

Proving Properties of Constraint Logic Programs by Eliminating Existential Variables

Alberto Pettorossi (Univ. Tor Vergata, Rome, Italy),

Maurizio Proietti (IASI-CNR, Rome, Italy),

Valerio Senni (Univ. Tor Vergata, Rome, Italy)

Theorem Proving via Program Transformation

- Transformation techniques for ATP:
 - proving equivalences via unfold/fold [Kott-82, PP-99, Roychoudhury et al-99]
 - proving first order formulas via unfold/fold [PP-00]
 - verifying temporal properties of infinite state systems via specialization and u/f [Leuschel 99,00, Roychoudhury et al-00, Fioravanti et al-01]
- This work: techniques for the elimination of intermediate data structures (*deforestation*) can be used as methods for quantifier elimination.
- Intermediate data structures can be avoided by eliminating *existential* (or *local*, or *redundant*) variables, i.e., variables *occurring in the body* of a clause and *not in the head*. For instance, X is an existential variable in

$$p \leftarrow q(X) \wedge r(X)$$

$$[\text{Recall: } \forall X (p \leftarrow q(X) \wedge r(X)) \equiv p \leftarrow \exists X (q(X) \wedge r(X))]$$

The Quantifier Elimination Method

- **QE**: Given a theory T and a formula φ , find a quantifier-free formula ψ s.t.
 - $\text{freevars}(\psi) \subseteq \text{freevars}(\varphi)$
 - $T \models \varphi$ iff $T \models \psi$
- If **QE** is **algorithmic** and the problem of checking $T \models A$, for any ground atom A , is **decidable**, then the problem of checking $T \models \varphi$, for any closed formula φ , is **decidable**.
- The theory \mathbf{R} of real numbers (as an ordered field) admits **QE** and for any ground atom A (*i.e.*, ground real arithmetic expression), $\mathbf{R} \models A$ can be decided by direct evaluation.
Thus, $\mathbf{R} \models \varphi$ is decidable [Tarski 48].

Quantifier Elimination for Linear Constraints

- For the first order theory R_{lin} of linear equations/inequations over the real numbers, **QE** makes use of Fourier-Motzkin elimination procedure.

- Example.

$$\forall x (x \geq 1 \rightarrow \exists y (x - y > 0 \wedge y > 0))$$

$$\neg \exists x (x \geq 1 \wedge \neg \exists y (x - y > 0 \wedge y > 0))$$

$$\neg \exists x (x \geq 1 \wedge \neg x > 0)$$

$$\neg \exists x (x \geq 1 \wedge 0 \geq x)$$

$$\neg 0 \geq 1$$

True

$$[\forall x (\alpha \rightarrow \beta) \equiv \neg \exists x (\alpha \wedge \neg \beta)]$$

[F-M: elimination of y]

[negation of $>$]

[F-M: elimination of x]

[evaluation of ground formula]

QE for $CLP(R_{lin})$ via Program Transformation

- We are given a $CLP(R_{lin})$ program

P : member $(X, [Y|L]) \leftarrow X=Y$
member $(X, [Y|L]) \leftarrow$ member (X, L)

and a closed first order formula

φ : $\forall L \exists U \forall X (\text{member}(X, L) \rightarrow X \leq U)$

- Step1.** By a variant of Lloyd-Topor transformation we get a *clause form* CF for φ s.t. $M(P) \models \varphi$ iff $M(P \cup CF) \models f$, where $M()$ denotes the *perfect R_{lin} -model*.

CF : $f \leftarrow \neg p$
 $p \leftarrow \text{list}(L) \wedge \neg q(L)$
 $q(L) \leftarrow \text{list}(L) \wedge \neg r(L, U)$
 $r(L, U) \leftarrow X > U \wedge \text{list}(L) \wedge \text{member}(X, L)$

...QE for $CLP(R_{lin})$ via Program Transformation

- **Step 2.** By *unfold/fold transformations and quantifier elimination from constraints in R_{lin}* , from $P \cup CF$ we derive a *propositional program Prop* s.t. $M(P \cup CF) \models f$ iff $M(Prop) \models f$.

Prop: $f \leftarrow \neg p$
 $p \leftarrow p1$
 $p1 \leftarrow p1$

- Thus, $M(P) \models \varphi$ iff $M(Prop) \models f$.
- The transformation from $P \cup CF$ to *Prop* consists in eliminating all *existential variables* from *CF*.

