Equational Unification and Symbolic Reachability in Maude

Santiago Escobar
Valencian Research Institute for Artificial Intelligence (VRAIN)
Universitat Politècnica de València Spain

SRI International®

Outline

Outline

Why rewriting logic?

1. Models and formal specification are easily written in Maude (simplicity, expressiveness, and performance)
2. Rewriting modulo associativity, commutativity and identity
3. Differentiation between concurrent and functional fragments of a model
4. Order-sorted and parameterized specifications
5. Infrastructure for formal analysis and verification (including search command, LTL model checker, theorem prover, etc.)
6. Reflection (meta-modeling, symbolic execution, building tools)
7. Application areas:
   - Models of computation (λ-calculi, π-calculus, petri nets, CCS),
   - Programming languages (C, Java, Haskell, Prolog),
   - Distributed algorithms and systems (security protocols, real-time, probabilistic),
   - Biological systems

Why adding logical features to Rewriting Logic?

1. Logical features were included in preliminary designs of the language (80's) but never implemented in Maude
2. Automated reasoning capabilities by adding logical variables
3. Differentiation between concurrent and functional fragments of a model is lifted to differentiation between symbolic models and equational reasoning.
4. Unification and Narrowing modulo combinations of A,C,U
5. Infrastructure for formal analysis and verification lifted:
   - from equational reduction to equational unification,
   - from search to symbolic reachability,
   - from LTL model checker to logical LTL model checker,
   - from theorem proving to narrowing-based theorem proving,
   - from SMT solving to variant-based SMT solving.

   ⋮

Outline

What have we done!!

- Maude 2.4 (2009)
  - Built-in Unification: free or associative-commutative (AC)
  - Narrowing-based search: rules modulo axioms (no equations).
- Maude 2.6 (2011)
  - Built-in Unification: free, C, AC, or ACU (AC + identity)
  - Variant Unification: Restricted equations modulo axioms.
  - Narrowing-based search: rules modulo equations and axioms.
- Maude 2.7 (2015)
  - Built-in Unification: free, C, AC, or ACU, CU, U, UI, Ur
  - Built-in Variant unification: wide class of equational theories.
  - Narrowing-based search: rules modulo equations and axioms.
- Maude 2.7.1 (2016)
  - Built-in Unification: previous cases + associativity
- Maude 3.0 (2019) Built-in Narrowing-based search: modulo all combinations
- Maude 3.1 (2020) Minimal (equational) unifiers, better unification modulo associavity
- Maude 3.2 (Now) Built-in Narrowing-based search with minimal unifiers

Outline

Rewriting logic in a nutshell

A rewrite theory is

$\mathcal{R} = (\Sigma, Ax \uplus E, R)$, with:

1. (Σ, R) a set of rewrite rules of the form $t \to s$
   (i.e., system transitions)
2. $(\Sigma, Ax \uplus E)$ a set of equational properties of the form $t = s$
   (i.e., E are equations and Ax are axioms such as ACU)

Intuitively, $\mathcal{R}$ specifies a concurrent system, whose states are elements of the initial algebra $T_{\Sigma/(Ax \uplus E)}$ specified by $(\Sigma, Ax \uplus E)$, and whose concurrent transitions are specified by the rules R.

Rewriting logic in a nutshell

```
mod VENDING-MACHINE is
  sorts Coin Item Marking Money State .
  subsort Coin < Money .
  op empty :  -> Money .
  op __ :  Money Money -> Money [assoc comm id:  empty] .
  subsort Money Item < Marking .
  op __ :  Marking Marking -> Marking [assoc comm id:  empty] .
  op <_> :  Marking -> State .
  ops $ q :  -> Coin .
  ops cookie cap :  -> Item .
  var M : Marking .
  rl [add-$] :  < M > => < M $ > .
  rl [add-q] :  < M > => < M q > .
  rl [buy-c] :  < M $ > => < M cap > .
  rl [buy-a] :  < M $ > => < M cookie q > .
  eq [change]:  q q q q = $ [variant] .
endm
```

Rewriting modulo

Rewriting is

Given $(\Sigma, Ax \uplus E, R)$, $t \rightarrow_{R,(Ax \uplus E)} s$ if there is

- a non-variable position $p \in Pos(t)$;
- a rule $l \rightarrow r$ in R;
- a matching σ (E-normalized and modulo Ax) such that $t|_p =_{(Ax \uplus E)} \sigma(l)$, and $s = t[\sigma(r)]_p$.

