Combining Theories Sharing Set Operations

Ruzica Piskac

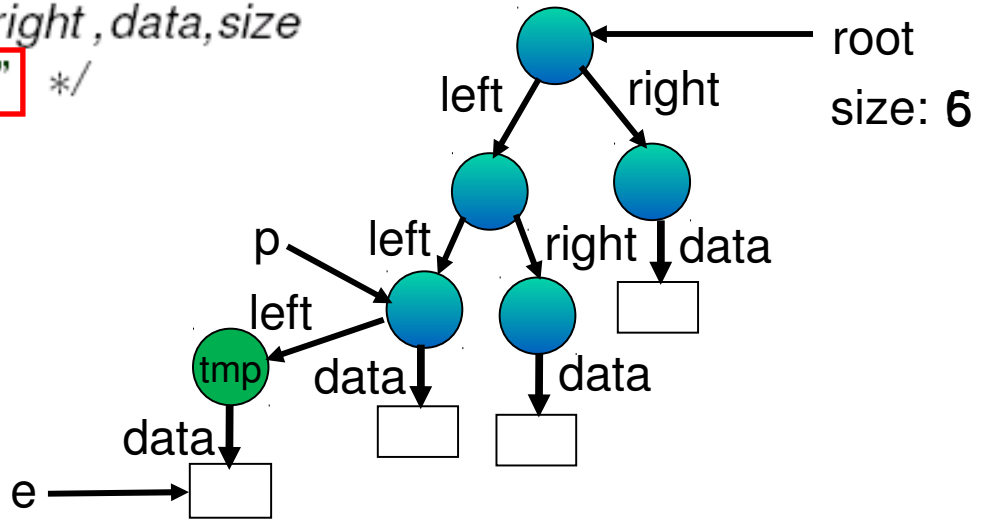
joint work with
Thomas Wies and Viktor Kuncak

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

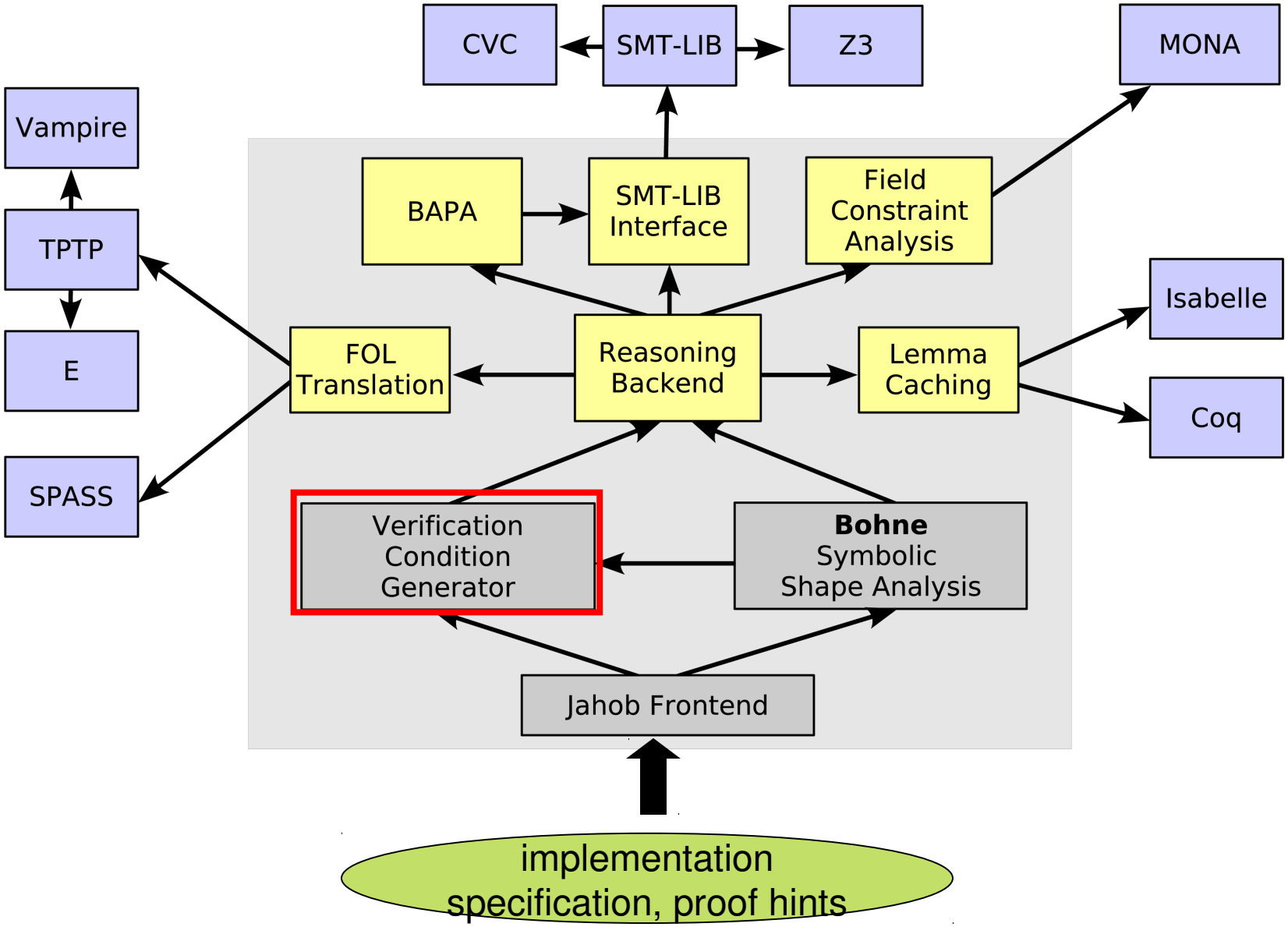
Fragment of Insertion into Tree

```
class Node {Node left,right; Object data;}
class Tree {
    private static Node root;
    private static int size; /*:
    private static specvar nodes :: objset;
    vardefs "nodes=={x. (root,x) ∈ {(x,y). left x = y ∨ right x = y}* }";
    private static specvar content :: objset;
    vardefs "content=={x. ∃ n. n ≠ null ∧ n ∈ nodes ∧ data n = x} " */

    private void insertAt(Node p, Object e) /*:
      requires "tree [ left , right ] ∧ nodes ⊆ Object.alloc ∧ size = card content ∧
                e ∉ content ∧ e ≠ null ∧ p ∈ nodes ∧ p ≠ null ∧ left p = null"
      modifies nodes,content,left, right ,data,size
      ensures "size = card content" */
    {
        Node tmp = new Node();
        tmp.data = e;
        p. left = tmp;
        size = size + 1;
    }
}
```



