 2\}^n$,
- with small products (i.e. $|\{i | \alpha_i \neq 0\}|$ as small as possible for each $\alpha \in A$.

Non empty sign conditions

(*Algorithms in real algebraic geometry*, Basu, Pollack, Roy)

Since

$$\mathrm{cnt}(P, \vec{Q}, \vec{\sigma}) = |\{x \in \mathrm{roots}(P) | \forall i, \mathrm{sign}(Q_i(x)) = \sigma_i\}|,$$

We have

$$\sum_{\vec{\sigma} \in \{0,1,-1\}^n} \mathrm{cnt}(P, \vec{Q}, \vec{\sigma}) \leq \deg P$$

Hence, **at most** $\deg P$ sign conditions $\vec{\sigma}$ are non empty. Let's call them Σ.

Reduction of the system

(*Algorithms in real algebraic geometry*, Basu, Pollack, Roy)

We have

$$\left(\mathrm{TaQ}(P, \vec{Q}^{\vec{\alpha}})\right)_{\vec{\alpha} \in \mathrm{Ada}(\Sigma)} = \left(\mathrm{cnt}(P, \vec{Q}, \vec{\sigma})\right)_{\vec{\sigma} \in \Sigma} \cdot M(\Sigma, \mathrm{Ada}(\Sigma))$$

where

- $\mathrm{Ada}(\Sigma)$ is a subset of $\{0, 1, 2\}^n$ which depends only on Σ,
- $\mathrm{Ada}(\Sigma)$ has small products, i.e. for all $\alpha \in \mathrm{Ada}(\Sigma)$,

$$|\{i | \alpha_i \neq 0\}| \leq \log |\Sigma|$$

- $M(\Sigma, A)$ is a submatrix of the tensor product, which depends only on Σ and A. More precisely $M(\Sigma, A)_{\vec{\sigma}, \vec{\alpha}} = \vec{\sigma}^{\vec{\alpha}}$
- $M(\Sigma, \mathrm{Ada}(\Sigma))$ is invertible (in particular $|\Sigma| = |\mathrm{Ada}(\Sigma)|$)

Definition of $M(\Sigma, A)$

We have:

$$M(\Sigma, A)_{\vec{\sigma},\vec{\alpha}} = \vec{\sigma}^{\vec{\alpha}}$$

We represent it using encodings between a set `s` and the finite type `'I_#|s|` of the same cardinality as `s`.

```
Definition sign (i : 'I_3) : int :=
  match val i with 0 => 0%R | 1 => 1%R | _ => -1%R end.

Definition expo (i : 'I_3) : nat :=
  match val i with 0 => 0%N | 1 => 1%N | _ => 2%N end.

Definition mat_coef n (i : 'I_3 ^ n) (j : 'I_3 ^ n) :=
  (\prod_k (sign (i k)) ^+ (expo (j k)))%:Q%R.

Definition mat n (s : {set 'I_3 ^ n}) (a : {set 'I_3 ^ n}) :
  'M[rat]_(#|s|, #|a|) := \matrix_(i,j) mat_coef (enum_val i) (enum_val j).

Definition adapted n (s : {set 'I_3 ^ n}) (a : {set 'I_3 ^ n}) :=
  (#|s| == #|a|) && row_free (mat s a).
```

Extension and restriction

Given $\vec{\sigma} \in \{0, 1, -1\}^{n+1}$ one can take the restriction $\vec{\sigma}'$ by taking out the last component:

```
Definition restrict n X (b : X ^ n.+1) : X ^ n :=
  [ffun i => b (lift ord_max i)].
```

Given $\vec{\sigma} \in \{0, 1, -1\}^n$ and $x \in \{0, 1, -1\}$, one can form the extension $(\sigma, x) \in \{0, 1, -1\}^{n+1}$:

```
Definition extelt n X (x : X) (s : X ^ n) : X ^ n.+1 :=
  [ffun i => if unlift ord_max i is Some j then s j else x].
```

Given $\Sigma \subset \{0, 1, -1\}^n$ and $x \in \{0, 1, -1\}$, one can form the extension $(\Sigma, x) \subset \{0, 1, -1\}^{n+1}$:

```
Definition extset n X (x : X) (S : {set X ^ n}) : {set X ^ n.+1} :=
  [set extelt x s | s in S].
```

Extensions

Given $\Sigma \subset \{0, 1, -1\}^{n+1}$ and a number m, one can form the set Ξ_m of restrictions of Σ which have at least m different extensions in Σ

```
Definition Xi n X (S : {set X ^ n.+1}) (m : nat) :=
  [set s : X ^ n | [exists E : {set X}, (#|E| == m)
                  && [forall x in E, extelt x s \in S]]].
```