Overview

- *LR-programs*: A class of CLP programs on lists of real numbers with linear constraints;
- *Clause form transformation* for first order formulas;
- *Unfold/fold* transformations of clause forms;
- An *automatic strategy* for deriving propositional programs by eliminating existential variables.

Programs on Lists of Real Numbers

- $a \in \mathbb{R}, X \in \text{Var}_{\mathbb{R}}, L \in \text{Var}_{\text{List}}$
- Polynomials: $p ::= a \mid X \mid p_1 + p_2 \mid a X$
- Constraints: $c ::= p_1 = p_2 \mid p_1 < p_2 \mid p_1 \leq p_2 \mid c_1 \wedge c_2$
- LR-programs:
 - head terms $h ::= X \mid [] \mid [X|L]$
 - body terms $b ::= p \mid L$
 - clauses $cl ::= r_1(h_1, \dots, h_n) \leftarrow c \mid$
 $r_1(h_1, \dots, h_n) \leftarrow c \wedge r_2(b_1, \dots, b_n) \mid$
 $r_1(h_1, \dots, h_n) \leftarrow c \wedge \neg r_2(b_1, \dots, b_n)$

where: (i) cl has *no existential variables*, (ii) $r_1(h_1, \dots, h_n)$ is a *linear atom*, and (iii) $\text{vars}(p) \neq \emptyset$.

LR-Programs: Examples

- LR-programs:

$\text{member}(X, [Y|L]) \leftarrow X=Y$

$\text{member}(X, [Y|L]) \leftarrow \text{member}(X, L)$

$\text{pos_sumlist}([], Y) \leftarrow Y=0$

$\text{pos_sumlist}([X|L], Y) \leftarrow X>0 \wedge \text{pos_sumlist}(L, Y-X)$

$\text{pos_sumlist}([X|L], Y) \leftarrow X\leq 0 \wedge \text{pos_sumlist}(L, Y)$

- Not an LR-program:

$\text{permutation}([], []) \leftarrow$

$\text{permutation}([X|L1], L2) \leftarrow \text{permutation}(L1, L3) \wedge \text{insert}(X, L3, L2)$

- $L2$ is neither $[]$ nor $[X|L]$
- $L3$ is existential
- the body has **two** literals

Properties of LR-Programs

- The problem of checking $M(P) \models \varphi$, for any LR-program P and closed formula φ , is **undecidable**. Peano arithmetic can be encoded via an LR-program.
- The transformation from $P \cup CF$ to $Prop$ cannot be algorithmic.
- If $P \cup CF$ is transformed into an LR-program T , then the 0-ary predicate f is defined by a set $Prop \subseteq T$ of propositional clauses. Thus, **quantifiers can be eliminated by deriving LR-programs**.

Clause-Form Transformation

$$\varphi: \forall L \exists U \forall X (\text{member}(X,L) \rightarrow X \leq U)$$

$$f \leftarrow \neg \exists L \neg \exists U \neg \exists X (\text{member}(X,L) \wedge \neg X \leq U)$$

- Lloyd-Topor transformation + addition of a `list(L)` atom for each list variable `L` (needed for unfolding).

CF: D4: $f \leftarrow \neg p$

D3: $p \leftarrow \text{list}(L) \wedge \neg q(L)$

D2: $q(L) \leftarrow \text{list}(L) \wedge \neg r(L, U)$

D1: $r(L, U) \leftarrow X > U \wedge \text{list}(L) \wedge \text{member}(X, L)$

- stratified program
- *not* LR-clauses (with **existential variables**)

- $D1, D2, D3, D4$ will be transformed to LR-clauses by unfold/fold transformations.

Eliminating Existential Variables via U/F

D1: $r(L,U) \leftarrow \exists X > U \wedge \text{list}(L) \wedge \text{member}(X,L)$

Eliminating Existential Variables via U/F

D1: $r(L,U) \leftarrow X > U \wedge \underline{\text{list}(L)} \wedge \underline{\text{member}(X,L)}$

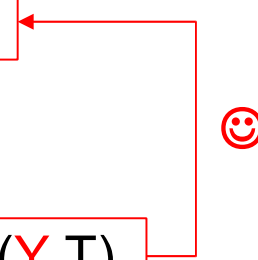
Unfold: $r([X|T],U) \leftarrow X > U \wedge \text{list}(T)$

$r([X|T],U) \leftarrow Y > U \wedge \text{list}(T) \wedge \text{member}(Y,T)$

Eliminating Existential Variables via U/F

D1: $r(L,U) \leftarrow \boxed{X > U \wedge \text{list}(L) \wedge \text{member}(X,L)}$