Program Verification with Jahob



Generated Verification Condition

$$\neg next0^*(root0,n) \wedge x \notin \{data0(v) \mid next0^*(root0,v)\} \wedge next = next0[n:=root0] \wedge data = data0[n:=x] \rightarrow |\{data(v) \,.\, next^*(n,v)\}| = |\{data0(v) \,.\, next0^*(root0,v)\}| + 1$$

"The number of stored objects has increased by one."

Expressing this VC requires a rich logic

- transitive closure * (in lists and also in trees)
- unconstraint functions (data, data0)
- cardinality operator on sets | ... |

Is there a decidable logic containing all this?

Outline

I. **Idea of decision procedure:**
reduction to a shared theory of sets

II. **BAPA-reducible theories**

III. **BAPA-reduction for WS1S**

Decomposing the Formula

Consider a (simpler) formula

$$|\{data(x).\ next^*(root,x)\}|=k+1$$

Introduce fresh variables denoting sets:

$$A = \{x.\ next^*(root,x)\} \wedge$$
$$B = \{y.\ \exists\ x.\ data(x,y) \wedge x \in A\} \wedge$$
$$|B|=k+1$$

1) WS2S
2) C^2
3) BAPA

Good news: conjuncts are in decidable fragments

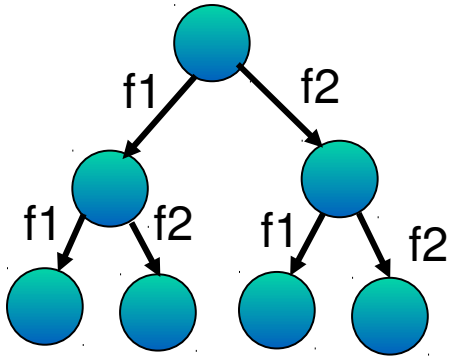
Bad news: conjuncts share more than just equality (they share set variables and set operations)