Given $\Sigma \subset \{0, 1, -1\}^n$ and an elements $\vec{\sigma}$, one can form the set of all possible extensions in Σ.

```
Definition exts X n (S : {set X ^ n.+1}) (s : X ^ n) :=
  [set (x : X ^ n.+1) ord_max | x in S & restrict x == s].


Lemma card_extsP (X : finType) n (S : {set X ^ n.+1}) (s : X ^ n) m :
  (s \in Xi S m) = (m <= #|exts S s|).
```

Adapted family

The adapted family $\mathrm{Ada}(\Sigma)$ is defined recursively as the disjoint union of $(\Xi_1, 0)$, $(\Xi_2, 1)$ and $(\Xi_3, 2)$.

```
Fixpoint adapt n (S : {set 'I_3 ^ n}) : {set 'I_3 ^ n} :=
  match n return {set 'I_3 ^ n} -> {set 'I_3 ^ n} with
    | 0 =>    fun S => S
    | n'.+1 => fun S => \bigcup_(i : 'I_3) extset i (adapt (Xi S i.+1))
  end S.
```

We prove the union is disjoint:

```
Lemma partition_adapt n (S : {set 'I_3 ^ n.+1}) :
  partition [set extset i (adapt (Xi S (i : 'I_3).+1))
          | i in 'I_3 & Xi S i.+1 != set0] (adapt S).
```

Intermediate results

```
Lemma Xi_monotonic n (X : finType) (S S' : {set X ^ n.+1}) m :
  S \subset S' -> Xi S m \subset Xi S' m.

Lemma leq_Xi n (X : finType) (S : {set X ^ n.+1}) :
  {homo Xi S : m p / (p <= m)%N >-> m \subset p}.

Lemma adapt_monotonic n (S S' : {set 'I_3 ^ n}) :
  S \subset S' -> adapt S \subset adapt S'.

Lemma adapt_down_closed n (S : {set 'I_3 ^ n}) (a b : Expos n) :
  (forall i, b i <= a i)%N -> a \in adapt S -> b \in adapt S.

Lemma partition_Signs n (S : {set 'I_3 ^ n.+1}) :
  partition [set reext S (i : 'I__) | i in 'I_3 & Xi S i.+1 != set0] S.
```

Main proofs

Completed:

```
Lemma prop1084 n (S : {set 'I_3 ^ n}) a :
 a \in adapt S -> 2 ^ #|[set i : 'I_n | a i != 0%R]| <= #|S|.

Lemma card_adapt n (S : {set 'I_3 ^ n}) : #|adapt S| = #|S|.
```

Ongoing:

```
Lemma adapt_adapted n (S : {set 'I_3 ^ n}) : adapted S (adapt S).
```

Difficulties

Encountered

- A lot of reindexing (kept implicit in the book)
- Many different partitioning of the same set (kept implicit in the book).

Avoided (so far):

- Using matrices with judgmentally different but propositionally identical indexes.
- Set extensionality problems, thanks to finite sets.

Conclusions

- The new formal proof of `prop1084` and the intermediate lemmas was backported to the future revision of the book.
- The new paper proof of `adapt_adapted` contains a pseudo-recurrence which was not in the first version.

We want to prove that all λ_τ are zero.

If $\sigma \in \mathrm{SIGN}(\mathcal{Q}, Z)_3$, we denote by $\sigma_1 <_{\mathrm{lex}} \sigma_2 <_{\mathrm{lex}} \sigma_3$ the sign conditions of $\mathrm{SIGN}(\mathcal{P}, Z)$ extending σ.

Similarly, if $\sigma \in \mathrm{SIGN}(\mathcal{Q}, Z)_2 \setminus \mathrm{SIGN}(\mathcal{Q}, Z)_3$, we denote by

$$\sigma_1 <_{\mathrm{lex}} \sigma_2$$

the sign conditions of $\mathrm{SIGN}(\mathcal{P}, Z)$ extending σ.

Finally if $\sigma \in \mathrm{SIGN}(\mathcal{Q}, Z) \setminus \mathrm{SIGN}(\mathcal{Q}, Z)_2$, we denote by σ_1 the sign condition of $\mathrm{SIGN}(\mathcal{P}, Z)$ extending σ.