Unfold: $r([X|T],U) \leftarrow X > U \wedge \text{list}(T)$
 $r([X|T],U) \leftarrow \boxed{Y > U \wedge \text{list}(T) \wedge \text{member}(Y,T)}$



Eliminating Existential Variables via U/F

D1: $r(L,U) \leftarrow X > U \wedge \text{list}(L) \wedge \text{member}(X,L)$

Unfold: $r([X|T],U) \leftarrow X > U \wedge \text{list}(T)$

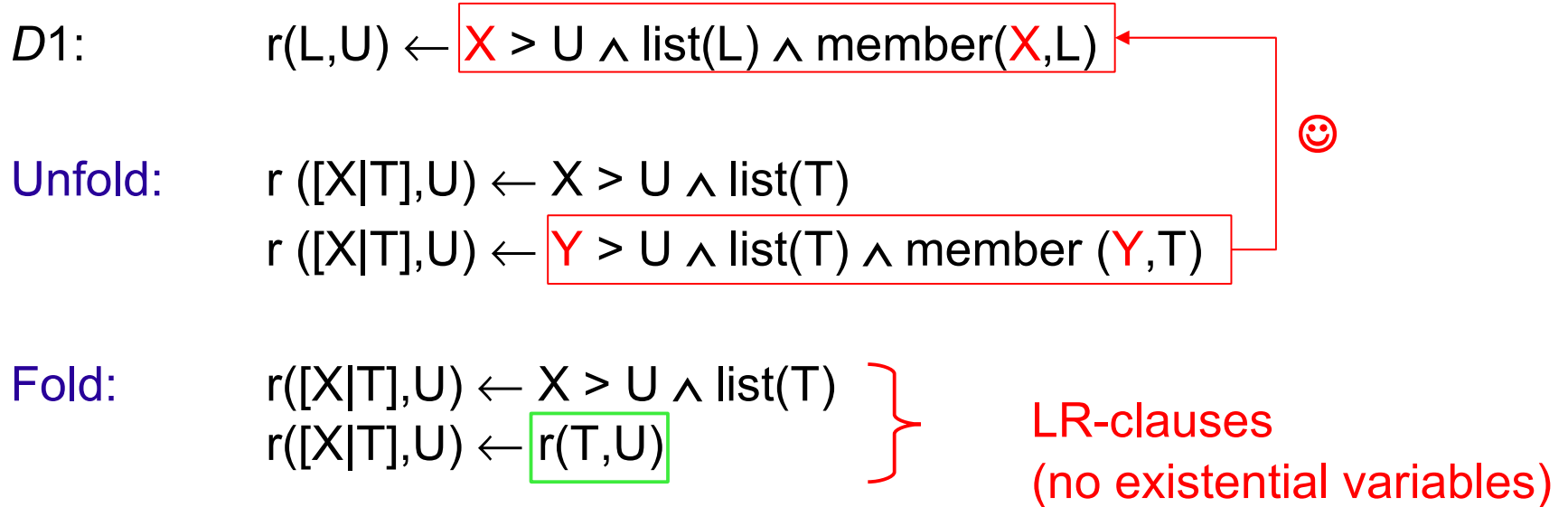
$r([X|T],U) \leftarrow Y > U \wedge \text{list}(T) \wedge \text{member}(Y,T)$

Fold: $r([X|T],U) \leftarrow X > U \wedge \text{list}(T)$

$r([X|T],U) \leftarrow r(T,U)$



Eliminating Existential Variables via U/F



Transformed program (1)

