Rewrite Rule Inference Using Equality Saturation

Chandrakana Nandi, Max Willsey, Amy Zhu, Yisu Remy Wang, Brett Saiki, Adam Anderson, Adriana Schulz, Dan Grossman, Zachary Tatlock

OOPSLA 2021
Rewrite Rules Are Ubiquitous!

CVC4

Z3

Halide

LLVM

HERBIE

tvm

Compilers

Program Synthesizers

Simplifiers / Optimizers

SMT Solvers

ML Frameworks

GCC
Rewrite Engines must be Efficient and Reliable!

Compilers

Program Synthesizers

Simplifiers / Optimizers

SMT Solvers

ML Frameworks

Performance and reliability are key for a TRS [Newcomb et al. OOPSLA'20]
But...Designing Rewrite Rules is still Hard!

Who *writes* the *rewrite* rules?
Typically hand written by experts
Time consuming, often takes years
Too few / too many rules
Unsound rules
A 3-Step Approach for Inferring Rewrite Rules

Joshi et al. 2002, Bansal et al. 2006, Singh et al. 2016, Menendez et al. 2017, ...
A 3-Step Approach for Inferring Rewrite Rules

Enumerate terms from a grammar

\[ a, b, 0, +, \ldots \]

A 3-Step Approach for Inferring Rewrite Rules

Enumerate terms from a grammar

Find candidates: interpret over concrete inputs

"Fingerprints"

A 3-Step Approach for Inferring Rewrite Rules

Enumerate terms from a grammar

\[ a, b, 0, +, \ldots \]

Find candidates: interpret over concrete inputs

\[ \text{"Fingerprints"} \]

\[(x + y) \leftrightarrow (y + x)\]

A 3-Step Approach for Inferring Rewrite Rules

Enumerate terms from a grammar

\[ a, b, 0, +, \ldots \]

Find candidates: interpret over concrete inputs

\[ \text{"Fingerprints"} \]

\[ (x + 0) \leftrightarrow x \]

A 3-Step Approach for Inferring Rewrite Rules

Enumerate terms from a grammar

- Enumerate terms from a grammar
  - `a, b, 0, +, ...`

Find candidates: interpret over concrete inputs

- Find candidates: interpret over concrete inputs
  - "Fingerprints"

Joshi et al. 2002, Bansal et al. 2006, Singh et al. 2016, Menendez et al. 2017, ...
A 3-Step Approach for Inferring Rewrite Rules

Enumerate terms from a grammar

Find candidates: interpret over concrete inputs

Filter candidates to get final ruleset

- Remove redundant rules

A 3-Step Approach for Inferring Rewrite Rules

Enumerate terms from a grammar
Exponentially many terms!

Find candidates: interpret over concrete inputs
Too many candidates, some potentially unsound!

Filter candidates to get final ruleset
Hard to find a small, useful ruleset

Equality Saturation for Inferring Rewrite Rules

This Talk:
Inferring **Small, Useful** Rulesets **Faster**
using **Equality Saturation**!
What is *Equality Saturation*?
What is *Equality Saturation*?

\[(a \times 2) / 2\]
What is *Equality Saturation*?

\[(a \times 2) / 2\]
What is *Equality Saturation*?

\[(a \times 2) / 2 \rightarrow a\]
What is *Equality Saturation*?

\[ (a \cdot 2) / 2 \rightarrow a \]

\[
\begin{align*}
(x \cdot y) / z & \leftrightarrow x \cdot (y / z) \\
y / y & \rightarrow 1 \\
x \cdot 1 & \leftrightarrow x
\end{align*}
\]

Rewrite rules!
How to Apply Rewrite Rules?

\[(a \times 2) / 2 \rightarrow a \times (2 / 2)\]

\[(x \times y) / z \leftrightarrow x \times (y / z)\]
How to Apply Rewrite Rules?

