Quantifier Elimination

Helpful lemmas

Recall that $\forall F := \forall x_1 \ldots \forall x_n\, F$ where $x_1, \ldots, x_n$ are the free variables of F.

Lemma

Let S be a set of sentences. $S \models F$ iff $S \models \forall F$

Proof. Exercise.

Lemma

Let S be a set of sentences. If $S \models F \leftrightarrow G$ then $S \models H \leftrightarrow H[G/F]$.

Proof. By structural induction on H. Exercise.

Quantifier elimination

Definition

Let T be a set of formulas. We say that F and F' are T-equivalent if $T \models F \leftrightarrow F'$

Definition

A theory T admits quantifier elimination if for every formula F there is a quantifier-free T-equivalent formula G such that $fv(G) \subseteq fv(F)$. We call G a quantifier-free T-equivalent of F.

Examples

Find quantifier-free equivalent formulas in linear real arithmetic for:

$$
\begin{aligned}
\exists x \exists y\, (3 * x + 5 * y = 7) &\quad \leftrightarrow \quad ? \\
\exists y\, (x < y \wedge y < z) &\quad \leftrightarrow \quad ? \\
\forall y\, (x < y \wedge y < z) &\quad \leftrightarrow \quad ?
\end{aligned}
$$

Quantifier elimination

A quantifier-elimination procedure (QEP) for a theory T and a set of formulas $\mathcal{F}$ is an algorithm that computes for every formula of $\mathcal{F}$ a quantifier-free T-equivalent.

Lemma

Let T be a theory such that

- *T has a QEP for all formulas and*
- *$T \models G$ or $T \models \neg G$ for all ground formulas G (quantifier-free formula without occurrences of variables), and it is decidable which is the case.*

Then T is decidable and complete.

Quantifier elimination

Proof. Decidability.

This algorithm decides whether $T \models F$ for given F (sentence or not):

> Compute a quantifier-free T-equivalent G of $\forall F$.
> Decide whether $T \models G$ or $T \models \neg G$.
> If $T \models G$ then answer $T \models F$, otherwise $T \not\models F$.

Correctness of the algorithm:

$$T \models F \quad \text{iff} \quad T \models \forall F \quad \text{iff} \quad T \models G$$

where we have made use of the lemmas.

Completeness. Exercise.

Simplifying quantifier elimination: one $\exists$

Fact

If T has a QEP for all formulas of the form $\exists x\, F$, where F is quantifier-free,
then T has a QEP for all formulas.

Essence: It is sufficient to be able to eliminate a single $\exists$

Construction:

Given: a QEP $qe1$ for formulas of the form $\exists x\, F$ where F is quantifier-free

Define: a QEP for all formulas
Method: Eliminate quantifiers bottom-up by $qe1$, use $\forall \equiv \neg\exists\neg$

Simplifying quantifier elimination: $\exists x \ \bigwedge$ *literals*

Fact

If T has a QEP for all $\exists x\, F$ where F is a conjunction of literals,
all of which contain x,
then T has a QEP for all $\exists x\, F$ where F is quantifier-free.

Construction:

Given: a QEP $qe1c$ for formulas of the form $\exists x\,(L_1 \wedge \cdots \wedge L_n)$
where each L_i is a literal that contains x

Define: $qe1(\exists x\, F)$ where F is quantifier-free
Method: Put F in DNF. Distribute $\exists$ over $\vee$. Apply $qe1c$.

This is the end of the generic part of quantifier elimination.
The rest is theory specific.

Simplifying quantifier elimination: Eliminating "$\neg$"

(Motivation: $\neg x < y \;\leftrightarrow\; y < x \lor y = x$ for linear orderings)

Fact

Assume that there is a computable function aneg that maps every negated atom to a quantifier-free and negation-free T-equivalent formula.

If T has a QEP for all $\exists x\, F$ where F is a conjunction of atoms, all of which contain x,
then T has a QEP for all $\exists x\, F$ where F is quantifier-free.

Construction:

Given: a QEP *qe1ca* for formulas of the form $\exists x\, (A_1 \land \cdots \land A_n)$ where each atom A_i contains x

Define: $qe1(\exists x\, F)$ where F quantifier-free
Method: Put F into NNF. Apply *aneg*. Put F in DNF. Distribute $\exists$ over $\lor$. Apply *qe1ca*.

Quantifier Elimination

Dense Linear Orders Without Endpoints

Dense Linear Orders Without Endpoints

Definition

Let $\Sigma = \{<, =\}$. The theory of dense linear order without endpoints (DLO) is the set of Σ-sentences that are consequences of the following set of axioms:

$$\forall x \forall y \forall z\, (x < y \wedge y < z \rightarrow x < z)$$

$$\forall x\, \neg(x < x)$$

$$\forall x \forall y\, (x < y \vee x = y \vee y < x)$$

$$\forall x \forall z\, (x < z \rightarrow \exists y\, (x < y \wedge y < z))$$

$$\forall x \exists y\; x < y$$

$$\forall x \exists y\; y < x$$

Models of DLO?

Theorem

All countable models of DLO are isomorphic.

Elimination of "$\neg$"

Fact

DLO has a computable function aneg that maps every negated atom to a quantifier-free and negation-free DLO-equivalent formula.

$$
\begin{aligned}
DLO &\models \neg(x = y) \leftrightarrow x < y \vee y < x \\
DLO &\models \neg(x < y) \leftrightarrow x = y \vee y < x
\end{aligned}
$$

$$
\begin{aligned}
aneg(\neg(x = y)) &= x < y \vee y < x \\
aneg(\neg(x < y)) &= x = y \vee y < x
\end{aligned}
$$

Quantifier elimination for conjunctions of atoms

We define a QEP for formulas of the form $\exists x\,(A_1 \wedge \cdots \wedge A_n)$, where x occurs in every A_i.