$D4: f \leftarrow \neg p$

$D3: p \leftarrow \text{list}(L) \wedge \neg q(L)$

$D2: q(L) \leftarrow \text{list}(L) \wedge \neg r(L, U)$

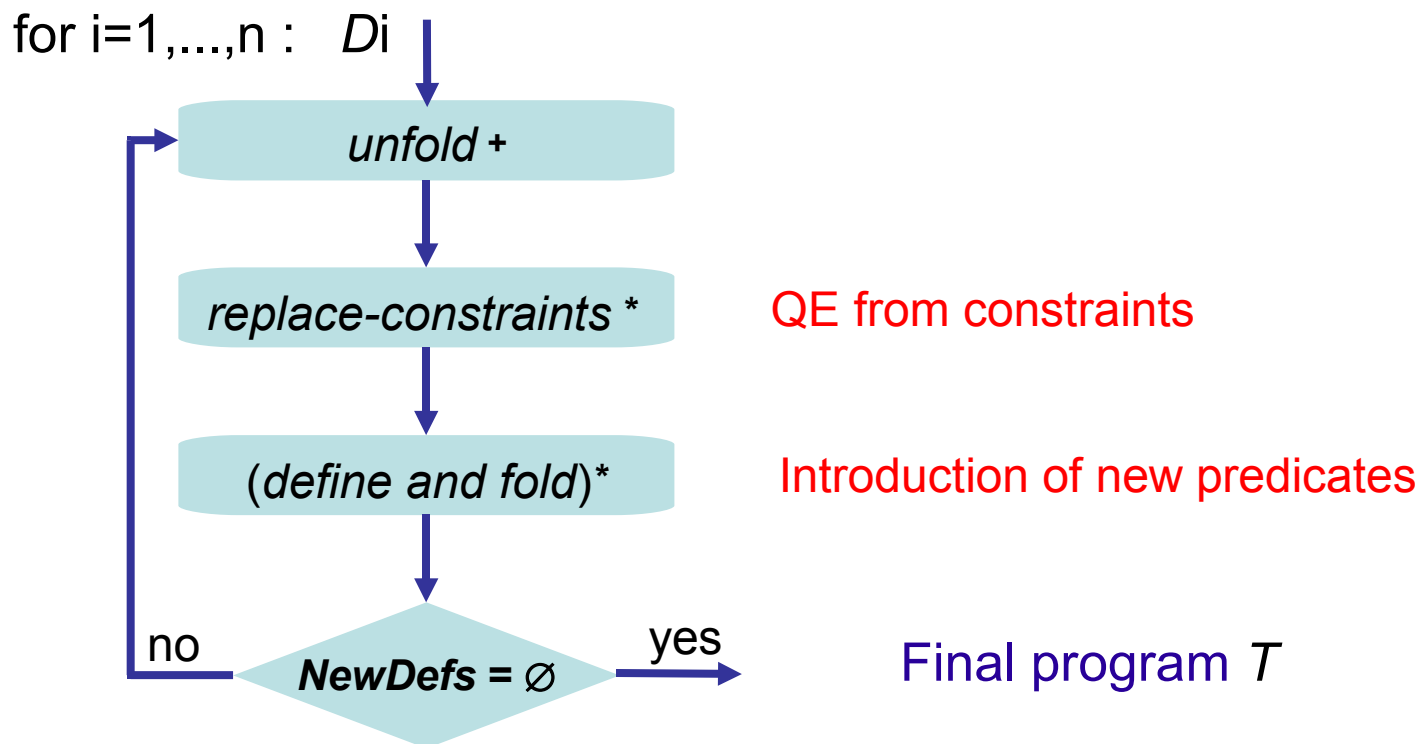
$D1: r([X|T], U) \leftarrow X > U \wedge \text{list}(T)$
 $r([X|T], U) \leftarrow r(T, U)$

No existential variables

An Unfold-Fold Strategy for Deriving LR-programs

Input: an LR-program P and the clause form $CF: D1, \dots, Dn$ of a closed first order formula φ

Output: an LR-program T s.t. f is defined by a propositional $Prop \subseteq T$ and $M(P) \models \varphi$ iff $M(Prop) \models f$.



Introducing New Definitions

$D2:$ $q(L) \leftarrow \text{list}(L) \wedge \neg r(L, U)$

Introducing New Definitions

D2: $q(L) \leftarrow \underline{\text{list}(L)} \wedge \neg \underline{r(L, U)}$

Unfold : $q([]) \leftarrow$
 $q([X|T]) \leftarrow X \leq U \wedge \text{list}(T) \wedge \neg r(T, U)$

Introducing New Definitions

D2:

$$q(L) \leftarrow \boxed{\text{list}(L) \wedge \neg r(L, U)}$$

Unfold :

$$q([\]) \leftarrow$$

$$q([X|T]) \leftarrow X \leq U \wedge \boxed{\text{list}(T) \wedge \neg r(T, U)}$$



Bad folding!
Existential variable
not eliminated.