Ex: `< $ q q q >` $\rightarrow$ `< $ cookie >`
using "`rl < M $ > => < M cookie q > .`"
modulo AC of symbol "`__`"

Ex: `< q q q q >` $\rightarrow$ `< cap >`
using "`rl < M $ > => < M cap > .`"
modulo simplification with `q q q q = $` and AC of symbol "`__`"

Narrowing modulo

Narrowing is

Given $(\Sigma, Ax \uplus E, R)$, $t \leadsto_{\sigma,R,(Ax \uplus E)} s$ if there is

- a non-variable position $p \in Pos(t)$;
- a rule $l \to r$ in R;
- a unifier σ (E-normalized and modulo Ax) such that $\sigma(t|_p) =_{(Ax \uplus E)} \sigma(l)$, and $s = \sigma(t[r]_p)$.

Ex: `< X q q >` $\leadsto$ `< $ cookie >`
using "`rl < M $ > => < M cookie q > .`"
using substitution $\{X \mapsto$ `$ q`$\}$ modulo AC of symbol "`__`"

Ex: `< X q q >` $\leadsto$ `< cap >`
using "`rl < M $ > => < M cap > .`"
using substitution $\{X \mapsto$ `q q`$\}$
modulo simplification with `q q q q = $` and AC of symbol "`__`"

Outline

Unification modulo axioms

Definition

Given equational theory (Σ, Ax), an Ax-unification problem is

$$t \stackrel{?}{=} t'$$

An Ax-unifier is an order-sorted substitution σ s.t.

$$\sigma(t) =_{Ax} \sigma(t')$$

Decidability

- at most one mgu (syntactic unification, i.e., empty theory)
- a finite number (associativity–commutativity)
- an infinite number (associativity)

Admissible Theories

Maude provides order-sorted Ax-unification algorithm for all order-sorted theories $(\Sigma, E \cup Ax, R)$ s.t. Σ is preregular modulo Ax and axioms Ax are:

1. arbitrary function symbols and constants with no attributes;
2. `iter` equational attribute declared for some unary symbols;
3. "`comm`", "`assoc`", "`assoc comm`", "`assoc comm id:`", "`comm id:`", "`assoc id:`", "`id:`", "`left id:`", or "`right id:`" attributes declared for some binary function symbols but no other equational attributes can be given for such symbols.

Unification Command in Maude

Maude provides a Ax-unification command of the form:

$$\texttt{unify [} n \texttt{] in } \langle ModId \rangle \texttt{ :}$$
$$\langle Term\text{-}1 \rangle \texttt{ =? } \langle Term'\text{-}1 \rangle \texttt{ /\ ... /\ } \langle Term\text{-}k \rangle \texttt{ =? } \langle Term'\text{-}k \rangle \texttt{ .}$$
$$\texttt{irredundant unify [} n \texttt{] in } \langle ModId \rangle \texttt{ :}$$
$$\langle Term\text{-}1 \rangle \texttt{ =? } \langle Term'\text{-}1 \rangle \texttt{ /\ ... /\ } \langle Term\text{-}k \rangle \texttt{ =? } \langle Term'\text{-}k \rangle \texttt{ .}$$

- ModId is the name of the module
- n is a bound on the number of unifiers
- new variables are created as `#n:Sort`
- Implemented at the core level of Maude (C++)

AC-Unification in Maude

```
Maude> unify [100] in NAT :
         X:Nat + X:Nat + Y:Nat =? A:Nat + B:Nat + C:Nat .

Solution 1
X:Nat --> #1:Nat + #2:Nat + #3:Nat + #5:Nat + #6:Nat + #8:Nat
Y:Nat --> #4:Nat + #7:Nat + #9:Nat
A:Nat --> #1:Nat + #1:Nat + #2:Nat + #3:Nat + #4:Nat
B:Nat --> #2:Nat + #5:Nat + #5:Nat + #6:Nat + #7:Nat
C:Nat --> #3:Nat + #6:Nat + #8:Nat + #8:Nat + #9:Nat
...

Solution 100
X:Nat --> #1:Nat + #2:Nat + #3:Nat + #4:Nat
Y:Nat --> #5:Nat
A:Nat --> #1:Nat + #1:Nat + #2:Nat
B:Nat --> #2:Nat + #3:Nat
C:Nat --> #3:Nat + #4:Nat + #4:Nat + #5:Nat
```

