Logic and Mechanized Reasoning

Decision Procedures for Linear Arithmetic

Marijn J.H. Heule

Carnegie Mellon University

Second Midterm Exam

The second midterm is on Tuesday, March 25, during class

- last name starts with A-H are in room GHC 4307
- last name starts with K-Z are in Doherty 2210

The exam will cover:

- DP and DPLL, following the slides from the 2/11 lecture
- Sections 8.2, 8.3, and 8.4 in the textbook
- Chapters 9-12 in the textbook
- Construct unifiers of terms by hand, but not the algorithm

Linear Real Arithmetic

Fourier-Motzkin

A Full Decision Procedure

Other Theories

Linear Real Arithmetic

Fourier-Motzkin

A Full Decision Procedure

Other Theories

Linear Expressions and Linear Constraints

A linear expression is of the form $a_1x_1 + a_2x_2 + \cdots + a_nx_n + b$

- a_i is a rational number
- b is a rational number
- x_i is a variable (ranging over the real numbers)

Linear Expressions and Linear Constraints

A linear expression is of the form $a_1x_1 + a_2x_2 + \cdots + a_nx_n + b$

- a_i is a rational number
- b is a rational number
- x_i is a variable (ranging over the real numbers)

A linear constraint is of the form $s < t$ or $s = t$

- s and t are linear expressions
- $s \leq t$ can be expressed as $(s < t) \vee (s = t)$
- $s \neq t$ can be expressed as $(s < t) \vee (t < s)$

Linear Expressions and Linear Constraints

A linear expression is of the form $a_1x_1 + a_2x_2 + \cdots + a_nx_n + b$

- a_i is a rational number
- b is a rational number
- x_i is a variable (ranging over the real numbers)

A linear constraint is of the form $s < t$ or $s = t$

- s and t are linear expressions
- $s \leq t$ can be expressed as $(s < t) \vee (s = t)$
- $s \neq t$ can be expressed as $(s < t) \vee (t < s)$

We use only $s < t$ and $s = t$ to simplify the presentation

Rewriting Linear Constraints

Any linear constraint can be turned into either $t = 0$ of $t < 0$

- Move everything to the left-hand side

Example

Consider the constraint: $3x + 2y < 3y + 4z$.
Which can be rewritten to: $3x - y - 4z < 0$.

A linear constraint with x can become $x = t$, $x < t$, or $t < x$

- Move everything (apart from x) to the right-hand side
- Divide the right-hand side by the left-hand side constant
- Do the reverse if the constant of x is negative

Example

Consider again the constraint: $3x + 2y < 3y + 4z$.
Which can be rewritten to: $x < \frac{1}{3}y + \frac{4}{3}z$.

Satisfiability of Linear Constraints is Decidable

Theorem

The question as to whether a finite set of linear constraints is satisfiable is decidable.

Proof.

Proof by induction on the number of variables

- Base case: only constraints $b_0 = b_1$ and $b_0 < b_1$
- Inductive case: eliminate a variable x
- Substitute an equality containing x
- Eliminate the inequalities containing x

□

Inductive Case: Eliminate a Variable by Substitution

If there is an equality containing variable x

- Rewrite the constraint to the form $x = t$
- Substitute all occurrences of x by t
- The resulting new problem is equisatisfiable
- Given a solution to the new problem, assign x the value of t

This reduces the number of variables by one and the number of constraints by one (possibly more by removing trivial ones)

Inductive Case: Eliminate Inequalities

Partition the inequalities in Γ:

- those that don't contain x at all
- those that can be expressed in the form $s_i < x$
- those that can be expressed in the form $x < t_j$

Inductive Case: Eliminate Inequalities

Partition the inequalities in Γ:

- those that don't contain x at all
- those that can be expressed in the form $s_i < x$
- those that can be expressed in the form $x < t_j$

Γ': Replace the last two parts with all inequalities $s_i < t_j$

- Any assignment that satisfies Γ, satisfied Γ'

Inductive Case: Eliminate Inequalities

Partition the inequalities in Γ:

- those that don't contain x at all
- those that can be expressed in the form $s_i < x$
- those that can be expressed in the form $x < t_j$

Γ': Replace the last two parts with all inequalities $s_i < t_j$

- Any assignment that satisfies Γ, satisfied Γ'

A solution to Γ' can be turned into a solution of Γ

- Determine the largest s_i and the smallest t_j
- Assign x to be a value somewhere in between
- If part s_i or t_j is missing make x sufficiently small or large

Complexity

What is the complexity of the procedure?