Next: explain these decidable fragments

WS2S: Monadic 2rd Order Logic

Weak Monadic 2rd-order Logic of 2 Successors

F ::= x=f1(y) | x=f2(y) |
x∈S | S⊆T | ∃S.F |
F_1 Æ F_2 | ¬F



- quantification is over finite sets of positions in a tree
- transitive closure encoded using set quantification

Decision procedure using tree automata (e.g. MONA)

C^2 : Two-Variable Logic w/ Counting

Two-Variable Logic with Counting

$F ::= P(v_1,...,v_n) \mid F_1 \text{ Æ } F_2 \mid \neg F \mid \exists^{\text{cnt}}\ v_i.F$

where

P : is a predicate symbol

v_i : is one of the ***two*** variable names x,y

cnt : is $=k$, $\leq k$, or $\geq k$ for nonnegative *constants* k

We can write $(\exists^{\leq k}\ v_i.F)$ as $|\{v_i.F\}| \leq k$

We can define $\exists, \forall$ and axiomatize total functions:
$\forall x \exists^{=1}\ y.R(x,y)$

Decidable sat. and fin-sat. (1997), NEXPTIME
even for binary-encoded k: Pratt-Hartman '05

BAPA (Kuncak et al. CADE'05): Boolean Algebra with Presburger Arithmetic

$$S ::= V \mid S_1 \cup S_2 \mid S_1 \cap S_2 \mid S_1 \setminus S_2$$

$$T ::= k \mid C \mid T_1 + T_2 \mid T_1 - T_2 \mid C \cdot T \mid |S|$$

$$A ::= S_1 = S_2 \mid S_1 \subseteq S_2 \mid T_1 = T_2 \mid T_1 < T_2$$

$$F ::= A \mid F_1 \wedge F_2 \mid F_1 \vee F_2 \mid \neg F \mid \exists S.F \mid \exists k.F$$

BAPA decidable in alternating time (V. Kuncak et al. JAR'06), QFBAPA decidable in NP (V. Kuncak et al. CADE'07)

Also decidable: qf fragment of multisets w/ cardinalities

(R. Piskac and V. Kuncak VMCAI'08,CAV'08,CSL'08)

New: role of BAPA in combination of theories sharing sets

Combining Theories by Reduction

Satisfiability problem expressed in HOL:

(all free symbols existentially quantified)

$\exists$ next,data,k,root. $\exists$A,B.

A = {x. next*(root,x)} $\wedge$

B = {y. $\exists$ x. data(x,y) $\wedge$ x $\in$ A} $\wedge$

|B|=k+1

1) WS2S

2) C^2

3) BAPA

We assume formulas share only:

- **set variables** (sets of uninterpreted elems)
- individual variables, as a special case - {x}

Combining Theories by Reduction

Satisfiability problem expressed in HOL,
after moving fragment-specific quantifiers

$\exists$ A,B.

$\exists$ next,root. A = {x. next*(root,x)} $\wedge$ — F_{WS2S}

$\exists$ data. B = {y. $\exists$ x. data(x,y) $\wedge$ x $\in$ A} $\wedge$ — F_{C2}

$\exists$ k. |B|=k+1 — F_{BAPA}

Extend decision procedures for fragments into

projection procedures that reduce each conjunct to a decidable shared theory

applies $\exists$ to all non-set variables

Combining Theories by Reduction

Satisfiability problem expressed in HOL,
after moving fragment-specific quantifiers

$\exists$ A,B.

$\exists$ next,root. A = {x. next*(root,x)} $\wedge$ — F_{WS2S}

$\exists$ data. B = {y. $\exists$ x. data(x,y) $\wedge$ x $\in$ A} $\wedge$ — F_{C2}

$\exists$ k. |B|=k+1 — F_{BAPA}

Check satisfiability of conjunction of projections

$$\exists A,B.\ F_{WS2S} \wedge F_{C2} \wedge F_{BAPA}$$

Conjunction of projections satisfiable → so is original formula

Decision Procedure for Combination

- Separate formula into WS2S, C^2, BAPA parts
- For each part, compute projection onto set vars
- Check satisfiability of conjunction of projections

What is the right target theory for expressing the projections onto set variables?