ACU-Unification in Maude

```
Maude> unify [100] in QID-SET : X:QidSet , X:QidSet , Y:QidSet =? A:QidSet , B:QidSet , C:QidSet .
unify [100] in QID-SET : X:QidSet, X:QidSet, Y:QidSet =? A:QidSet, B:QidSet, C:QidSet .
Decision time: 0ms cpu (1ms real)

Solution 1
X:QidSet --> empty
Y:QidSet --> empty
A:QidSet --> empty
B:QidSet --> empty
C:QidSet --> empty

Solution 2
X:QidSet --> #1:QidSet
Y:QidSet --> empty
A:QidSet --> #1:QidSet, #1:QidSet
B:QidSet --> empty
C:QidSet --> empty
```

Irredundant Unification in Maude

```
Maude> unify in UNIF-VENDING-MACHINE :
        < q q X:Marking > =? < $ Y:Marking > .

Unifier 1
X:Marking --> $
Y:Marking --> q q

Unifier 2
X:Marking --> $ #1:Marking
Y:Marking --> q q #1:Marking


Maude> irredundant unify in UNIF-VENDING-MACHINE :
        < q q X:Marking > =? < $ Y:Marking > .

Unifier 1
X:Marking --> $ #1:Marking
Y:Marking --> q q #1:Marking
```

Outline

From equational reduction to variants (1/4)

E,Ax-variant

Given a term t and an equational theory $Ax \uplus E$, (t', θ) is an E,Ax-variant of t if $\theta(t)\downarrow_{E,Ax} =_{Ax} t'$ [Comon-Delaune-RTA05]

Exclusive Or

$$X \oplus 0 \to X \qquad X \oplus (Y \oplus Z) = (X \oplus Y) \oplus Z$$
$$X \oplus X \to 0 \qquad X \oplus Y = Y \oplus X$$
$$X \oplus X \oplus Y \to Y \qquad \text{(axioms: } Ax\text{)}$$

Computed Variants

For $X \oplus X$: $(0, id), (0, \{X \mapsto a\}), (0, \{X \mapsto a \oplus b\}), \ldots$

From equational reduction to variants (2/4)

Finite and complete set of E,Ax-variants

A preorder relation of generalization between variants provides a notion of most general variant.

Computed Variants

For $X \oplus Y$ there are 7 most general E,Ax-variants

1. $(X \oplus Y, id)$
2. $(0, \{X \mapsto U, Y \mapsto U\})$
3. $(Z, \{X \mapsto 0, Y \mapsto Z\})$
4. $(Z, \{X \mapsto Z \oplus U, Y \mapsto U\})$
5. $(Z, \{X \mapsto Z, Y \mapsto 0\})$
6. $(Z, \{X \mapsto U, Y \mapsto Z \oplus U\})$

From equational reduction to variants (3/4)

Finite Variant Property

Theory has FVP if finite number of most general variants for every term.

Common

- Cryptographic Security Protocols: Public or shared encryption, Exclusive Or, Abelian groups, Diffie-Hellman, Bilinear Pairings
- Satisfiability Modulo Theories Natural Presburger Arithmetic, Integer Presburger Arithmetic, Lists, Sets

Used in application areas

Equational Unification, Logical Model Checking, Cyber-Physical systems, Partial evaluation, Confluence tools, Termination tools, Theorem provers

From equational reduction to variants (4/4)

Test for FVP

Whether a theory has FVP is undecidable in general, though there are approximations techniques.

Computing most general variants

Given a theory that has FVP, it is possible to compute all the most general variants by using the Folding Variant Narrowing Strategy (Escobar et al. 2012)

Variant Command in Maude

Maude provides variant generation:

$$\texttt{get variants [} n \texttt{] in } \langle ModId \rangle \texttt{ : } \langle Term \rangle \texttt{ .}$$
$$\texttt{get irredundant variants [} n \texttt{] in } \langle ModId \rangle \texttt{ : } \langle Term \rangle \texttt{ .}$$

