Blocking and Other Enhancements for Bottom-Up Model Generation Methods

Peter Baumgartner
National ICT Australia

Renate A. Schmidt
The University of Manchester
Motivation: Disproving

• **Disproving**
  – Show that a given (first-order) formula (with equality) is not valid
  – This can be done by **computing a model**, i.e. a counterexample

• **Applications of Disproving**
  – Mathematics
    • Refute conjectures
    • Finite group existence
  – Verification: disproving verification conditions
  – Knowledge representation
    • Knowledge base is consistent
    • Speculated subsumption relation does not hold

Existing methods? Limits? What’s new here?
Disproving Methods (1)

• **Finite model building**
  – Assume a fixed, finite domain \{d_1, ..., d_n\}
  – Decide if there is a model of the given formula over that domain
  – If not, add a new domain element and repeat

• **Methods**
  – MACE-style: by reduction to
    • propositional SAT (Paradox, Mace2) or
    • function-free clause logic (FM-Darwin)
  – SEM-style: guess function tables and check for model
  – (Tableaux) algorithms by Bry&Torge, Bezem, Nivelle&Meng

✓ No syntactic restrictions on input formula
✗ Finite models sometimes not sufficient
✗ Poor scaling
Consider the clause set consisting of the $O(n^2)$ unit clauses

$$p(c_1, \ldots, c_n)$$

$$\neg p(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n) \quad \text{for all } 1 \leq i < j \leq n$$

Second clause says no $c_i$ and $c_j$ can be mapped to the same element

- Therefore, smallest model has $n$ domain elements

10$^9$ instances for $n=10$

- For which $n$ do current finite model finders give up?

Any resolution method will terminate here

Finite model builders / (our) resolution methods are rather different

Our approach doesn’t iterate on domain size
Our approach doesn’t identify different constants
Disproving Methods (2)

• **Identify decidable fragment of FOL**
  – Guarded Fragment
  – Description and modal logics
  – Positive-variable dominated clauses
  – Prefix-classes: $\exists^*\forall^*$, $\exists^*\forall\exists^*$, ...

• **Design decision procedure for it**
  – From scratch. E.g. tableaux algorithms for description logics
  – By showing that a certain (resolution) refinement decides it. E.g. with axiomatic translation [Schmidt&Hustadt 2005], ordered resolution + splitting decides many modal logics

✔ Powerful
✗ Resolution not practical for $\exists^*\forall^*$
Really?
Problems of Using Resolution for \( \exists^* \forall^* \)

- \( \exists^* \forall^* \) fragment corresponds to function-free clause logic
  - Important for many database-like applications (Datalog)
- Pathological example for resolution:

\[
\text{Res} \quad \frac{p(x, y) \lor p(y, z) \leftarrow p(x, z) \quad p(a, z) \lor p(z, a)}{p(a, y) \lor p(y, z) \lor p(z, a)}
\]

Derived clauses pattern:

\[
\begin{align*}
p(a, z) \lor p(z, a) \\
p(a, y) \lor p(y, z) \lor p(z, a) \\
p(a, x) \lor p(x, y) \lor p(y, z) \lor p(z, a) \\
\vdots
\end{align*}
\]

- \( \times \) Refinements like subsumption, condensing, splitting don’t help
- \( \times \) Hyperresolution + range restriction works, but inefficiently
- \( \checkmark \) (One) contribution here: improved range restriction
Classical Approach

Input Clause Set

Standard Range Restriction

Bottom-Up Model Generation

Transformation

“Standard” prover
Classical Approach

Input Clause Set

Standard Range Restriction

Bottom-Up Model Generation

Transformation

“Standard” prover
Standard Range Restriction

[Manthey & Bry 88]