Outline

I. **Idea of decision procedure:**
reduction to a shared theory of sets

II. **BAPA-reducible theories**

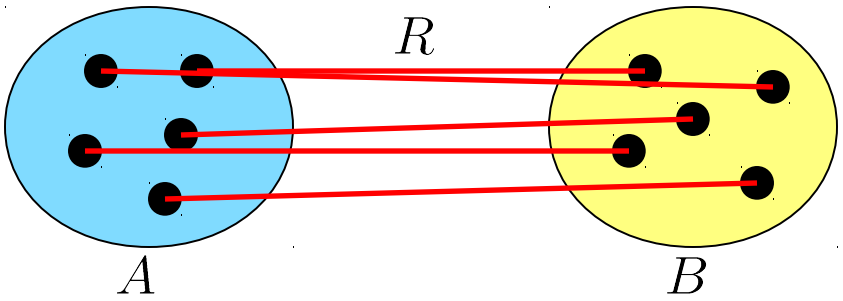
III. **BAPA-reduction of WS1S**

Reduction to BAPA

Consider the C² formula

$$F \equiv (\forall x.\exists^{=1} y.R(x,y)) \wedge (\forall x.\exists^{=1} y.R(y,x)) \wedge (\forall x.A(x) \leftrightarrow (\exists y.B(y) \wedge R(x,y)))$$

F expresses "R is bijection between A and B"



Projection of F onto A and B gives

$$(\exists R.F) \equiv (|A| = |B|)$$

Cardinalities are needed to express projections → BAPA

BAPA-Reducibility

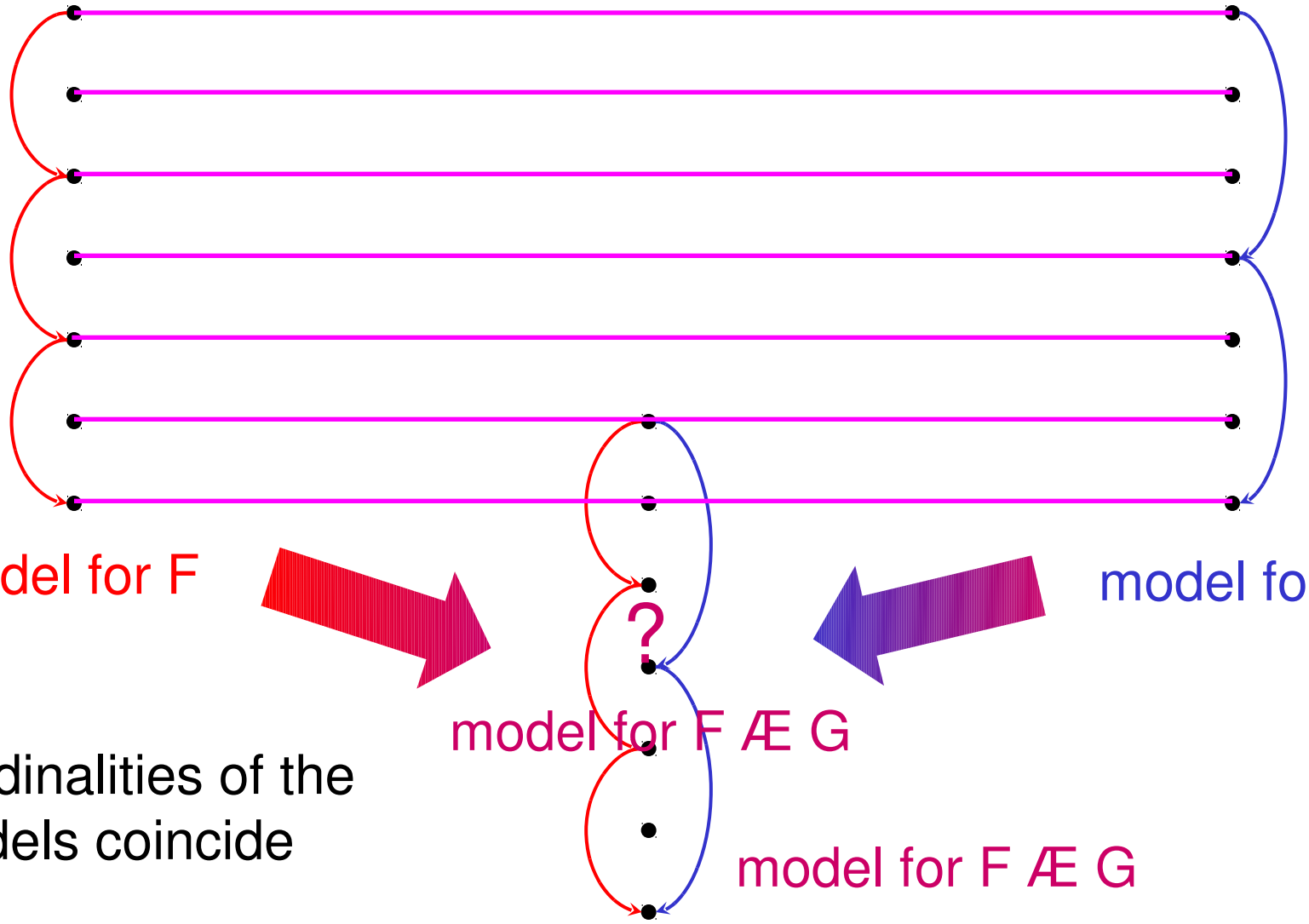
Definition: Logic is **BAPA-reducible** iff there is an algorithm that computes projections of formulas onto set variables, and these projections are BAPA formulas.