- Consider a formula with n variables and m inequalities

Complexity

What is the complexity of the procedure?

- Consider a formula with n variables and m inequalities

Costs of eliminating a single variable:

- A variable may occur in $\frac{m}{2}$ inequalities of the form $s_i < x$
- A variable may occur in $\frac{m}{2}$ inequalities of the form $x < t_j$
- Eliminating such a variable increases the size from m to $\frac{m^2}{4}$

Complexity

What is the complexity of the procedure?

- Consider a formula with n variables and m inequalities

Costs of eliminating a single variable:

- A variable may occur in $\frac{m}{2}$ inequalities of the form $s_i < x$
- A variable may occur in $\frac{m}{2}$ inequalities of the form $x < t_j$
- Eliminating such a variable increases the size from m to $\frac{m^2}{4}$

Total costs: $\mathcal{O}(m^{2^n})$

Why Doesn't This Work for the Integers?

Why Doesn't This Work for the Integers?

There may be an integer between s_i and t_j

Why Doesn't This Work for the Integers?

There may be an integer between s_i and t_j

Consider the following inequalities:

$$\begin{aligned} y &< 6 - x \\ 2x - 4 &< y \\ 2 &< x \end{aligned}$$

Why Doesn't This Work for the Integers?

There may be an integer between s_i and t_j

Consider the following inequalities:

$$\begin{aligned} y &< 6 - x \\ 2x - 4 &< y \\ 2 &< x \end{aligned}$$

Eliminating y results in the following inequality:

$$2x - 4 < 6 - x \;\equiv\; x < \frac{10}{3}$$

Why Doesn't This Work for the Integers?

There may be an integer between s_i and t_j

Consider the following inequalities:

$$\begin{array}{rcl} y & < & 6 - x \\ 2x - 4 & < & y \\ 2 & < & x \end{array}$$

Eliminating y results in the following inequality:

$$2x - 4 < 6 - x \;\equiv\; x < \frac{10}{3}$$

So $x = 3$ is a solution, but there is no solution for y:

- $y < 6 - 3 \equiv y < 3$
- $2 \cdot 3 - 4 < y \equiv 2 < y$

Linear Real Arithmetic

Fourier-Motzkin

A Full Decision Procedure

Other Theories

Fourier and Motzkin

Jean-Baptiste Joseph Fourier (1768 - 1830)

- French mathematician
- Many scientific contributions, including Fourier Series, Fourier Transformation, and FM Elimination

Theodore Motzkin (1908 - 1970)

- Israeli-American mathematician
- Influenced linear programming, optimization, combinatorics, and algebraic geometry
- Rediscovered FM Elimination

Fourier-Motzkin Example (1)

Consider the following inequalities:

$$\begin{aligned} x + y &< 7 \\ y + z &< 6 \\ x - z &< 4 \end{aligned} \qquad \begin{aligned} z &< x - 2 \\ 1 &< y \\ 0 &< z \end{aligned}$$

Fourier-Motzkin Example (1)

Consider the following inequalities:

$$\begin{array}{rcl} x+y & < & 7 \\ y+z & < & 6 \\ x-z & < & 4 \end{array} \qquad \begin{array}{rcl} z & < & x-2 \\ 1 & < & y \\ 0 & < & z \end{array}$$

Eliminate z: Rewriting the inequalities to $s < z$ or $z < t$:

$$\begin{array}{rcl} z & < & 6-y \\ x-4 & < & z \end{array} \qquad \begin{array}{rcl} z & < & x-2 \\ 0 & < & z \end{array}$$

Fourier-Motzkin Example (1)

Consider the following inequalities:

$$\begin{aligned} x+y &< 7 \\ y+z &< 6 \\ x-z &< 4 \end{aligned} \qquad \begin{aligned} z &< x-2 \\ 1 &< y \\ 0 &< z \end{aligned}$$