- A clause is **range restricted** iff each variable in its head also occurs in its body, as in \( p(x, z) \lor p(z, a) \leftarrow q(x, z) \)
- Every clause (set) can be made range restricted:
  - Restrict all extra variables in head in all input clauses to \( \text{dom} \), e.g.
    \[
    p(x, z) \lor p(z, a) \quad \text{becomes} \quad p(x, z) \lor p(z, a) \leftarrow \text{dom}(x) \land \text{dom}(z)
    \]
  - Add "dom"-clauses to enumerate Herbrand universe:
    \[
    \begin{align*}
    \text{dom}(a) \\
    \text{dom}(b) \\
    \text{dom}(f(x_1, \ldots, x_n)) & \leftarrow \text{dom}(x_1) \land \cdots \land \text{dom}(x_n)
    \end{align*}
    \]

**All positive clauses derived by hyperresolution are ground**
Classical Approach

- Input Clause Set
  - Standard Range Restriction
    - Bottom-Up Model Generation
      - "Standard" prover
        - E.g. based on **hyperresolution**
Hyperresolution + Range Restriction + Splitting

Given clause set:
\[ p(a, z) \lor p(z, a) \]
\[ p(x, y) \lor p(y, z) \quad \leftarrow \quad p(x, z) \]

Derivation by hyperresolution + splitting after range restriction

\[
\text{H-Res} \quad \frac{p(a, z) \lor p(z, a) \leftarrow \text{dom}(z) \quad \text{dom}(b)}{p(a, b) \lor p(b, a)}
\]

\[
\text{Split} \quad \frac{p(a, b) \lor p(b, a)}{p(a, b) \mid p(b, a)}
\]

\[
\text{H-Res} \quad \frac{p(x, y) \lor p(y, z) \leftarrow \text{dom}(y) \land p(x, z) \quad \text{dom}(a) \quad p(a, b)}{p(a) \lor p(a, b)}
\]

\[
\text{Split} \quad \frac{p(a, a) \lor p(a, b)}{p(a, a) \mid p(a, b)}
\]

\[ \checkmark \quad \text{Decides function-free clause logic (BS class)} \]

\[ \times \quad \text{Search space too big} \]

Improvement?
Our Approach

Input Clause Set

Improved

Standard
Range Restriction

Bottom-Up
Model Generation

Transformation

“Standard” prover
**Improved Range Restriction**

**Here:** idea by means of an example, see paper for details

(1) **Domain elements from clause heads**

| Clause          | Transformation                        |
|-----------------|----------------------------------------|
| \( P(x) \lor Q(b) \) | \( P(x) \lor Q(b) \leftarrow \text{dom}(x) \) |
|                 | \( \text{dom}(x) \leftarrow P(x) \)     for each head |
|                 | \( \text{dom}(x) \leftarrow Q(x) \)     predicate symbol |

✓ May yield smaller domain, depending on splits chosen
(2) Domain elements from clause bodies

| Clause          | Transformation                                      |
|-----------------|-----------------------------------------------------|
| $P(x)$          | $P(x) \leftarrow \text{dom}(x)$                    |
|                 | $\text{dom}(x) \leftarrow P(x)$                    |
| $\bot \leftarrow P(a) \land P(b)$ | $\bot \leftarrow P(a) \land P(b)$ |
|                 | $\text{dom}(a) \leftarrow P(x)$ for each body      |
|                 | $\text{dom}(b) \leftarrow P(x)$ predicate symbol   |

- May yield smaller domain, depending on satisfied literals
## Improved Range Restriction

### (2) Domain elements from clause bodies

| Clause | Transformation |
|--------|----------------|
| $P(x)$ | $P(x) \leftarrow \text{dom}(x)$  
$\text{dom}(x) \leftarrow P(x)$ |
| $\bot \leftarrow Q(a) \land Q(b)$ | $\bot \leftarrow Q(a) \land Q(b)$  
$\text{dom}(a) \leftarrow Q(x)$  
$\text{dom}(b) \leftarrow Q(x)$ for each body | predicate symbol |

- **May yield smaller domain, depending on satisfied literals**
Soundness and Completeness

- \( \text{rr}(M) := \) transformation of clause set \( M \) into range-restricted form