\[(a \times 2) / 2 \rightarrow a \times (2 / 2)\]

\[(x \times y) / z \leftrightarrow x \times (y / z)\]

\[a \times (2 / 2) \rightarrow a \times 1\]

\[y / y \rightarrow 1\]
How to Apply Rewrite Rules?

\[
\begin{align*}
(a \times 2) / 2 &\Rightarrow a \times (2 / 2) \\
\end{align*}
\]

\[
\begin{align*}
(a \times (2 / 2)) &\Rightarrow a \times 1 \\
\end{align*}
\]

\[
\begin{align*}
a \times 1 &\Rightarrow a \\
\end{align*}
\]
Destructively, In a Specific Order

Order of rule application affects result

Missed opportunities for optimizations

Same order may not work for all inputs

Old expression is lost

e.g., supporting commutativity is hard without additional tricks to ensure termination!
Equality Saturation Mitigates Phase Ordering!

Initial term → E-graph → Optimized term

Apply all rewrite rules!

Extract e.g., AST size
How Does Equality Saturation Work?

\[(a * 2) / 2\]
How Does Equality Saturation Work?

\[(a \times 2) / 2\]
How Does Equality Saturation Work?

\[(a \times 2) / 2\]
How Does Equality Saturation Work?

\((a \times 2) / 2\)

\((x \times y) / z \leftrightarrow x \times (y / z)\)
How Does Equality Saturation Work?

\[(a * 2) / 2\]

E-node

E-class

\[(x * y) / z \leftrightarrow x * (y / z)\]

\[(a * 2) / 2, a * (2 / 2)\]
How Does Equality Saturation Work?

(a * 2) / 2

E-node

E-class

(x * y) / z \leftrightarrow x * (y / z)

(a * 2) / 2, a * (2 / 2)

Represents both terms!
Equality Saturation for Inferring Rewrite Rules

Equality Saturation for not just applying rewrites, but to also *infer* them!
Ruler

Grammar

```
e ::= x, 0, e + e, e * e, ...
```

Interpreter

```
match e {
    | const => const
    | var (v) => lookup (v)
    | e1 + e2 => eval(e1) + eval(e2)
    | e1 * e2 => eval(e1) * eval(e2)
    ...
}
```

Validator

```
SMT / model check / fuzz
```

Rewrites

```
x + 0 = x
x * 1 = x
x - 0 = x
x / 1 = x
x + y = y + x
x + (y + z) = (x + y) + z
x * (y * z) = (x * y) * z
```

Enumeration

Candidate Generation

Rule Selection
Ruler

Grammar

\[ e ::= x, 0, e + e, e \ast e, \ldots \]

Interpreter

\[
\text{match } e \{
\text{const } \Rightarrow \text{const} \\
\text{var } (v) \Rightarrow \text{lookup } (v) \\
\text{e1 + e2 } \Rightarrow \text{eval } (e1) + \text{eval } (e2) \\
\text{e1 * e2 } \Rightarrow \text{eval } (e1) * \text{eval } (e2) \\
\ldots
\}
\]

Validator

SMT / model check / fuzz

Enumeration

Candidate Generation

Rule Selection

Rewrites

\[
\begin{align*}
x + 0 &= x \\
x + 1 &= x \\
x - 0 &= x \\
x / 1 &= x \\
x + y &= y + x \\
x + (y + z) &= (x + y) + z \\
x \ast (y \ast z) &= (x \ast y) \ast z
\end{align*}
\]
Enumeration Modulo Equality Saturation

Exponentially many terms!
Exponentially many terms!

Enumerate over an E-graph
Enumeration Modulo Equality Saturation

Exponentially many terms!

Enumerate over an E-graph

Apply current ruleset

\[(x + y) \leftrightarrow (y + x)\]
Enumeration Modulo Equality Saturation

Exponentially many terms!