- ModId is the name of the module
- n is a bound on the number of variants
- new variables are created as `#n:Sort` and `%n:Sort`
- Implemented at the core level of Maude (C++)
- Folding variant narrowing strategy is used internally
- Terminating if Finite Variant Property
- Incremental output if not Finite Variant Property
- Irredundant version only if Finite Variant Property

Exclusive-or Variants

```
fmod EXCLUSIVE-OR is
  sorts Nat NatSet .  subsort Nat < NatSet .
  op 0 : -> Nat .
  op s : Nat -> Nat .
  op mt : -> NatSet .
  op _*_ : NatSet NatSet -> NatSet [assoc comm] .
  vars X Z : [NatSet] .
  eq [idem] :     X * X = mt     [variant] .
  eq [idem-Coh] : X * X * Z = Z  [variant] .
  eq [id] :       X * mt = X     [variant] .
endfm
```

```
Maude> get variants in EXCLUSIVE-OR : X * Y .
Variant 1                                          Variant 7
[NatSet]: #1:[NatSet] * #2:[NatSet]   .........   [NatSet]: %1:[NatSet]
X --> #1:[NatSet]                                  X -->  %1:[NatSet]
Y --> #2:[NatSet]                                  Y --> mt
```

Abelian Group Variants

```
fmod ABELIAN-GROUP is
  sorts Elem .
  op _+_ : Elem Elem -> Elem [comm assoc] .
  op -_ : Elem -> Elem .
  op 0 : ->  Elem .
  vars X Y Z : Elem  .
  eq X + 0 = X [variant] .
  eq X + (- X) = 0 [variant] .
  eq X + (- X) + Y = Y [variant] .
  eq - (- X) = X [variant] .
  eq - 0 = 0 [variant] .
  eq (- X) + (- Y) = -(X + Y) [variant] .
  eq -(X + Y) + Y = - X [variant] .
  eq -(- X + Y) = X + (- Y) [variant] .
  eq (- X) + (- Y) + Z  = -(X + Y) + Z [variant] .
  eq -(X + Y) + Y + Z = (- X) + Z [variant] .
endfm
```

```
Maude> get variants in ABELIAN-GROUP : X + Y .
Variant 1                                    Variant 47
Elem: #1:Elem + #2:Elem    ..............    Elem: - (%2:Elem + %3:Elem)
X --> #1:Elem                                X --> %4:Elem + - (%1:Elem + %2:Elem)
Y --> #2:Elem                                Y --> %1:Elem + - (%3:Elem + %4:Elem)
```

Outline

Admissible Theories

Maude provides order-sorted $Ax \uplus E$-unification algorithm for all order-sorted theories $(\Sigma, Ax, \vec{E})$ s.t.

1. Maude has an Ax-unification algorithm,
2. E equations specified with the `eq` and `variant` keywords.
3. E is unconditional, convergent, sort-decreasing and coherent modulo Ax.
4. The `owise` feature is not allowed.

Equational Unification Command in Maude

Maude provides a $(Ax \uplus E)$-unification command of the form:

variant unify [n **] in** $\langle ModId \rangle$ **:**
$\langle Term\text{-}1 \rangle$ **=?** $\langle Term'\text{-}1 \rangle$ **/** ... **/** $\langle Term\text{-}k \rangle$ **=?** $\langle Term'\text{-}k \rangle$.
filtered variant unify [n **] in** $\langle ModId \rangle$ **:**
$\langle Term\text{-}1 \rangle$ **=?** $\langle Term'\text{-}1 \rangle$ **/** ... **/** $\langle Term\text{-}k \rangle$ **=?** $\langle Term'\text{-}k \rangle$.

- ModId is the name of the module
- n is a bound on the number of unifiers
- new variables are created as `#n:Sort` and `%n:Sort`
- Implemented at the core level of Maude (C++)
- Terminating if Finite Variant Property
- Incremental output if not Finite Variant Property

Filtered Variant-based Unification in Maude

```
Maude> variant unify in VARIANT-VENDING-MACHINE :
        < q q X:Marking > =? < $ Y:Marking > .

Unifier 1
X:Marking --> $ %1:Marking
Y:Marking --> q q %1:Marking

Unifier 2
X:Marking --> q q #1:Marking
Y:Marking --> #1:Marking

Maude> filtered variant unify in VARIANT-VENDING-MACHINE :
        < q q X:Marking > =? < $ Y:Marking > .

Unifier 1
X:Marking --> q q #1:Marking
Y:Marking --> #1:Marking
```

Outline

Symbolic reachability analysis in rewrite theories

- Given $(\Sigma, E \cup Ax, R)$ as a concurrent system, a symbolic reachability problem is

$$(\exists X)\ t \longrightarrow^* t'$$

- Narrowing provides a sound and complete method for topmost theories.
- Narrowing with R modulo $Ax \uplus E$ requires $Ax \uplus E$-unification at each narrowing step
- Narrowing can be also used for logical model checking

Narrowing in Maude

Narrowing generalizes term rewriting by allowing free variables in terms and by performing unification instead of matching in order to (non–deterministically) reduce a term.

