Automatic and Practical Equational Deduction

William McCune

Mathematics and Computer Science Division
Argonne National Laboratory
U.S.A.

Presented at the 5th AI & Math Symposium
Fort Lauderdale, Florida
January 4, 1998

Knuth-Bendix Completion

A minimal basis for group theory

$$e * x = x$$
$$x^{-1} * x = e$$
$$(x * y) * z = x * (y * z)$$

A good set of rules for simplifying group theory expressions

$$e * x \longrightarrow x$$
$$x * e \longrightarrow x$$
$$x^{-1} * x \longrightarrow e$$
$$x * x^{-1} \longrightarrow e$$
$$e^{-1} \longrightarrow e$$
$$(x^{-1})^{-1} \longrightarrow x$$
$$(x * y)^{-1} \longrightarrow y^{-1} * x^{-1}$$
$$(x * y) * z \longrightarrow x * (y * z)$$
$$x * (x^{-1} * y) \longrightarrow y$$
$$x^{-1} * (x * y) \longrightarrow y$$

The Knuth-Bendix Procedure

- Infer new equations (*superposition*).
- Try to make all equations into (terminating) rewrite rules.
- Keep all equations reduced w.r.t. each other.

If the procedure succeeds, we have a

complete set of reductions,

which is a decision procedure for equality of terms in the equational theory.

Failure of K-B Completion

- A complete set of reductions might not exist.
  - undecidablility of the word problem
- It might exist, but be hard to find.
  - a bad term ordering
  - hard to prove termination
  - take a long time
- What if a derived equation cannot be made into a rewrite rule?
  - For example, $x*y = y*x$ does not terminate.

Extensions of K-B Completion

- Build equations into the inference and simplification processes.
  - commutativity (C)
  - associativity and commutativity (AC)
  - arbitrary equational theory?
- Unfailing completion
  - Use nonorientable equations as constrained rewrite rules.
- Nonequational theories
- Search for proofs
  - Deny the goal, and search for a contradiction.
  - Don't worry about nonorientable equations.
  - This can be a complete proof procedure for equational logic.

How do We Search for Proofs?

- Start with a procedure similar to K-B completion.
- For theories with associative-commutative (AC) operators, use AC unification instead of AC equations.
  - *with a heuristic*
- Don't worry too much about completeness.
- Use *basic paramodulation*, a refinement (constraint) of the inference process.
- Select *carefully* the inferences to make.
  - Combine breadth-first and shortest-first search.
- If the search fails, examine the output, and try again with a different search strategy.

AC Unification

Paramodulation: Given $t = s,\qquad f[t'] = r,$
where t and t' can be made identical by unification (instantiating variables),
infer the corresponding instance of $f[s] = r$.

With AC unification, t and t' can be made identical in many (doubly exponential) different ways.

The pair of terms $x + x + x$ and $y + z + u + v$ has 1,044,569 most general AC unifiers.

A heuristic:

- Use only the simplest AC unifiers.
- This is an incomplete (and severe) restriction.
- Applied to the prolific pair above, we get 139 AC unifiers.
- In practice, it doesn't block many proofs.

The Robbins Problem

- A well-known basis for Boolean algebra [Huntington, 1933]:

$x + y = y + x$	(C) Commutativity
$(x + y) + z = x + (y + z)$	(A) Associativity
$\overline{(\overline{x} + y)} + \overline{(\overline{x} + \overline{y})} = x$	(H) Huntington

- The Robbins Problem:

$$\overline{(\overline{(x + \overline{y})} + \overline{(x + y)})} = x \qquad \text{(R) Robbins}$$

Is $\{C, A, R\}$ basis for Boolean algebra?

- A proof of the following would solve the problem.

Robbins $\Longrightarrow_{\text{AC}}$ Huntington.