Eliminate z: Rewriting the inequalities to $s < z$ or $z < t$:

$$\begin{aligned} z &< 6-y \\ x-4 &< z \end{aligned} \qquad \begin{aligned} z &< x-2 \\ 0 &< z \end{aligned}$$

Compute all pairs $s < t$:

$$\begin{aligned} x-4 &< x-2 && \equiv && -4 < -2 \\ x-4 &< 6-y && \equiv && x+y < 10 \\ 0 &< x-2 && \equiv && 2 < x \\ 0 &< 6-y && \equiv && y < 6 \end{aligned}$$

Fourier-Motzkin Example (2)

Example after eliminating z and simplification:

$$\begin{aligned} x + y &< 7 \qquad & 2 &< x \\ 1 &< y \qquad & y &< 6 \end{aligned}$$

Eliminate y: Rewriting the inequalities to $s < y$ or $y < t$:

$$\begin{aligned} y &< 7 - x \\ y &< 6 \\ 1 &< y \end{aligned}$$

Compute all pairs $s < t$:

$$\begin{aligned} 1 &< 7 - x & \equiv \quad & x < 6 \\ 1 &< 6 & \equiv \quad & 1 < 6 \end{aligned}$$

Fourier-Motzkin Example (3)

Example after eliminating z and y and simplification:

$$\begin{array}{rcl} x & < & 6 \\ 2 & < & x \end{array}$$

Which is satisfiable.

Pick a value for x within the range, say 4. Determine y:

$$\begin{array}{rcl} y & < & 7-4 \\ y & < & 6 \\ 1 & < & y \end{array}$$

Now, we can pick a value for y to determine z, etc.

Heuristics

Although the worst-case complexity is double exponential, Fourier-Motzkin Elimination can be quite efficient in practice

Heuristics can limit the number of inequalities in practice

- Remove in each step the least occurring variable
- Only make elimination steps that keep the constants at 1

Implemented in various automated reasoning tools

- Some SAT solvers using FME preprocessing
- Also used in some SMT solvers

Fourier-Motzkin in Lean

```
-- first, eliminate all the equations
partial def elimEqConstraints : List LinearExp → List LinearExp → Option (List LinearExp)
  | []       , gts => some gts
  | eq :: eqs, gts => Id.run do
    let (x, a) := eq.getTerm
    let u      := eq.erase x
    let newEqs := substituteEqConstraints a x u eqs
    match substituteGtConstraints a x u gts with
      | some newGts => elimEqConstraints newEqs newGts
      | none        => none

-- then eliminate variables from the `e > 0` constraints
partial def elimGtConstraints : List LinearExp → Bool
  | []        => true
  | gt :: gts => Id.run do
      let x := gt.getTerm.1
      match elimVarGtConstraints x (gt :: gts) with
        | some gts => elimGtConstraints gts
        | none     => false

def FourierMotzkin (eqs gts : List LinearExp) : Bool :=
match elimEqConstraints eqs gts with
  | some gts => elimGtConstraints gts
  | none     => false
```

Linear Real Arithmetic

Fourier-Motzkin

A Full Decision Procedure

Other Theories

Clean Constraints

First, consider a problem in linear arithmetic

- Variables (in the reals) are labeled $x_1, x_2, \ldots, x_n$
- Constraints are labeled $c_1, c_2, \ldots, c_m$
- $\exists x_1, x_2, \ldots, x_n . c_1 \wedge c_2 \wedge \ldots c_m$

And consider the structure $(\mathbb{R}, 0, 1, +, <)$.

- Express $3x$ by $x + x + x$
- Express $x - (1/2)y + (4/3)z < 0$ by $6x + 8z < 3y$

We apply Fourier-Motzkin if all constraints are $s < t$ or $s = t$

- How to deal with constraints of a different form?