Theorem:

1) WS2S, **2)** C^2, **3)** BAPA, **4)** BSR, **5)** qf-multisets are all BAPA-reducible.

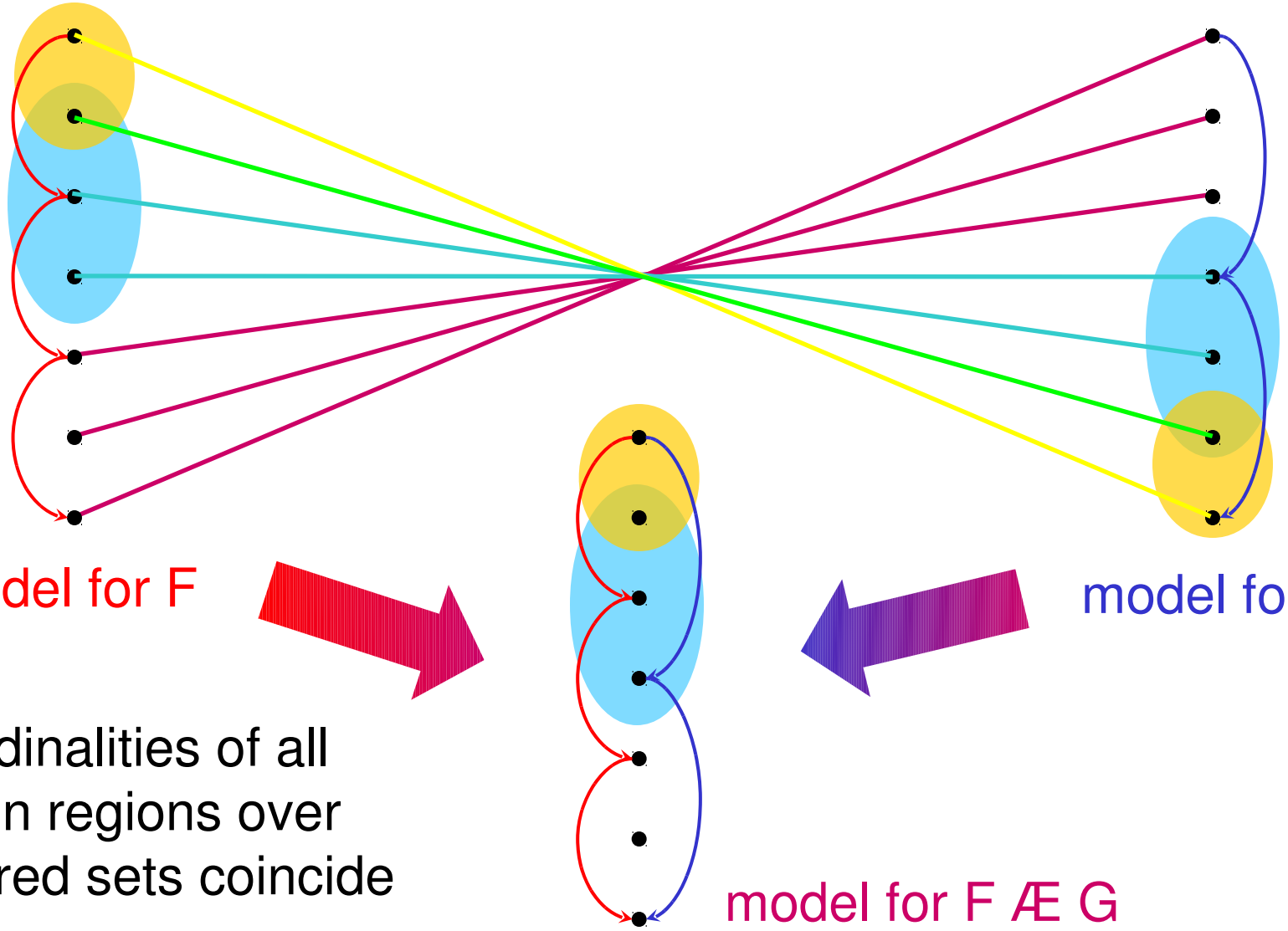
Thus, their set-sharing combination is decidable.