Easy Lemmas About the Robbins Equation

A very easy one:

$$\left\{ \begin{array}{l} \text{Robbins} \\ \forall x\ (\overline{\overline{x}} = x) \end{array} \right\} \Longrightarrow_{\mathrm{AC}} \text{Huntington.}$$

Another easy one (but not first-order):

$$\left\{ \begin{array}{l} \text{Robbins} \\ \text{over a finite set} \end{array} \right\} \Longrightarrow_{\mathrm{AC}} \text{Huntington.}$$

Existence of a zero:

$$\left\{ \begin{array}{l} \text{Robbins} \\ \exists x \forall y\ (x + y = y) \end{array} \right\} \Longrightarrow_{\mathrm{AC}} \text{Huntington.}$$

Idempotence:

$$\left\{ \begin{array}{l} \text{Robbins} \\ \forall x\ (x + x = x) \end{array} \right\} \Longrightarrow_{\mathrm{AC}} \text{Huntington.}$$

Existence of an idempotent element:

$$\left\{ \begin{array}{l} \text{Robbins} \\ \exists x\ (x + x = x) \end{array} \right\} \Longrightarrow_{\mathrm{AC}} \text{Huntington.}$$

Hard Lemmas About the Robbins Equation

These were formulated and proved by Steve Winker c. 1980 [J. Algebra 1992].

Lemma 1 (An absorption):

$$\left\{ \begin{array}{l} \text{Robbins} \\ \exists x \exists y\ (x + y = x) \end{array} \right\} \Longrightarrow_{\mathrm{AC}} \text{Huntington.}$$

Lemma 2 (An absorption within complement):

$$\left\{ \begin{array}{l} \text{Robbins} \\ \exists x \exists y\ (\overline{x + y} = \overline{x}) \end{array} \right\} \Longrightarrow_{\mathrm{AC}} \text{Huntington.}$$

Winker's proofs are not first-order (induction).

These lemmas were proved automatically at Argonne in 1992 and 1995.

The Final Assault

- About 50 searches with various strategies, including 14 multiday searches.
  - Many small attempts to achieve *well-behaved search*.
- Success: Robbins $\Longrightarrow_{\mathrm{AC}} \exists x \exists y\ (\overline{x + y} = \overline{x})$
  - 8 day search, 30 megabytes of memory.
  - Basic paramod., AC heuristic.
  - Small inference tasks.
  - Equal parts breadth-first, shortest-first search.
  - Discard equations with more than 70 symbols.
- 5 weeks of real time, 5 weeks of CPU time.
- Quicker proofs and shorter proofs were found later.

Solutions of the Robbins Problem

First proof (EQP, 8 days of search):

$$\text{Robbins} \Longrightarrow_{\mathrm{AC}} \exists x \exists y\ (\overline{x + y} = \overline{x})$$

Another proof (EQP, a bit more search):

$$\text{Robbins} \Longrightarrow_{\mathrm{AC}} \exists x \exists y\ (x + y = x)$$

Yet another proof (EQP, 10%–90% more search):

$$\text{Robbins} \Longrightarrow_{\mathrm{AC}} \text{Huntington}$$

A Proof that Solves the Robbins Problem

Theorem. *Robbins* $\Longrightarrow_{\mathsf{AC}} \exists c \exists d\ (c+d=c)$.

7	$\overline{\overline{\overline{x}+y}+\overline{x+y}} = y$	[Robbins Equation]
10	$\overline{\overline{\overline{x+y}+\overline{x}+y}+y} = \overline{x+y}$	[7 → 7]
11	$\overline{\overline{\overline{\overline{x}+y}+x+y}+y} = \overline{\overline{x}+y}$	[7 → 7]
29	$\overline{\overline{\overline{\overline{x}+y}+x+2y}+\overline{\overline{x}+y}} = y$	[11 → 7]
54	$\overline{\overline{\overline{\overline{\overline{x}+y}+x+2y}+\overline{\overline{x}+y}+z}+\overline{y+z}} = z$	[29 → 7]
217	$\overline{\overline{\overline{\overline{\overline{x}+y}+x+2y}+\overline{\overline{x}+y}+\overline{y+z}+z}+z} = \overline{y+z}$	[54 → 7]
674	$\overline{\overline{\overline{\overline{\overline{\overline{x}+y}+x+2y}+\overline{\overline{x}+y}+\overline{y+z}+z}+z+u}+\overline{\overline{y+z}+u}} = u$	[217 → 7]
6736	$\overline{\overline{\overline{\overline{3x}+x}+\overline{3x}}+\overline{\overline{\overline{3x}+x}+5x}} = \overline{\overline{3x}+x}$	[10 → 674]
8855	$\overline{\overline{\overline{3x}+x}+5x} = \overline{3x}$	[6736 → 7 :54]
8865	$\overline{\overline{\overline{\overline{3x}+x}+\overline{3x}+2x}+\overline{3x}} = \overline{\overline{3x}+x}+2x$	[8855 → 7]
8866	$\overline{\overline{\overline{3x}+x}+\overline{3x}} = x$	[8855 → 7: 11]
8870	$\overline{\overline{\overline{\overline{3x}+x}+\overline{3x}+y}+\overline{x+y}} = y$	[8866 → 7]
8871	$\overline{\overline{3x}+x}+2x = 2x$	[8865 :8870]