Arbitrary Constraints

First, turn the formula in negation normal form

- replace $\neg(s < t)$ by $(t < s) \vee (s = t)$
- (in practice, it is better to include $\leq$ in the language)
- replace $s \neq t$ by $(s < t) \vee (t < s)$

Second, turn the NNF in disjunctive normal form

- Solve each cube using Fourier-Motzkin elimination.
- Satisfiable if one of the cubes is satisfiable

$(\mathbb{Q}, 0, 1, +, <, \leq)$ is Decidable

Note that in all reasoning so far, we only required that we can always find a number in between two other numbers

- This does not only hold for $\mathbb{R}$, but also for $\mathbb{Q}$

The same procedure decides questions in $(\mathbb{Q}, 0, 1, +, <, \leq)$

Presburger Arithmetic

What happens if we replace the real numbers with integers?

Presburger Arithmetic

What happens if we replace the real numbers with integers?

$(\mathbb{Z}, 0, 1, +, <)$, or Presburger arithmetic, is also decidable

- First proven by Presburger in 1926
- Also known as linear integer arithmetic
- Integers are discrete: no number between x and $x + 1$

Presburger Arithmetic

What happens if we replace the real numbers with integers?

$(\mathbb{Z}, 0, 1, +, <)$, or Presburger arithmetic, is also decidable

- First proven by Presburger in 1926
- Also known as linear integer arithmetic
- Integers are discrete: no number between x and $x + 1$

The decision procedure is more complicated

- SMT solvers have an efficient algorithm
  - for the quantifier-free fragment

$(\mathbb{R}, 0, 1, +, \times, <)$ is Decidable

What about adding multiplication to the language?

- Still decidable
- Extending linear arithmetic with $p = 0$ and $p < 0$ for arbitrary polynomials p
- Known as Real closed fields
- Decidability proved by Alfred Tarski before World War II, but only published in 1948

$(\mathbb{Z}, 0, 1, +, \times)$ is Undecidable

Is $(x + y = z) \wedge (x \neq 0) \wedge (y \neq 0)$ satisfiable?

$(\mathbb{Z}, 0, 1, +, \times)$ is Undecidable

Is $(x + y = z) \wedge (x \neq 0) \wedge (y \neq 0)$ satisfiable?

- Trivial: $x = 1$, $y = 2$, $z = 3$

$(\mathbb{Z}, 0, 1, +, \times)$ is Undecidable

Is $(x + y = z) \wedge (x \neq 0) \wedge (y \neq 0)$ satisfiable?

- Trivial: $x = 1$, $y = 2$, $z = 3$

Is $(x^2 + y^2 = z^2) \wedge (x \neq 0) \wedge (y \neq 0)$ satisfiable?

$(\mathbb{Z}, 0, 1, +, \times)$ is Undecidable

Is $(x + y = z) \land (x \neq 0) \land (y \neq 0)$ satisfiable?

- Trivial: $x = 1$, $y = 2$, $z = 3$

Is $(x^2 + y^2 = z^2) \land (x \neq 0) \land (y \neq 0)$ satisfiable?

- Easy (Pythagorean triple): $x = 3$, $y = 4$, $z = 5$

$(\mathbb{Z}, 0, 1, +, \times)$ is Undecidable

Is $(x + y = z) \wedge (x \neq 0) \wedge (y \neq 0)$ satisfiable?

- Trivial: $x = 1$, $y = 2$, $z = 3$

Is $(x^2 + y^2 = z^2) \wedge (x \neq 0) \wedge (y \neq 0)$ satisfiable?

- Easy (Pythagorean triple): $x = 3$, $y = 4$, $z = 5$

Is $(x^4 + y^4 = z^4) \wedge (x \neq 0) \wedge (y \neq 0)$ satisfiable?

$(\mathbb{Z}, 0, 1, +, \times)$ is Undecidable

Is $(x + y = z) \wedge (x \neq 0) \wedge (y \neq 0)$ satisfiable?

- Trivial: $x = 1$, $y = 2$, $z = 3$

Is $(x^2 + y^2 = z^2) \wedge (x \neq 0) \wedge (y \neq 0)$ satisfiable?

- Easy (Pythagorean triple): $x = 3$, $y = 4$, $z = 5$

Is $(x^4 + y^4 = z^4) \wedge (x \neq 0) \wedge (y \neq 0)$ satisfiable?

- Non-trivial, unsatisfiable