Enumerate over an E-graph

Merge equivalent terms

Apply current ruleset

\[(x + y) \leftrightarrow (y + x)\]
Shrinks the term space by applying rewrites as they are learned!
Ruler

Grammar
\[ e ::= x, 0, e + e, e * e, \ldots \]

Interpreter
\[
\text{match } e \{
\text{  | } \text{const } \Rightarrow \text{const}
\text{  | } \text{var } (v) \Rightarrow \text{lookup } (v)
\text{  | } e1 + e2 \Rightarrow \text{eval } (e1) + \text{eval } (e2)
\text{  | } e1 * e2 \Rightarrow \text{eval } (e1) * \text{eval } (e2)
\ldots
\}
\]

Validator
SMT / model check / fuzz

Enumeration

Candidate Generation

Rule Selection

Rewrites
\[
\begin{align*}
x + 0 &= x \\
x * 1 &= x \\
x - 0 &= x \\
x / 1 &= x \\
x + y &= y + x \\
x + (y + z) &= (x + y) + z \\
x * (y * z) &= (x * y) * z
\end{align*}
\]
Candidate Generation by Characteristic Vector Matching

Seed initial E-classes with concrete values (cvecs) from the domain
Candidate Generation by Characteristic Vector Matching

Seed initial E-classes with concrete values (cvecs) from the domain

Compute the cvecs for newly enumerated E-classes
Candidate Generation by Characteristic Vector Matching

Seed initial E-classes with concrete values (cvecs) from the domain

Compute the cvecs for newly enumerated E-classes
Candidate Generation by Characteristic Vector Matching

Seed initial E-classes with concrete values (cvecs) from the domain

Compute the cvecs for newly enumerated E-classes

(x + y) ←→ (y + x)

(x + 0) ←→ x
Candidate Generation by Characteristic Vector Matching

Seed initial E-classes with concrete values (cvecs) from the domain

Compute the cvecs for newly enumerated E-classes

Validate candidates using SMT, fuzzing, model checking

\[ (x + y) \quad \leftrightarrow \quad (y + x) \]

\[ (x + 0) \quad \leftrightarrow \quad x \]
Rule Selection with Equality Saturation

\[
C = \begin{align*}
(x + y) & \iff (y + x) \\
(x + 0) & \iff (0 + x) \\
(y + 0) & \iff (0 + y) \\
(x \times y) & \iff (y \times x) \\
(x \times 1) & \iff (1 \times x) \\
(y \times 1) & \iff (1 \times y)
\end{align*}
\]
Rule Selection with Equality Saturation

<table>
<thead>
<tr>
<th>Equation</th>
<th>Transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>x + y</td>
<td>x + y</td>
</tr>
<tr>
<td>x * y</td>
<td>x * y</td>
</tr>
<tr>
<td>x + 0</td>
<td>x + 0</td>
</tr>
<tr>
<td>y + 0</td>
<td>y + 0</td>
</tr>
<tr>
<td>x * 1</td>
<td>x * 1</td>
</tr>
<tr>
<td>y * 1</td>
<td>y * 1</td>
</tr>
</tbody>
</table>

C = R

Rank sound candidates based on generality and pick top-k (2)
Rule Selection with Equality Saturation

Rank sound candidates based on generality and pick top-k (2)

(x + y) ↔ (y + x)
(x * y) ↔ (y * x)
(x + 0) ↔ (0 + x)
(y + 0) ↔ (0 + y)
(x * 1) ↔ (1 * x)
(y * 1) ↔ (1 * y)

Instantiate and add to rule E-graph
Rule Selection with Equality Saturation

- Rank sound candidates based on generality and pick top-k (2)

- Instantiate and add to rule E-graph

(x + y) $\leftrightarrow$ (y + x)
(x * y) $\leftrightarrow$ (y * x)
(x + 0) $\leftrightarrow$ (0 + x)
(y + 0) $\leftrightarrow$ (0 + y)
(x * 1) $\leftrightarrow$ (1 * x)
(y * 1) $\leftrightarrow$ (1 * y)
Rule Selection with Equality Saturation