Equation 8871 asserts the existence of an object c, namely $2x$, and an object d, namely $\overline{\overline{3x}+x}$, such that $c+d=c$. Q.E.D. (Note that 8855 satisfies $\exists c \exists d(\overline{c+d}=\overline{c})$.)

Step 6736 in Detail

Premises:

$$\overline{\overline{\overline{x+y}+\overline{x}+y}+y} = \overline{x+y} \qquad (10)$$

$$\overline{\overline{\overline{\overline{\overline{\overline{x}+y}+x+2y}+\overline{\overline{x}+y}+\overline{y+z}+z}+z+u}+\overline{\overline{y+z}+u}} = u \qquad (674)$$

With (10), let x be $3v$, and y be $\overline{\overline{3v}+v}+2v$:

$$\underbrace{\overline{\overline{\overline{3v+\overline{\overline{3v}+v}+2v}+\overline{3v}+\overline{\overline{3v}+v}+2v}+\overline{\overline{3v}+v}+2v}} = \overline{3v+\overline{\overline{3v}+v}+2v} \qquad (10')$$

With (674), let x be $3v$, y be v, z be $2v$, and u be $\overline{\overline{3v}+v}$:

$$\overline{\underbrace{\overline{\overline{\overline{\overline{\overline{3v}+v}+5v}+\overline{\overline{3v}+v}+\overline{3v}+2v}+2v+\overline{\overline{3v}+v}}}+\overline{\overline{3v}+\overline{\overline{3v}+v}}} = \overline{\overline{3v}+v} \qquad (674')$$

Replace the braced term of (674′), which is AC-identical to the left-hand side of (10′), with the right-hand side of (10′):

$$\overline{\overline{3v+\overline{\overline{3v}+v}+2v}+\overline{\overline{3v}+\overline{\overline{3v}+v}}} = \overline{\overline{3v}+v}.$$

Rename the variable and rearrange:

$$\overline{\overline{\overline{\overline{3x}+x}+\overline{3x}}+\overline{\overline{\overline{3x}+x}+5x}} = \overline{\overline{3x}+x}. \qquad [10{\rightarrow}674]\ (6736)$$

Achieving Well Behaved Search

Some Examples

- Limit size of equations.
  - Incomplete, but necessary in practice.
  - Find a *good* limit.
- Compare results of different trial searches.
  - If a search is more economical and it gives a similar set of equations, use it.
- Restart the search with new lemmas.
  - Clean out the deadwood and focus on the lemmas.
- Observe derived equations, try to get more syntactic variety, iterate.

Quantum Logic

An *Ortholattice* (OL) satisfies

$x \vee y = y \vee x$	$x \wedge y = y \wedge x$
$(x \vee y) \vee z = x \vee (y \vee z)$	$(x \wedge y) \wedge z = x \wedge (y \wedge z)$
$x \vee (x \wedge y) = x$	$x \wedge (x \vee y) = x$
$x'' = x$	
$(x \wedge y)' = x' \vee y'$	
$x \vee x' = 1$	$x \wedge x' = 0$
$x \vee 1 = 1$	$x \wedge 0 = 0$

An *orthomodular lattice* (OML) also satisfies

$$x \vee (x' \wedge (x \vee y)) = x \vee y$$

- Ordinary probability: Boolean algebra over a set of events.
- Probability in quantum dynamics: Orthomodular lattice over a Hilbert space.