- **Proposition**
  A clause set \( M \) is satisfiable iff \( \text{rr}(M) \) is satisfiable
  Proof (completeness):
  - Given a Herbrand model \( I_{\text{rr}} \) of \( \text{rr}(M) \).
  - Define Interpretation \( I \) for \( M \):
    - Domain \( |I| = \{ t \mid I_{\text{rr}} \models \text{dom}(t) \} \)
    - Terms in \( |I| \) evaluate to themselves ("Quasi-Herbrand")
  - Show that \( I \) is a model of \( M \):

- **Corollary**
  A clause set \( M \) is E-satisfiable iff
  \( \text{rr}(M) \cup \{ x \approx x \leftarrow \text{dom}(x) \} \) is E-satisfiable
  Proof: Use equality axioms. Only equality axiom affected is reflexivity
Problem with Improved Range Restriction

- **Problem**: function symbols in clause bodies may lead to non-termination of BUMG

- **Example**:
  
  From \( r(x) \leftarrow q(x) \land p(f(x)) \)
  
  obtain \( \text{dom}(f(x)) \leftarrow p(y) \)
  
  and finally \( \text{dom}(f(x)) \leftarrow \text{dom}(x) \land p(y) \)

- Together with \( p(b), q(a), \text{dom}(a) \)
  
  derive \( \text{dom}(f(a)) \)
  
  \( \text{dom}(f(f(a))) \)
  
  \( \text{dom}(f(f(f(a)))) \)
  
  ...
Our Approach

Input Clause Set

Shifting

Improved Range Restriction

Bottom-Up Model Generation

Transformations

“Standard” prover
Shifting

- Moves body literals with function terms into the head:
  \[ r(x) \leftarrow q(x) \land p(f(x)) \]

  Shifting:
  \[ r(x) \lor \neg p(f(x)) \leftarrow q(x) \]
  \[ \bot \leftarrow p(x) \land \neg p(x) \]

- Advantage: critical head atom \( \neg p(f(x)) \) possibly avoided now:

- With, say \( p(b), q(a) \)
  \( \text{dom}(a) \)

  derive \( r(a) \lor \neg p(f(a)) \)

  then split \[ r(a) \]

  and done

\[ \checkmark \text{ Improved RR + Shifting already quite effective} \]

Next: going beyond \( \exists^* \forall^* \) by “blocking”
Our Approach

Input Clause Set

- Shifting
- Improved Range Restriction
- Blocking

Transformations

"Standard" prover

Bottom-Up Model Generation
Blocking: Idea

- Detect periodicity in models and achieve termination by exploiting standard redundancy criteria
- Example from Tambis KB

```
Chapter(a)
Book(f_{Book}(x)) \leftarrow \text{Chapter}(x)
Chapter(f_{Chapter}(x)) \leftarrow \text{Book}(x)
\bot \leftarrow \text{Chapter}(x) \land \text{Book}(x)
```

- BUMG without blocking derives infinite model:
  
  `{\text{Chapter}(a), \text{Book}(f_{Book}(a)), \text{Chapter}(f_{Chapter}(f_{Book}(a))), \ldots}`

- But same model represented finitely by
  
  `{\text{Chapter}(a), \text{Book}(f_{Book}(a))}` and `f_{Chapter}(f_{Book}(a)) \approx a`

Blocking transformation encodes this search for equations
Blocking

• If $y$ is a subterm of $x$ then speculate $x \approx y$ - or not:

$$x \approx y \lor x \not\approx y \leftarrow \text{sub}(x, y)$$
$$\bot \leftarrow x \approx y \land x \not\approx y$$

To be effective, BUMG must consider the case $x \approx y$ first

• The subterm relation, restricted to dom elements:

$$\text{sub}(x, x) \leftarrow \text{dom}(x)$$
$$\text{sub}(x, f(x_1, \ldots, x_n)) \leftarrow \text{sub}(x, x_i) \land \text{dom}(x) \land \text{dom}(f(x_1, \ldots, x_n))$$

for every $n$-ary function symbol $f \in \Sigma_f$ and all $i \in \{1, \ldots, n\}$