1. Narrowing + simplification (for built-in operators and equational simplification)
2. Frozen arguments, similar to the context-sensitive narrowing
3. Extra variables in right hand sides of the rules for functional logic programming features (e.g. constraint programming and instantiation search).

Narrowing Search Command in Maude

Narrowing-based search command of the form:

`vu-narrow [` n`,` m `] in` $\langle ModId \rangle$ `:` $\langle Term\text{-}1 \rangle$ $\langle SearchArrow \rangle$ $\langle Term\text{-}2 \rangle$ `.`

- n is the bound on the desired reachability solutions
- m is the maximum depth of the narrowing tree
- `Term-1` is not a variable but may contain variables
- `Term-2` is a pattern to be reached
- `SearchArrow` is either =>1, =>+, =>*, =>!
- =>! denotes strongly irreducible terms or rigid normal forms.
- Implemented at the core level of Maude (C++)
- "`vu-narrow {filter}`" for filtered variant unification

Variant-based unification in Narrowing Search Command

```
mod NARROWING-VENDING-MACHINE is
  sorts Coin Item Marking Money State .
  subsort Coin < Money .
  op empty : -> Money .
  op __ : Money Money -> Money [assoc comm id: empty] .
  subsort Money Item < Marking .
  op __ : Marking Marking -> Marking [assoc comm id: empty] .
  op <_> : Marking -> State .
  ops $ q : -> Coin .
  ops a c : -> Item .
  var M : Marking .
  rl [buy-c] : < M $ > => < M c > [narrowing] .
  rl [buy-a] : < M $ > => < M a q > [narrowing] .
  eq [change] : q q q q M = $ M [variant] .
endm

Maude> vu-narrow [1] in NARROWING-VENDING-MACHINE : < M:Money > =>* < a c > .

Solution 1
state: < a c #1:Money >
accumulated substitution:
M:Money --> $ q q q #1:Money
variant unifier:
#1:Money --> empty
```

Variant-based unification in Narrowing Search Command

```
mod AG-VENDING is
  sorts Item Items State Coin Money .
  subsort Item < Items . subsort Coin < Money .
  op __ : Items Items -> Items [assoc comm id: mt] .
  op <_|_> : Money Items -> State .
  ops a c : -> Item . ops q $ : -> Coin .
  rl < M:Money | I:Items > => < M:Money + - $          | I:Items c > [narrowing] .
  rl < M:Money | I:Items > => < M:Money + - q + - q + - q | I:Items a > [narrowing] .
  eq $ = q + q + q + q [variant] . --- Property of the original vending machine example
  op _+_ : Money Money -> Money [comm assoc] .
  op -_ : Money -> Money .
  op 0 : -> Money .
  vars X Y Z : Money .
  ... (here come the variant equations shown before for Abelian Group)
endm

Maude> vu-narrow [1] in AG-VENDING : < M:Money | mt > =>* < 0 | a c > .
Solution 1
rewrites: 32032 in 247478ms cpu (272327ms real) (129 rewrites/second)
state: < %1:Money + - (q + q + q + q + q + q + q) | a c >
accumulated substitution:
M:Money --> %1:Money
variant unifier:
%1:Money --> q + q + q + q + q + q + q

Maude> vu-narrow {filter} [1] in AG-VENDING : < M:Money | mt > =>* < 0 | a c > .
Solution 1
rewrites: 510 in 236ms cpu (274ms real) (2160 rewrites/second)
state: < %1:Money + - (q + q + q + q + q + q + q) | a c >
accumulated substitution:
M:Money --> %1:Money
variant unifier:
%1:Money --> q + q + q + q + q + q + q
```

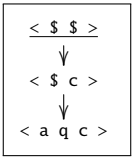
Outline

Model Checking

- Model checking techniques effective in verification of concurrent systems
- However, standard techniques only work for:
  - specific initial state (or finite set of initial states)
  - the set of states reachable from the initial state is finite
  - abstraction techniques
- Various model checking techniques for infinite-state systems exist, but they are less developed
  - Stronger limitations on the kind of systems and/or the properties that can be model checked

VENDING Example (1/6)

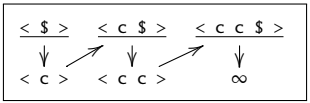
Terminating theory without rules adding money ($ and q).