Ortholattice Questions
(from Norman Megill)

The following three equations hold for OML. Do they hold also for OL?

$$((a \wedge b') \vee a')' \vee ((a \wedge b') \vee ((a' \wedge ((a \vee b') \wedge (a \vee b))) \vee (a' \wedge ((a \vee b') \wedge (a \vee b))'))) = 1 \qquad \text{(E1)}$$

$$(a \vee ((a' \wedge ((a \vee b') \wedge (a \vee b))) \vee (a' \wedge ((a' \wedge b) \vee (a' \wedge b'))))) = 1 \qquad \text{(E2)}$$

$$(((a' \wedge b) \vee (a' \wedge b')) \vee (a \wedge (a' \vee b)))' \vee (a' \vee b) = 1 \qquad \text{(E3)}$$

The following fails for OL.
Does it hold for OML?

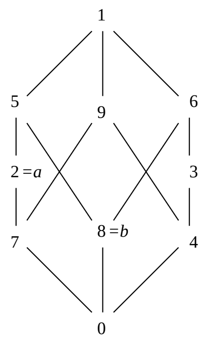
$$a' \vee ((b' \vee a) \wedge (b \vee (b' \wedge a))) = 1 \qquad \text{(E7)}$$

Searching for Counterexamples

- Look for finite models of first-order statements.
  - constraint satisfaction problem
- Existing programs
  - MACE (Argonne National Laboratory)
  - SEM (Jian Zhang, Hantao Zhang, U Iowa)
  - FINDER (John Slaney, ANU)
- Results with these programs
  - quasigroup existence problems
  - axiom systems for various theories
  - test conjectures

Ortholattice Answers

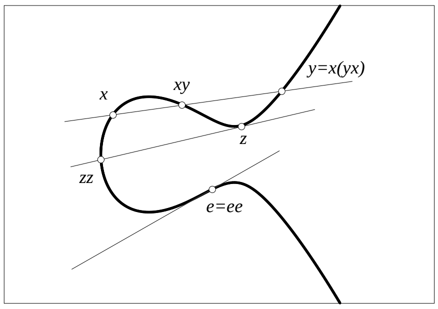
- E1: counterexample (by MACE, in 20 min.):



An ortholattice violating E1.

- E2: proof (by EQP, in 4 sec.), OL $\Rightarrow$ E2.
- E3: proof (by EQP, in 23 hr.), OL $\Rightarrow$ E3.
- E7: proof (by EQP, in 13 min.), OML $\Rightarrow$ E7.

Algebraic Geometry
(with R. Padmanabhan)



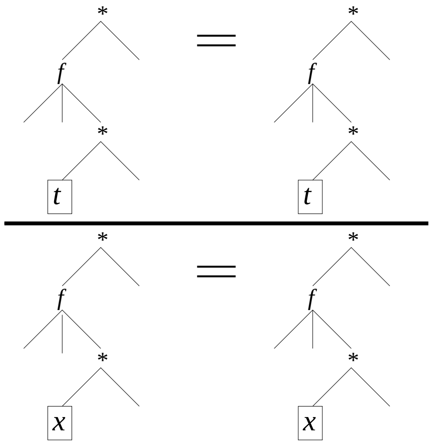
Consider a nonsingular cubic curve in the complex projective plane. Define a binary operation on the points of the curve.

- If a line intersects the curve in 3 points, the product of any two points is the third point.
- zz is where the tangent at z intersects the curve.
- Inflection points are idempotent elements ($ee = e$).
- The operation is commutative, and the Steiner law holds: $x(yx) = y$.

All other operations are defined in terms of this one.