Introducing New Definitions

D2: $q(L) \leftarrow \text{list}(L) \wedge \neg r(L, U)$

Unfold : $q([]) \leftarrow$
 $q([X|T]) \leftarrow X \leq U \wedge \text{list}(T) \wedge \neg r(T, U)$

Define : $q1(X, T) \leftarrow X \leq U \wedge \text{list}(T) \wedge \neg r(T, U)$

Introducing New Definitions

D2: $q(L) \leftarrow \text{list}(L) \wedge \neg r(L, U)$

Unfold : $q([]) \leftarrow$
 $q([X|T]) \leftarrow X \leq U \wedge \text{list}(T) \wedge \neg r(T, U)$

Define : $q1(X, T) \leftarrow X \leq U \wedge \text{list}(T) \wedge \neg r(T, U)$

Fold : $q([]) \leftarrow$
 $q([X|T]) \leftarrow q1(X, T)$ } LR-clauses ☺
(no existential variables)

Existential variables to be eliminated from the new definition

Transforming New Definitions

We transform the new definition into a set of LR-clauses

New Def.: $q1(X,T) \leftarrow X \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$

Transforming New Definitions

We transform the new definition into a set of LR-clauses

New Def.: $q1(X,T) \leftarrow X \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$

Unfold: $q1(X,[]) \leftarrow$

$q1(X,[Y|T]) \leftarrow X \leq U \wedge Y \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$



Bad folding!
Existential variable
not eliminated

Transforming New Definitions

We transform the new definition into a set of LR-clauses

New Def.: $q1(X,T) \leftarrow X \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$

Unfold: $q1(X,[]) \leftarrow$

$q1(X,[Y|T]) \leftarrow X \leq U \wedge Y \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$

linear order

$\equiv (X > Y \wedge X \leq U) \vee (X \leq Y \wedge Y \leq U)$

Replace-constraints:

$q1(X,[Y|T]) \leftarrow X > Y \wedge X \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$

$q1(X,[Y|T]) \leftarrow X \leq Y \wedge Y \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$

Transforming New Definitions

We transform the new definition into a set of LR-clauses

New Def.: $q1(X,T) \leftarrow X \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$

Unfold: $q1(X,[]) \leftarrow$

$q1(X,[Y|T]) \leftarrow X \leq U \wedge Y \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$

Replace-constraints:

$q1(X,[Y|T]) \leftarrow X > Y \wedge X \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$

$q1(X,[Y|T]) \leftarrow X \leq Y \wedge Y \leq U \wedge \text{list}(T) \wedge \neg r(T,U)$

Fold: $q1(X,[]) \leftarrow$

$q1(X,[Y|T]) \leftarrow X > Y \wedge q1(X,T)$

$q1(X,[Y|T]) \leftarrow X \leq Y \wedge q1(Y,T)$

} LR-clauses
(no existential variables)

Transformed program (2)

D4: $f \leftarrow \neg p$

D3: $p \leftarrow \text{list}(L) \wedge \neg q(L)$

D2: $q([\])$ \leftarrow
 $q([X|T]) \leftarrow q1(X,T)$
 $q1(X,[Y|T]) \leftarrow X > Y \wedge q1(X, L)$
 $q1(X,[Y|T]) \leftarrow X \leq Y \wedge q1(Y, L)$

No existential variables

D1: $r([X|T],U) \leftarrow X > U \wedge \text{list}(T)$
 $r([X|T],U) \leftarrow r(T,U)$

Deriving a Propositional Program

D3: $p \leftarrow \text{list}(L) \wedge \neg q(L)$

Deriving a Propositional Program

D3: $p \leftarrow \underline{\text{list(L)}} \wedge \underline{\neg q(L)}$

Unfold: $p \leftarrow \text{list(T)} \wedge \neg q1(X, T)$ Folding not possible

Deriving a Propositional Program

D3: $p \leftarrow \text{list}(L) \wedge \neg q(L)$

Unfold: $p \leftarrow \text{list}(T) \wedge \neg q1(X, T)$

Define: $p1 \leftarrow \text{list}(T) \wedge \neg q1(X, T)$



Deriving a Propositional Program

D3: $p \leftarrow \text{list}(L) \wedge \neg q(L)$

Unfold: $p \leftarrow \text{list}(T) \wedge \neg q1(X, T)$

Define: $p1 \leftarrow \text{list}(T) \wedge \neg q1(X, T)$

Fold: $p \leftarrow p1$

LR-clause
(no existential variables,
propositional)

Deriving a Propositional Program

New Def.: $p1 \leftarrow \text{list}(T) \wedge \neg q1(X, T)$

Deriving a Propositional Program

New Def.: $p1 \leftarrow \underline{\text{list}(T)} \wedge \underline{\neg q1(X,T)}$