• This way, say, $\text{dom}(f(a))$ will be simplified to $\text{dom}(a)$ when equation $f(a) \approx a$ has been speculated

Never equates two constants, search limited to subterms in domain
Experiments

- TPTP Version 3.1.1, tried all 514 clausal satisfiable problems
- Main prover: slightly modified superposition prover MSPASS
- Environment: Linux PC, Intel Pentium 4, 3.8 GHz, 1 GByte
- Timeout 5 minutes
  Memory limit 300 MByte (never a problem for MSPASS and KRHyper)
- Results: MSPASS + our transformations vs. ...
  - ... SPASS auto mode: orthogonal
  - ... Paradox: about 20 problems unsolvable for Paradox that can be solved by our methods
- Next slides:
  - Detailed evaluation of MSPASS + our transformations
  - On non-equational problems also tried KRHyper
# MSPASS on Satisfiable TPTP Problems

| Category | # | rr −sp | rr +sp | sh ◦ rr −sp | sh ◦ rr +sp | rr ◦ bl +sp | sh ◦ rr ◦ bl +sp | crr ◦ bl +sp |
|----------|---|--------|--------|-------------|-------------|-------------|----------------|-------------|
| ALG      | 1 | 0      | 0      | 0           | 0           | 1           | 0              | 0           |
| BOO      | 13| 0      | 0      | 0           | 0           | 2           | 3              | 2           |
| COL      | 5 | 0      | 0      | 0           | 0           | 0           | 0              | 0           |
| GEO      | 1 | 0      | 0      | 0           | 0           | 0           | 0              | 0           |
| GRP      | 25| 7      | 7      | 7           | 8           | 15          | 14             | 12          |
| KRS      | 8 | 1      | 1      | 4           | 8           | 4           | 6              | 4           |
| LAT      | 1 | 0      | 0      | 0           | 0           | 1           | 1              | 0           |
| LCL      | 4 | 0      | 1      | 1           | 1           | 1           | 1              | 1           |
| MGT      | 10| 1      | 1      | 3           | 4           | 4           | 5              | 0           |
| MSC      | 1 | 1      | 1      | 1           | 1           | 1           | 1              | 1           |
| NLP      | 236| 49    | 79    | 68          | 96          | 87          | 160            | 68          |
| NUM      | 1 | 1      | 1      | 1           | 1           | 1           | 1              | 1           |
| PUZ      | 20| 6      | 6      | 6           | 6           | 10          | 10             | 9           |
| RNG      | 4 | 0      | 0      | 0           | 0           | 0           | 0              | 0           |
| SWV      | 8 | 0      | 0      | 0           | 0           | 1           | 1              | 0           |
| SYN      | 176| 20   | 50    | 20          | 52          | 124         | 125            | 120         |
| All      | 514| 86   | 147   | 111         | 177         | 252         | 328            | 218         |
## MSPASS on Satisfiable TPTP Problems