(one initial state - finite space)

VENDING Example (2/6)

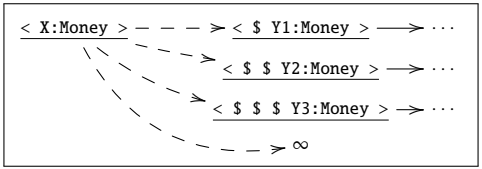
Non-terminating theory with rules adding money ($ and q).



(one initial state - infinite space)

VENDING Example (3/6)

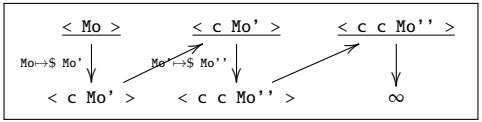
Instantiation is another source of infinity.



(infinite number of initial states)

VENDING Example (4/6)

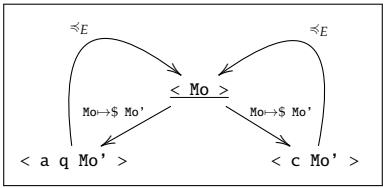
Narrowing usually provides an infinite space due to instantiation even for terminating theories (e.g. without rules adding money ($ and q)).



(one initial state - infinite space)

VENDING Example (5/6)

Narrowing-based state space can be treated in new ways and folded into a finite space in many cases



Narrowing + folding relation $\Rightarrow$ (multiple initial states - finite space)
(equality $=_E$)
(renaming $\approx_E$)
(instantiation $\preccurlyeq_E$)

VENDING Example (6/6)

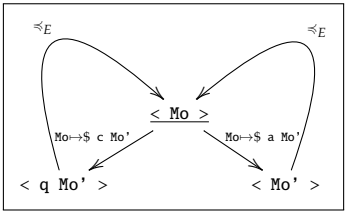
```
Maude> fvu-narrow in NARROWING-VENDING-MACHINE : < M:Marking > =>* < a c > .

Solution 1
state: < #1:Marking >
accumulated substitution:
M:Marking --> #1:Marking
variant unifier:
#1:Marking --> a c

No more solutions.
```

FVU-VENDING Example

```
mod FOLDING-NARROWING-VENDING-MACHINE is
  sorts Coin Item Marking Money State .
  subsort Coin < Money .
  op empty : -> Money .
  op __ : Money Money -> Money [assoc comm id: empty] .
  subsort Money Item < Marking .
  op __ : Marking Marking -> Marking [assoc comm id: empty] .
  op <_> : Marking -> State .
  ops $ q : -> Coin .
  ops a c : -> Item .
  var M : Marking .
  rl [buy-c] : < M $ c > => < M > [narrowing] .
  rl [buy-a] : < M $ a > => < M q > [narrowing] .
  eq [change] : q q q q M = $ M [variant] .
endm
```




```
Maude> fvu-narrow in FOLDING-NARROWING-VENDING-MACHINE : < M:Marking a c > =>* < empty > .

Solution 1
state: < #1:Marking >
accumulated substitution:
M:Marking --> $ q q q #1:Marking
variant unifier:
#1:Marking --> empty
```

Outline

Applications

- Variant-based unification itself
- Formal reasoning tools :
  - Relying on unification capabilities:
    - termination proofs
    - proofs of local confluence and coherence
  - Relying on narrowing capabilities:
    - narrowing-based theorem proving
    - testing
- Logical model checking (model checking with logical variables)
- Cryptographic protocol analysis:
  - the Maude-NPA tool (narrowing + unification in Maude)
  - the Tamarin and AKISS protocol analyzers also use Maude capabilities
- Program transformation: partial evaluation, slicing
- SMT based on narrowing or by variant generation.

Thank you!

More information in the Maude webpage.