Unfold: $p1 \leftarrow X > Y \wedge \text{list}(T) \wedge \neg q1(X,T)$

$p1 \leftarrow X \leq Y \wedge \text{list}(T) \wedge \neg q1(Y,T)$

Deriving a Propositional Program

New Def.: $p1 \leftarrow \text{list}(T) \wedge \neg q1(X, T)$

Unfold: $p1 \leftarrow X > Y \wedge \text{list}(T) \wedge \neg q1(X, T)$

$p1 \leftarrow X \leq Y \wedge \text{list}(T) \wedge \neg q1(Y, T)$

$\exists Y X > Y \equiv \text{true}$

$\exists X X \leq Y \equiv \text{true}$

Replace-Constraints (variable elimination):

$p1 \leftarrow \text{list}(T) \wedge \neg q1(Y, T)$

Deriving a Propositional Program

New Def.:

$$p1 \leftarrow \boxed{\text{list}(T) \wedge \neg q1(X, T)}$$

Unfold:

$$p1 \leftarrow \boxed{X > Y} \wedge \text{list}(T) \wedge \neg q1(X, T)$$

$$p1 \leftarrow \boxed{X \leq Y} \wedge \text{list}(T) \wedge \neg q1(Y, T)$$

Replace-Constraints (variable elimination):

$$p1 \leftarrow \boxed{\text{list}(T) \wedge \neg q1(Y, T)}$$

Fold:

$$p1 \leftarrow \boxed{p1}$$

LR-clause
(no existential variables,
propositional)



The Final LR-Program

T {

Prop

D4: $f \leftarrow \neg p$

D3: $p \leftarrow p1$
 $p1 \leftarrow p1$

D2: $q([\])$ \leftarrow
 $q([X|T]) \leftarrow q1(X,T)$
 $q1(X,[Y|T]) \leftarrow X > Y \wedge q1(X, L)$
 $q1(X,[Y|T]) \leftarrow X \leq Y \wedge q1(Y, L)$

D1: $r([X|T],U) \leftarrow X > U \wedge \text{list}(T)$
 $r([X|T],U) \leftarrow r(T,U)$

No existential variables

Proving Propositional Formulas

T	$D4:$	$f \leftarrow \neg p$	
	$D3:$	$p \leftarrow p1$	
		$p1 \leftarrow p1$	tautology
	$D2:$	$q([\])$	
		$q([X T]) \leftarrow q1(X,T)$	
		$q1(X,[Y T]) \leftarrow X > Y \wedge q1(X, L)$	
		$q1(X,[Y T]) \leftarrow X \leq Y \wedge q1(Y, L)$	
	$D1:$	$r([X T],U) \leftarrow X > U \wedge \text{list}(T)$	
		$r([X T],U) \leftarrow r(T,U)$	

Proving Propositional Formulas

T	$D4:$	$f \leftarrow \neg p$
	$D3:$	$p \leftarrow p1$
	$D2:$	$q([\])$
		$q([X T]) \leftarrow q1(X,T)$
		$q1(X,[Y T]) \leftarrow X > Y \wedge q1(X, L)$
		$q1(X,[Y T]) \leftarrow X \leq Y \wedge q1(Y, L)$
	$D1:$	$r([X T],U) \leftarrow X > U \wedge \text{list}(T)$
		$r([X T],U) \leftarrow r(T,U)$