| Category | # | rr | rr | sh \(\circ\) rr | sh \(\circ\) rr | rr \(\circ\) bl | sh \(\circ\) rr \(\circ\) bl | crr \(\circ\) bl |
|----------|---|----|----|-----------------|-----------------|-----------------|--------------------|-----------------|
| ALG      | 1 | 0  | 0  | 0               | 0               | 1               | 1                  | 0               |
| BOO      | 13| 0  | 0  | 0               | 0               | 0               | 0                  | 0               |
| COL      | 5 | 0  | 0  | 0               | 0               | 0               | 0                  | 0               |
| GEO      | 1 | 0  | 0  | 0               | 0               | 0               | 0                  | 0               |
| GRP      | 25| 7  | 7  | 7               | 8               | 15              | 14                 | 12              |
| KRS      | 8 | 1  | 1  | 4               | 8               | 4               | 6                  | 4               |
| LAT      | 1 | 0  | 0  | 0               | 0               | 1               | 1                  | 0               |
| LCL      | 4 | 0  | 1  | 1               | 1               | 1               | 1                  | 1               |
| MGT      | 10| 1  | 1  | 3               | 4               | 4               | 5                  | 0               |
| MSC      | 1 | 1  | 1  | 1               | 1               | 1               | 1                  | 1               |
| NLP      | 236| 49 | 79 | 68              | 96              | 87              | 160                 | 68              |
| NUM      | 1 | 1  | 1  | 1               | 1               | 1               | 1                  | 1               |
| PUZ      | 20| 6  | 6  | 6               | 6               | 10              | 10                 | 9               |
| RNG      | 4 | 0  | 0  | 0               | 0               | 0               | 0                  | 0               |
| SWV      | 8 | 0  | 0  | 0               | 0               | 1               | 1                  | 0               |
| SYN      | 176| 20 | 50 | 20              | 52              | 124             | 125                 | 120             |
| All      | 514| 86 | 147| 111             | 177             | 252             | 328                 | 218             |

*Splitting is advisable*
## MSPASS on Satisfiable TPTP Problems

| Category | # rr | # rr | # sh o rr | sh o rr | rr o bl | sh o rr o bl | crr o bl |
|----------|------|------|-----------|---------|---------|--------------|---------|
|          | -sp | +sp  | -sp       | +sp     | +sp     | +sp          | +sp     |
| ALG      | 1    | 0    | 0         | 0       | 1       | 0            | 0       |
| BOO      | 3    | 0    | 398       |         | 3       | 2            |         |
| COL      | 5    | 0    | 0         | 0       | 0       | 0            | 0       |
| GEO      | 1    | 0    | 0         | 0       | 0       | 0            | 0       |
| GRP      | 25   | 7    | 7         |         | 14      | 12           |         |
| KRS      | 8    | 1    | 1         | 4       | 8       | 4            | 6       |
| LAT      | 1    | 0    | 0         | 0       | 0       | 1            | 1       |
| LCL      | 4    | 0    | 1         | 1       | 1       | 1            | 1       |
| MGT      | 10   | 1    | 1         | 3       | 4       | 4            | 5       |
| MSC      | 1    | 1    | 1         | 1       | 1       | 1            | 1       |
| NLP      | 236  | 49   | 79        | 68      | 96      | 87           | 160     |
| NUM      | 1    | 1    | 1         | 1       | 1       | 1            | 1       |
| PUZ      | 20   | 6    | 6         | 6       | 6       | 10           | 9       |
| RNG      | 4    | 0    | 0         | 0       | 0       | 0            | 0       |
| SWV      | 8    | 0    | 0         | 0       | 0       | 1            | 0       |
| SYN      | 176  | 20   | 50        | 20      | 52      | 124          | 125     |
| All      | 514  | 86   | 147       | 111     | 177     | 252          | 328     |