(x + y) ≡ (y + x)
(x * y) ≡ (y * x)
(x + 0) ≡ (0 + x)
(y + 0) ≡ (0 + y)
(x * 1) ≡ (1 * x)
(y * 1) ≡ (1 * y)

Rank sound candidates based on generality and pick top-k (2)

Instantiate and add to rule E-graph

+  a  0  b  1

R
Rule Selection with Equality Saturation

Instantiate and add to rule E-graph

Run equality saturation
Rule Selection with Equality Saturation

All four rules are redundant and therefore discarded!

Instantiate and add to rule E-graph

Run equality saturation

(x + y) ←→ (y + x)
(x * y) ←→ (y * x)
Rule Selection with Equality Saturation

Continue processing until candidate set is empty or has only unsound ones left!

All four rules are redundant and therefore discarded!

Run equality saturation

(x + y) ↔ (y + x)
(x * y) ↔ (y * x)

Instantiate and add to rule E-graph

(x + 0) ↔ (0 + x)
(y + 0) ↔ (0 + y)
(x * 1) ↔ (1 * x)
(y * 1) ↔ (1 * y)
Rule Selection with Equality Saturation

Larger top-k makes Ruler faster
Smaller top-k gives smaller rulesets
See paper for detailed comparison!

Instantiate and add to rule E-graph

Run equality saturation

\((x + 0) \leftrightarrow (0 + x)\)
\((y + 0) \leftrightarrow (0 + y)\)
\((x \times 1) \leftrightarrow (1 \times x)\)
\((y \times 1) \leftrightarrow (1 \times y)\)
Rule Selection with Equality Saturation

Shrinks the candidate space by applying rewrites as they are learned!
Ruler

Grammar

\[ e ::= x, 0, e + e, e \times e, \ldots \]

Interpreter

\[
\text{match } e \{
\text{  | const } \Rightarrow \text{ const} \\
\text{  | var } (v) \Rightarrow \text{ lookup } (v) \\
\text{  | e1 + e2 } \Rightarrow \text{ eval(e1) + eval(e2)} \\
\text{  | e1 \times e2 } \Rightarrow \text{ eval(e1) \times eval(e2)} \\
\ldots
\}
\]

Validator

SMT / model check / fuzz

Rewrites

\[
\begin{align*}
\text{x + 0} &= x \\
\text{x \times 1} &= x \\
\text{x - 0} &= x \\
\text{x / 1} &= x \\
\text{x + y + z} &= (x + y) + z \\
\text{x \times (y \times z)} &= (x \times y) \times z
\end{align*}
\]
Equality Saturation “Soundiness”

Equality Saturation amplifies unsoundness!
Equality Saturation “Soundiness”

Equality Saturation *amplifies* unsoundness!

```
*  
/  
a 0 1
```
Equality Saturation “Soundiness”

Equality Saturation amplifies unsoundness!

```
(a * 0) ←→ 0
(y * 0) ←→ 1
```
Equality Saturation “Soundiness”

Equality Saturation *amplifies* unsoundness!

Run equality saturation on term E-graph
Equality Saturation “Soundiness”

Equality Saturation amplifies unsoundness!

Run equality saturation on term E-graph

(current ruleset)
Equality Saturation "Soundiness"

Equality Saturation amplifies unsoundness!