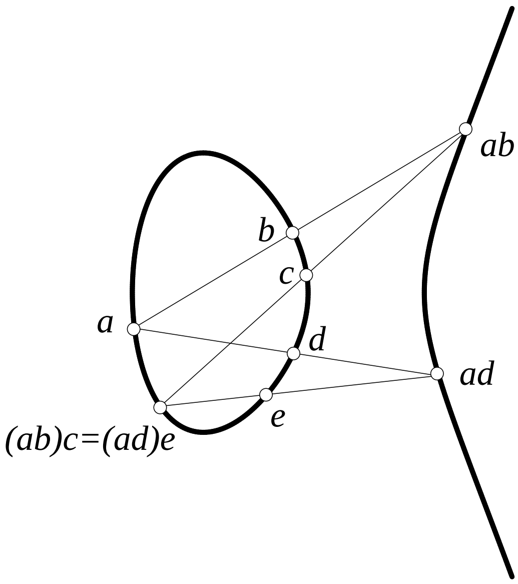
The Rule $=(gL)\Rightarrow$

(R. Padmanabhan)



- If identical terms (t) occur in corresponding positions, and in the same nest of function symbols, replace both occurrences with a new variable (x).

A Geometric Example of $=(gL)\Rightarrow$



- $[(ab)c = (ad)e]$ $=(gL)\Rightarrow$ $[\forall x\ (xb)c = (xd)e]$.
- Select a point x and see if it works.

Strategies for Applying $=(gL)\Rightarrow$

It is incompatible with Knuth-Bendix.

- $=(gL)\Rightarrow$ likes to apply to *simplifiable* terms, so we cannot always rewrite.
- We must paramodulate from smaller sides into smaller sides.
- We must paramodulate from variables and into variables.

We must relax (or eliminate) the K-B constraints.

New Results with $=(gL)\Rightarrow$

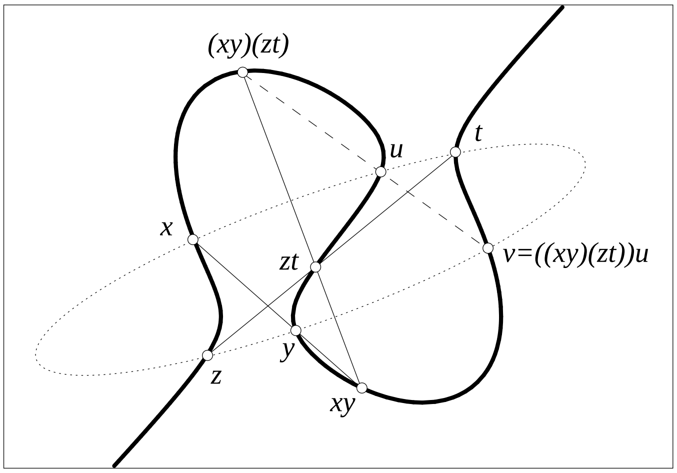
- A cancellative semigroup on a cubic curve must be commutative.
- Uniqueness inversive groupoid on a cubic curve.
- General exploration of $=(gL)\Rightarrow$ algebras.
- Uniqueness of the 5-ary Steiner Law:

$$\left\{\begin{array}{l} \text{full permutability of } f \\ f(e,e,e,e,e) = e \\ f(e,e,e,x,f(e,e,e,x,y)) = y \\ \text{full permutability of } g \\ g(e,e,e,e,e) = e \\ g(e,e,e,x,g(e,e,e,x,y)) = y \end{array}\right\} =(gL)\Rightarrow$$

$$g(x,y,z,u,v) = f(x,y,z,u,v).$$

A corollary: given 5 points of intersection of a cubic and a conic, the sixth point can be found be a simple ruler construction.

Conic-cubic Construction



Automated Equational Deduction Types of Use

- Stand-alone (Robbins problem)
  - try to get the programs to prove the interesting theorems
  - TRC – combinatory logic - new foundations
- Support for higher-level work (quantum logic)
  - semigroup conjecture of Padmanabhan
  - formal methods, verification systems
- Exploration of new methods (cubic curves)

What About More Interesting Theories?

- Set theory can axiomatized equationally:
  - *is our work done???*
- The practical examples are in simple theories
  - first-order
  - few symbols
  - few axioms
  - no defined concepts

Reference

Available at

`http://www.mcs.anl.gov/home/mccune/`

- these slides
- information on other applications
- the theorem provers Otter and EQP
- the model searcher MACE
- Son of BirdBrain
  - try Otter and MACE via the Web