**sh o rr and rr o bl orthogonal**
## MSPASS on Satisfiable TPTP Problems

| Category | # | rr | rr | sh ○ rr | sh ○ rr | rr ○ bl | sh ○ rr ○ bl | crr ○ bl |
|----------|---|----|----|---------|---------|---------|--------------|---------|
| ALG      | 1 | 0  | 0  | 0       | 0       | 1       | 0            | 0       |
| BOO      | 13| 0  | 0  | 0       | 0       | 2       | 3            | 2       |
| COL      | 5 | 0  | 0  | 0       | 0       | 0       | 0            | 0       |
| GEO      | 1 | 0  | 0  | 0       | 0       | 0       | 0            | 0       |
| GRP      | 25| 7  | 7  | 7       | 8       | 15      | 14           | 12      |
| KRS      | 8 | 1  | 1  | 4       | 8       | 4       | 6            | 4       |
| LAT      | 1 | 0  | 0  | 0       | 0       | 1       | 1            | 0       |
| LCL      | 4 | 0  | 1  | 1       | 1       | 1       | 1            | 1       |
| MGT      | 10| 1  | 1  | 3       | 4       | 4       | 5            | 0       |
| MSC      | 1 | 1  | 1  | 1       | 1       | 1       | 1            | 1       |
| NLP      | 236| 49 | 79 | 68      | 96      | 87      | 160          | 68      |
| NUM      | 1 | 1  | 1  | 1       | 1       | 1       | 1            | 1       |
| PUZ      | 20| 6  | 6  | 6       | 6       | 10      | 10           | 9       |
| RNG      | 4 | 0  | 0  | 0       | 0       | 0       | 0            | 0       |
| SWV      | 8 | 0  | 0  | 0       | 0       | 1       | 1            | 0       |
| SYN      | 176| 20 | 50 | 20      | 52      | 124     | 125          | 120     |
| All      | 514| 86 | 147| 111     | 177     | 252     | 328          | 218     |

Shifting is generally best.
## MSPASS on Satisfiable TPTP Problems

| Category | # | rr | rr | sh ∘ rr | sh ∘ rr | rr ∘ bl | sh ∘ rr ∘ bl | crr ∘ bl |
|----------|---|----|----|---------|---------|---------|-------------|---------|
|          |   | −sp | +sp | −sp | +sp | +sp | −sp | +sp |
| ALG      | 1 | 0   | 0   | 0   | 0   | 1   | 0   | 0   |
| BOO      | 13 | 0   | 0   | 0   | 0   | 2   | 3   | 2   |
| COL      | 5 | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| GEO      | 1 | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| GRP      | 25 | 7   | 7   | 3   | 3   | 5   | 14  | 12  |
| KRS      | 8 | 1   | 1   | 4   | 8   | 4   | 6   | 4   |
| LAT      | 1 | 0   | 0   | 0   | 0   | 1   | 1   | 0   |
| LCL      | 4 | 0   | 1   | 1   | 1   | 1   | 1   | 1   |
| MGT      | 10 | 1   | 1   | 3   | 4   | 4   | 5   | 0   |
| MSC      | 1 | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| NLP      | 236 | 49  | 79  | 68  | 96  | 87  | 160 | 68  |
| NUM      | 1 | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| PUZ      | 20 | 6   | 6   | 6   | 6   | 10  | 10  | 9   |
| RNG      | 4 | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| SWV      | 8 | 0   | 0   | 0   | 0   | 1   | 1   | 0   |
| SYN      | 176 | 20  | 50  | 20  | 52  | 124 | 125 | 120 |
| **All**  | 514 | 86  | 147 | 111 | 177 | 252 | 328 | 218 |

**Note:** sh ∘ rr ∘ bl much better than crr ∘ bl.
## KRHyper on Satisfiable non-Equational Problems

| Rating | #  | MSPASS | KRHyper additional | Rating | #  | MSPASS | KRHyper additional |
|--------|----|--------|--------------------|--------|----|--------|--------------------|
| 1.00   | 4  | 0      | 0                  | 0.40   | 47 | 26     | 1                  |
| 0.80   | 57 | 24     | 4                  | 0.33   | 8  | 4      | 1                  |
| 0.67   | 26 | 5      |                    | 0.20   | 70 | 50     |                   |
| 0.60   | 44 | 23     | 10                 | 0.17   | 31 | 10     |                   |
| 0.50   | 5  | 0      |                    | 0.00   | 223| 198    | 1                  |
Conclusions

• Various improvements to BUMG paradigm, based on
  – Shifting, improved range restriction, blocking
  – Hyperresolution + splitting
    • State-of-the-art equality inference rules
    • Standard notion of redundancy
• Improves model building capabilities of standard BUMG provers
  – E.g. MSPASS, KRHyper, but not limited to these
  – Method generates domain elements on a by need basis
  – Never identifies constants (unlike finite model finders)
• Future work
  – Sorts
  – Nonmonotonic reasoning