Amalgamation of Models:
The Disjoint Case



Cardinalities of the models coincide

Amalgamation of Models:
The Set-Sharing Case



Cardinalities of all
Venn regions over
shared sets coincide

BAPA-reducible Theories

For a set of formulas $\mathcal{F}$ the following are equivalent:

- theory $\mathcal{T} \subseteq \mathcal{F}$ is **BAPA-reducible**
- for every $F \in \mathcal{F}$ the set of vectors

$$\mathcal{V}(\mathcal{T}, F) = \{(|M(V_1)|, \ldots, |M(V_n)|) \mid M \models \mathcal{T} \cup \{F\}\}$$

  is **semilinear** and **effectively computable**,
  where the $M(V_i)$ are the Venn regions over
  the free set variables of F in its $\mathcal{T}$-models M

Projections onto the set variables characterize the sets $\mathcal{V}(\mathcal{T}, F)$.

Outline

I. **Idea of decision procedure:**
reduction to a shared theory of sets

II. **BAPA-reducible theories**

III. **BAPA-reduction of WS1S**

BAPA-reduction for WS1S

WS1S formula for a regular language

$$F = ((A \text{ Æ} \neg B)(B \text{ Æ} \neg A))^* (\neg B \text{ Æ} \neg A)^*$$

Formulas are interpreted over finite words

Symbols in alphabet correspond to

$$(\neg A \text{ Æ} \neg B), (A \text{ Æ} \neg B), (\neg A \text{ Æ} B), (A \text{ Æ} B)$$

00 10 01 11

Model of formula F

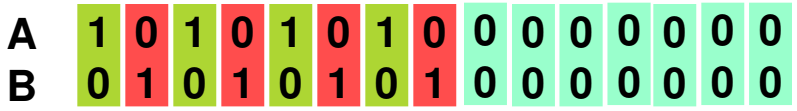
A	1	0	1	0	1	0	1	0	0	0	0	0	0	0
B	0	1	0	1	0	1	0	1	0	0	0	0	0	0

BAPA-reduction for WS1S

WS1S formula for a regular language

F = ((A Æ¬B)(B Æ¬A))* (¬B Æ¬A)*

Model of formula F



} w

A,B denote sets of positions in the word w.

00, 10, 01, 11 denote Venn regions over A,B

Parikh image gives card.s of Venn regions

Parikh(w) = { 00 7, 10 4, 01 4, 11 0}

BAPA-reduction for WS1S

Decision procedure for sat. of WS1S:

- construct finite word automaton A from F
- check emptiness of L(A)

Parikh 1966:

Parikh image of a regular language is semilinear and effectively computable from the finite automaton

Construct BAPA formula from Parikh image of the reg. lang.

BAPA-reduction for WS1S

WS1S formula for a regular language

$$F = ((A \text{ Æ} \neg B)(B \text{ Æ} \neg A))^* (\neg B \text{ Æ} \neg A)^*$$

Parikh image of the models of F:

$$\text{Parikh}(F) = \{(q,p,p,0) \mid q,p \geq 0\}$$

00 10 01 11

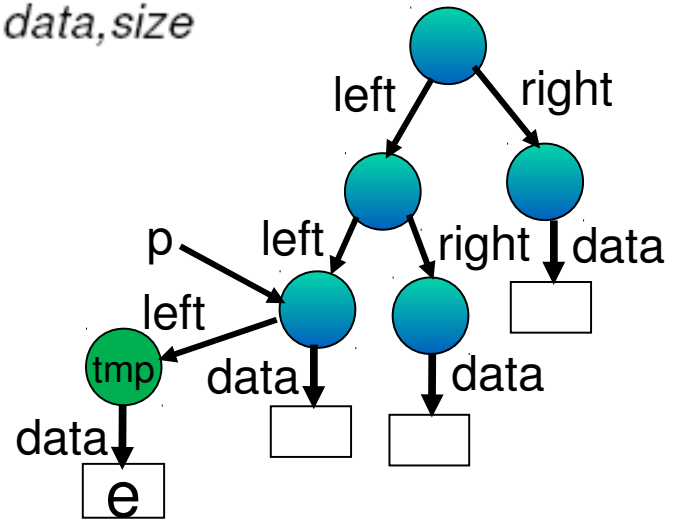
BAPA formula for projection of F onto A,B:

$$|A \text{ Å } B^c| = |A^c \text{ Å } B| \text{ Æ } |A \text{ Å } B| = 0$$

Fragment of Insertion into Tree

```
class Node {Node left,right; Object data;}
class Tree {
    private static Node root;
    private static int size; /*:
    private static specvar nodes :: objset;
    vardefs "nodes=={x. (root,x) ∈ {(x,y). left x = y ∨ right x = y}* }";
    private static specvar content :: objset;
    vardefs "content=={x. ∃ n. n ≠ null ∧ n ∈ nodes ∧ data n = x} " */

    private void insertAt(Node p, Object e) /*:
      requires "tree [ left , right ] ∧ nodes ⊆ Object.alloc ∧ size = card content ∧
                e ∉ content ∧ e ≠ null ∧ p ∈ nodes ∧ p ≠ null ∧ left p = null"
      modifies nodes,content,left, right ,data,size
      ensures "size = card content" */
    {
        Node tmp = new Node();
        tmp.data = e;
        p. left = tmp;
        size = size + 1;
    }
}
```



Reduction of VC for insertAt

SHARED SETS: nodes, nodes1, content, content1, {e}, {tmp}

WS2S FRAGMENT:

$tree[\ left\ ,\ right\ ]\ \wedge\ left\ p = null\ \wedge\ p \in nodes\ \wedge\ left\ tmp = null\ \wedge\ right\ tmp = null\ \wedge$

$nodes = \{x.\ (root,x) \in \{(x,y).\ left\ x = y \mid right\ x = y\}^{*}\}\ \wedge$

$nodes1 = \{x.\ (root,x) \in \{(x,y).\ (left\ (p:=tmp))\ x = y)\ \mid\ right\ x = y\}$

CONSEQUENCE: $nodes1 = nodes \cup \{tmp\}$

C2 FRAGMENT:

$data\ tmp = null\ \wedge\ (\forall\ y.\ data\ y \neq tmp)\ \wedge\ tmp \notin alloc\ \wedge\ nodes \subseteq alloc\ \wedge$

$content = \{x.\ \exists\ n.\ n \neq null\ \wedge\ n \in nodes\ \wedge\ data\ n = x\}\ \wedge$

$content1 = \{x.\ \exists\ n.\ n \neq null\ \wedge\ n \in nodes1\ \wedge\ (data(tmp:=e))\ n = x\}$

CONSEQUENCE: $nodes1 \neq nodes \cup \{tmp\} \vee content1 = content \cup \{e\}$

BAPA FRAGMENT: $e \notin content\ \wedge\ card\ content1 \neq card\ content + 1$

CONSEQUENCE: $e \notin content\ \wedge\ card\ content1 \neq card\ content + 1$

Conjunction of projections unsatisfiable → so is original formula

Related Work on Combination

Nelson-Oppen, 1980 – disjoint theories

- reduces to equality logic (finite # of formulas)

Tinelli, Ringeissen, 2003 – general non-disjoint

- we consider the particular case of sets

Ghilardi – sharing locally finite theories

- cardinality on sets needed, not locally finite

Fontaine – gentle theories (BSR, …)

- disjoint case only

Ruess, Klaedtke – WS2S + cardinality (no C^2)

Reduction procedures to SAT (UCLID)

- we reduce to (QF)BAPA (NP-complete)
- reduction QFBAPA → QFPA → SAT non-trivial

Summary

Presented new combination technique for theories sharing sets by reduction to a common shared theory (BAPA).

Identified an expressive decidable set-sharing combination of theories by extending their decision procedures to BAPA-reductions

1) WS2S, **2)** C^2 **3)** BSR, **4)** BAPA, **5)** qf-multisets

Resulting theory is useful for automated verification of complex properties of data structure implementations.