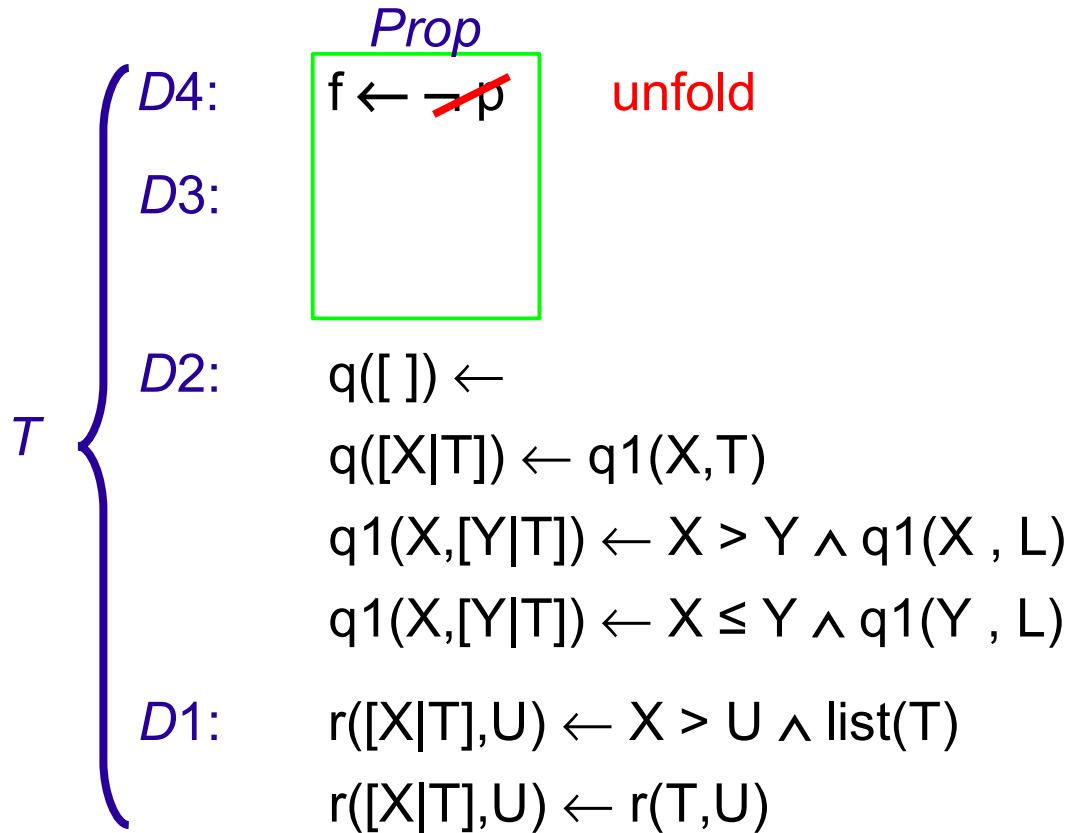
Proving Propositional Formulas

{	T	D4:	$f \leftarrow \neg p$		
		D3:	$p \leftarrow p1$	unfold	
		D2:	$q([\])$	\leftarrow	
			$q([X T])$	\leftarrow	$q1(X,T)$
			$q1(X,[Y T])$	\leftarrow	$X > Y \wedge q1(X, L)$
		$q1(X,[Y T])$	\leftarrow	$X \leq Y \wedge q1(Y, L)$	
		D1:	$r([X T],U)$	\leftarrow	$X > U \wedge \text{list}(T)$
			$r([X T],U)$	\leftarrow	$r(T,U)$

Proving Propositional Formulas

T	$D4:$	$f \leftarrow \neg p$
	$D3:$	
	$D2:$	$q([\])$
		$q([X T]) \leftarrow q1(X,T)$
		$q1(X,[Y T]) \leftarrow X > Y \wedge q1(X, L)$
		$q1(X,[Y T]) \leftarrow X \leq Y \wedge q1(Y, L)$
	$D1:$	$r([X T],U) \leftarrow X > U \wedge list(T)$
		$r([X T],U) \leftarrow r(T,U)$

Proving Propositional Formulas



Proving Propositional Formulas

$$\begin{array}{l} T \left\{ \begin{array}{l} D4: \quad \boxed{\begin{array}{l} \text{Prop} \\ f \leftarrow \end{array}} \Rightarrow M(\text{Prop}) \models f \Rightarrow M(P) \models \varphi \\ D3: \\ D2: \quad \begin{array}{l} q([\]) \leftarrow \\ q([X|T]) \leftarrow q1(X,T) \\ q1(X,[Y|T]) \leftarrow X > Y \wedge q1(X, L) \\ q1(X,[Y|T]) \leftarrow X \leq Y \wedge q1(Y, L) \end{array} \\ D1: \quad \begin{array}{l} r([X|T],U) \leftarrow X > U \wedge \text{list}(T) \\ r([X|T],U) \leftarrow r(T,U) \end{array} \end{array} \end{array}$$