Run equality saturation on term E-graph

Unsound merge, 0 ≠ 1
Implementation

https://github.com/uwplse/ruler

Implemented in Rust

Uses egg for equality saturation
Evaluation

Ruler vs Other tools (CVC4)
How do the rulesets compare?
## Comparison with CVC4

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Ruler</th>
<th>CVC4</th>
<th>Ruler / CVC4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time (s)</td>
<td># Rules</td>
<td>Drv</td>
</tr>
<tr>
<td>Domain</td>
<td># Conn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bool</td>
<td>2</td>
<td>0.01</td>
<td>20</td>
</tr>
<tr>
<td>bool</td>
<td>3</td>
<td>0.06</td>
<td>28</td>
</tr>
<tr>
<td>bv4</td>
<td>2</td>
<td>0.14</td>
<td>49</td>
</tr>
<tr>
<td>bv4</td>
<td>3</td>
<td>4.30</td>
<td>272</td>
</tr>
<tr>
<td>bv32</td>
<td>2</td>
<td>13.00</td>
<td>46</td>
</tr>
<tr>
<td>bv32</td>
<td>3</td>
<td>630.09</td>
<td>188</td>
</tr>
</tbody>
</table>
## Comparison with CVC4

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Ruler</th>
<th>CVC4</th>
<th>Ruler / CVC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>Time (s)</td>
<td># Rules</td>
<td>Drv</td>
</tr>
<tr>
<td>bool</td>
<td>0.01</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>bool</td>
<td>0.06</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>bv4</td>
<td>0.14</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>bv4</td>
<td>4.30</td>
<td>272</td>
<td>1</td>
</tr>
<tr>
<td>bv32</td>
<td>13.00</td>
<td>46</td>
<td>0.97</td>
</tr>
<tr>
<td>bv32</td>
<td>630.09</td>
<td>188</td>
<td>0.98</td>
</tr>
</tbody>
</table>
## Comparison with CVC4

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Ruler</th>
<th>CVC4</th>
<th>Ruler / CVC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td># Conn</td>
<td>Time (s)</td>
<td># Rules</td>
</tr>
<tr>
<td>bool</td>
<td>2</td>
<td>0.01</td>
<td>20</td>
</tr>
<tr>
<td>bool</td>
<td>3</td>
<td>0.06</td>
<td>28</td>
</tr>
<tr>
<td>bv4</td>
<td>2</td>
<td>0.14</td>
<td>49</td>
</tr>
<tr>
<td>bv4</td>
<td>3</td>
<td>4.30</td>
<td>272</td>
</tr>
<tr>
<td>bv32</td>
<td>2</td>
<td>13.00</td>
<td>46</td>
</tr>
<tr>
<td>bv32</td>
<td>3</td>
<td>630.09</td>
<td>188</td>
</tr>
</tbody>
</table>
## Comparison with CVC4

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Ruler</th>
<th>CVC4</th>
<th>Ruler / CVC4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time (s)</td>
<td># Rules</td>
<td>Drv</td>
</tr>
<tr>
<td>Domain</td>
<td># Conn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bool</td>
<td>2</td>
<td>0.01</td>
<td>20</td>
</tr>
<tr>
<td>bool</td>
<td>3</td>
<td>0.06</td>
<td>28</td>
</tr>
<tr>
<td>bv4</td>
<td>2</td>
<td>0.14</td>
<td>49</td>
</tr>
<tr>
<td>bv4</td>
<td>3</td>
<td>4.30</td>
<td>272</td>
</tr>
<tr>
<td>bv32</td>
<td>2</td>
<td>13.00</td>
<td>46</td>
</tr>
<tr>
<td><strong>bv32</strong></td>
<td>3</td>
<td><strong>630.09</strong></td>
<td>188</td>
</tr>
</tbody>
</table>
## Comparison with CVC4

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Ruler</th>
<th>CVC4</th>
<th>Ruler / CVC4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time (s)</td>
<td># Rules</td>
<td>Drv</td>
</tr>
<tr>
<td>Domain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bool</td>
<td>0.01</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>bool</td>
<td>0.06</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>bv4</td>
<td>0.14</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>bv4</td>
<td>4.30</td>
<td>272</td>
<td>1</td>
</tr>
<tr>
<td>bv32</td>
<td>13.00</td>
<td>46</td>
<td>0.97</td>
</tr>
<tr>
<td>bv32</td>
<td>630.09</td>
<td>188</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Fraction of the 1782 rules from CVC4 that the 188 rules from Ruler can derive via equality saturation.
Comparison with CVC4