Since by induction hypothesis, the matrix

$$\mathrm{Mat}(\mathrm{Ada}(\mathcal{Q}, Z), \mathrm{SIGN}(\mathcal{Q}, Z))$$

is invertible,

$$\begin{aligned} \lambda_{\sigma_1} &= 0, && \text{for } \sigma \in \mathrm{SIGN}(\mathcal{Q}, Z) \setminus \mathrm{SIGN}(\mathcal{Q}, Z)_2, \\ \lambda_{\sigma_1} + \lambda_{\sigma_2} &= 0, && \text{for } \sigma \in \mathrm{SIGN}(\mathcal{Q}, Z)_2 \setminus \mathrm{SIGN}(\mathcal{Q}, Z)_3, \\ \lambda_{\sigma_1} + \lambda_{\sigma_2} + \lambda_{\sigma_3} &= 0, && \text{for } \sigma \in \mathrm{SIGN}(\mathcal{Q}, Z)_3. \end{aligned}$$

Thus $\lambda_{\sigma_1} = 0$ for every $\sigma \in \mathrm{SIGN}(\mathcal{Q}, Z) \setminus \mathrm{SIGN}(\mathcal{Q}, Z)_2$.

Using again the induction hypothesis, the matrix

$$\mathrm{Mat}(\mathrm{Ada}(\mathcal{Q}, Z_2), \mathrm{SIGN}(\mathcal{Q}, Z_2))$$

is invertible, so

$$\begin{aligned} \sigma_1(P)\, \lambda_{\sigma_1} - \sigma_2(P)\, \lambda_{\sigma_2} &= 0, && \text{for } \sigma \in \mathrm{SIGN}(\mathcal{Q}, Z)_2 \setminus \mathrm{SIGN}(\mathcal{Q}, Z)_3, \\ \lambda_{\sigma_2} - \lambda_{\sigma_3} &= 0, && \text{for } \sigma \in \mathrm{SIGN}(\mathcal{Q}, Z)_3. \end{aligned}$$

Thus $\lambda_{\sigma_1} = \lambda_{\sigma_2} = 0$, for every $\sigma \in \mathrm{SIGN}(\mathcal{Q}, Z)_2 \setminus \mathrm{SIGN}(\mathcal{Q}, Z)_3$.

Finally, using once more the induction hypothesis, the matrix

$$\mathrm{Mat}(\mathrm{Ada}(\mathcal{Q}, Z_3), \mathrm{SIGN}(\mathcal{Q}, Z_3))$$

...

i) We prove the statement in the special case when $\Xi^{(0)} = \Xi^{(2)}$. The statement is clear since $\mathrm{Mat}(\mathrm{Ada}(\Sigma), \Sigma)$ is the tensor product of $\mathrm{Mat}(\mathrm{Ada}(\Xi^{(2)}), \Xi^{(2)})$ by

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix},$$

and is invertible, because $\mathrm{Mat}(\mathrm{Ada}(\Xi^{(2)}), \Xi^{(2)})$ is.

ii) We now prove the statement in the special case when $\Xi^{(0)} = \Xi^{(1)}$. We first prove that all λ_τ for τ having a restriction to $\Xi^{(1)} \setminus \Xi^{(2)}$ are zero. Indeed, using that $\mathrm{Mat}(\mathrm{Ada}(\Xi^{(1)}), \Xi^{(1)})$ is invertible

(a) $\lambda_{\sigma_1} + \lambda_{\sigma_2} = 0$ and $\sigma_1(P_i)\lambda_{\sigma_1} + \sigma_2(P_i)\lambda_{\sigma_2} = 0$, for every $\sigma \in \Xi^{(1)} \setminus \Xi^{(2)}$,
(b) $\lambda_{\sigma_1} + \lambda_{\sigma_2} + \lambda_{\sigma_3} = 0$ and $\lambda_{\sigma_2} - \lambda_{\sigma_3} = 0$, for every $\sigma \in \Xi^{(2)}$.

We conclude by i), removing from Σ the elements having a restriction in $\Xi^{(1)} \setminus \Xi^{(2)}$.

iii) We finally prove the statement for a general $\Xi^{(0)}$. We want to prove that all λ_τ are zero. We first prove that all λ_τ for τ having a restriction in $\Xi^{(0)} \setminus \Xi^{(1)}$ are zero. Indeed, using that $\mathrm{Mat}(\mathrm{Ada}(\Xi^{(0)}), \Xi^{(0)})$ is invertible,

(a) $\lambda_{\sigma_1} = 0$, for every $\sigma \in \Xi^{(0)} \setminus \Xi^{(1)}$,
(b) $\lambda_{\sigma_1} + \lambda_{\sigma_2} = 0$, for every $\sigma \in \Xi^{(2)} \setminus \Xi^{(1)}$,
(c) $\lambda_{\sigma_1} + \lambda_{\sigma_2} + \lambda_{\sigma_3} = 0$, for every $\sigma \in \Xi^{(2)}$.

We conclude by ii), removing from Σ the elements having a restriction in $\Xi^{(0)} \setminus \Xi^{(1)}$. □

Future work

- Finish `adapt_adapted`
- Reintegration into the previous development.
- Efficient computation using refinements.

Thanks for your attention