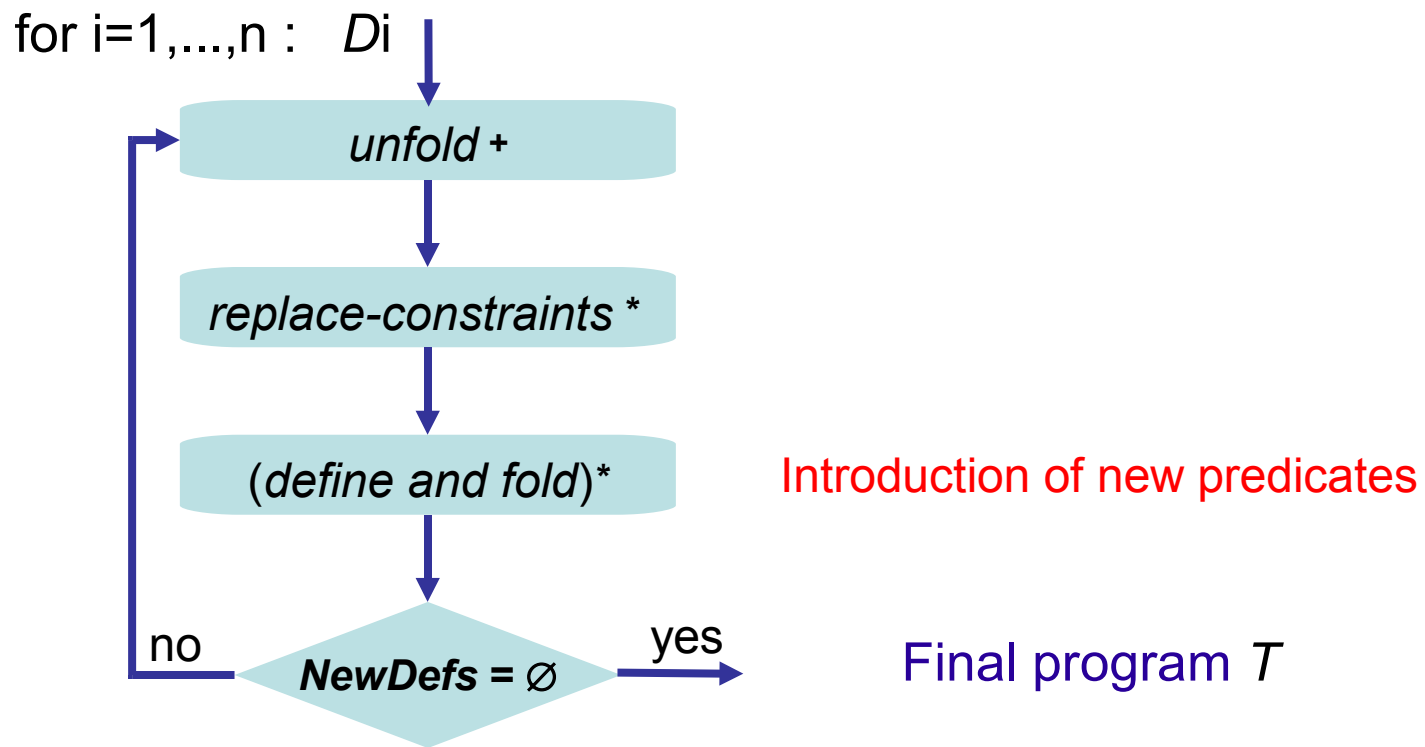
Experimental Results

- The Unfold/Fold transformation strategy for elimination of existential variables is implemented on the MAP system (www.iasi.cnr.it/~proietti/system.html).
- Constraints are handled using the clp(r) module of SICStus Prolog (implementing a variant of Fourier-Motzkin variable elimination)
- Proven formulas in the theory of linear orders, lists, and addition

Property	Time (PM 1.73)
$\forall L \exists U \forall Y (\text{member}(Y,L) \rightarrow Y \leq U)$	31 ms
$\forall L \forall Y ((\text{sumlist}(L,Y) \wedge Y > 0) \rightarrow \exists X (\text{member}(X,L) \rightarrow X > 0))$	15 ms
$\forall L \forall M \forall N ((\text{ord}(L) \wedge \text{ord}(M) \wedge \text{sumzip}(L,M,N)) \rightarrow \text{ord}(N))$	16 ms
$\forall L \forall M \forall X \forall Y ((\text{leqlist}(L,M) \wedge \text{sumlist}(L,X) \wedge \text{sumlist}(M,Y)) \rightarrow X \leq Y)$	16 ms

Termination of the Unfold-Fold Strategy

The only source for nontermination is the possible introduction of infinitely many new predicates.



Conclusions

- Program transformations to eliminate existential variables (deforestation) can be used for automated theorem proving;
- **Future work**
 - Identify theories of interest for which the unfold/fold strategy succeeds and, thus, works as a decision procedure;
 - Extend to other data structures (e.g. trees) and other domains closed under projection;
 - Investigate generalization techniques to avoid non-termination, with partial elimination of existential variables.