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Ruler</th>
<th>CVC4</th>
<th>Ruler / CVC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>Time (s)</td>
<td># Rules</td>
<td>Drv</td>
</tr>
<tr>
<td>bool</td>
<td>0.01</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>bool</td>
<td>0.06</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>bv4</td>
<td>0.14</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>bv4</td>
<td>4.30</td>
<td>272</td>
<td>1</td>
</tr>
<tr>
<td>bv32</td>
<td>13.00</td>
<td>46</td>
<td>0.97</td>
</tr>
<tr>
<td>bv32</td>
<td>630.09</td>
<td>188</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Ruler infers a **smaller, useful ruleset faster**
Evaluation

Ruler vs Other tools (CVC4)
How do the rule sets compare?

Ruler vs Humans (Herbie)
Can Ruler compete with experts?
Comparison with Human-written Rules

\[ \sqrt{x+1} - \sqrt{x} \rightarrow \frac{1}{\sqrt{x+1} + \sqrt{x}} \]

Herbie detects inaccurate expressions and finds more accurate replacements. The red expression is inaccurate when \( x > 1 \); Herbie's replacement, in blue, is accurate for all \( x \).
Comparison with Human-written Rules

52 rational rules, designed by the developers over 6 years

55 / 155 benchmarks are purely over rational arithmetic

\[ \sqrt{x+1} - \sqrt{x} \rightarrow \frac{1}{\sqrt{x+1} + \sqrt{x}} \]

Herbie detects inaccurate expressions and finds more accurate replacements. The red expression is inaccurate when \( x > f \); Herbie's replacement, in blue, is accurate for all \( x \).
Comparison with Human-written Rules

52 rational rules, designed by the developers over 6 years

55 / 155 benchmarks are purely over rational arithmetic

Herbie can generate more-complex expressions that aren't more precise #261

Closed  nbraud opened this issue on Aug 31, 2019 · 4 comments
Comparison with Human-written Rules

52 rational rules, designed by the developers over 6 years

55 / 155 benchmarks are purely over rational arithmetic

Herbie can generate more-complex expressions that aren't more precise #261

Discovered by Ruler, resolved the GitHub issue!
End-to-End: Rational Herbie

None: Remove all rules
Herbie: Herbie without any changes
Ruler: Herbie with Ruler’s rules
Both: Herbie with both original and Ruler’s rules
Rational Herbie: Comparing Accuracy

None: Remove all rules
Herbie: Herbie without any changes
Ruler: Herbie with Ruler’s rules
Both: Herbie with both original and Ruler’s rules

Ruler’s rules are at least as good as the original Herbie rules
Rational Herbie: Comparing AST Size

None: Remove all rules

Herbie: Herbie without any changes

Ruler: Herbie with Ruler’s rules

Both: Herbie with both original and Ruler’s rules

Ruler’s rules are at least as good as the original Herbie rules
Rational Herbie: Comparing AST Size

None: Remove all rules
Herbie: Herbie without any changes
Ruler: Herbie with Ruler’s rules
Both: Herbie with both original and Ruler’s rules

Ruler’s rules are at least as good as the original Herbie rules

See paper for more results!
Rewrite Rule Inference Using Equality Saturation

- **Enumerate terms from a grammar**
  - a, b, 0, +, ...

- **Find candidates: interpret over concrete inputs**
  - "Fingerprints"

- **Filter candidates to get final ruleset**
  - Remove redundant rules

  \[
  x + 0 \leftrightarrow 0 + x \\
  y + 0 \leftrightarrow 0 + y \\
  x + y \leftrightarrow y + x
  \]

Equality Saturation improves all three steps!

**Ruler:** [https://github.com/uwplse/ruler](https://github.com/uwplse/ruler)