$qe1ca(\exists x\,(A_1 \wedge \cdots \wedge A_n)$:

- If some A_i is of the form $x = y$ (x and y different), apply:
  $$\exists x\,(x = t \wedge F) \;\equiv\; F[t/x] \qquad (x \text{ does not occur in } t)$$
  and return $F[y/x]$.
- Drop all A_i of the form $y = y$. If no A_i left return $\top$.
- If some A_i is of the form $y < y$, return $\bot$.
- Separate the A_i into lower and upper bounds for x.
  If no lower and/or no upper bounds, return $\top$. Otherwise use
  $$DLO \models \exists x \left(\bigwedge_{i=1}^{m} l_i < x \wedge \bigwedge_{j=1}^{n} x < u_j \right) \leftrightarrow \bigwedge_{i=1}^{m} \bigwedge_{j=1}^{n} l_i < u_j$$
  and return $\bigwedge_{i=1}^{m} \bigwedge_{j=1}^{n} l_i < u_j$.

Quantifier elimination for conjunctions of atoms

Example

$$\exists x\,(x < z \land y < x \land x < w) \quad \equiv_{DOL} \quad y < z \land y < w$$

$$\begin{aligned}
\forall x\,\forall y\,(x < y) \quad &\equiv_{DOL} \quad \forall x\,\neg\exists y\,\neg(x < y) \\
&\equiv_{DOL} \quad \forall x\,\neg\exists y\,(y < x \lor x = y) \\
&\equiv_{DOL} \quad \forall x\,\neg(\exists y\, y < x \lor \exists y\, x = y) \\
&\equiv_{DOL} \quad \forall x\,\neg(\top \lor x = x) \\
&\equiv_{DOL} \quad \forall x\,\bot \\
&\equiv_{DOL} \quad \bot
\end{aligned}$$

$$\begin{aligned}
\exists x\,\exists y\,\exists z\,(x < y \land y < z \land z < x) \quad &\equiv_{DOL} \quad \exists x\,\exists y\,(x < y \land y < x) \\
&\equiv_{DOL} \quad \exists x\, x < x \\
&\equiv_{DOL} \quad \bot
\end{aligned}$$

Complexity

Quadratic blow-up with each elimination step

Therefore, eliminating all $\exists$ from

$$\exists x_1 \ldots \exists x_m \; F$$

where F has length n needs $O(n^{2^m})$ time and space, assuming F is DNF.

More efficient algorithms exist.

Consequences

- *DLO* has quantifier elimination.
- *DLO* is decidable and complete.
- All models of DLO (for example $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$) are elementarily equivalent:

  We cannot distinguish models of DLO by first-order formulas.

Quantifier Elimination
Linear real arithmetic

Linear real arithmetic

Definition

Let $\mathcal{R}_+ = (\mathbb{R}, 0, 1, +, <, =)$. Linear real arithmetic is the theory $R_+ = Th(\mathcal{R}_+)$.

We allow the following additional function symbols:

For every $c \in \mathbb{Q}$:

- c is a constant symbol
- $c\cdot$, multiplication with c, is a unary function symbol

Example

$$\forall x\, \exists y\, \big(2.1x - y < 4.6 \land \forall z\, (7.3 < 2z + 4.7y \lor 3.25z > 2x)\big)$$

Linear real arithmetic

Fact

Every formula with these additional function symbols can be effectively transformed into an $\mathcal{R}_+$-equivalent formula of R_+.

For example:

$$2.1(x - y) < 4.6$$

is $\mathcal{R}_+$-equivalent to

$$\underbrace{(x + \cdots + x)}_{21 \ \textit{times}} < \underbrace{1 + \cdots + 1}_{46 \ \textit{times}} + \underbrace{y + \cdots + y}_{21 \ \textit{times}}$$

Linear real arithmetic

Definition

A term t is in normal form if $t = c_1 \cdot x_1 + \ldots + c_n \cdot x_n + c$ where $c_i \neq 0$ and $x_i \neq x_j$ for every $1 \leq i \neq j \leq n$.

An atom A is in normal form (NF) if $A = 0 \bowtie t$ for some term t in normal form, where $\bowtie \in \{<, =\}$.

An atom is solved for x if it is of the form $x < t$, $x = t$ or $t < x$, where x does not occur in t.

Fact

Every atom is $\mathcal{R}_+$-equivalent to an atom in normal form.

Any atom in normal form that contains x can be effectively transformed into a $\mathcal{R}_+$-equivalent atom solved for x.

We let $sol_x(A)$ denote the result of solving an atom A for x.

Elimination of "$\neg$"

Fact

R_+ has a computable function aneg that maps every negated atom to a quantifier-free and negation-free $\mathcal{R}_+$-equivalent formula.

$$R_+ \models \neg(t = t') \leftrightarrow t < t' \vee t' < t$$
$$R_+ \models \neg(t < t') \leftrightarrow t = t' \vee t' < t$$

$$aneg(\neg(t = t')) = t < t' \vee t' < t$$
$$aneg(\neg(t < t')) = t = t' \vee t' < t$$

Fourier-Motzkin Elimination

We define a QEP for formulas of the form $\exists x\, (A_1 \wedge \cdots \wedge A_n)$, where all A_i are atoms in NF and x occurs in every A_i.

$qe1ca(\exists x\, (A_1 \wedge \cdots \wedge A_n)$:

- Let $S = \{sol_x(A_1), \ldots, sol_x(A_n)\}$
- If $(x = t) \in S$ for some t, apply:

  $$\exists x\, (x = t \wedge F) \;\equiv\; F[t/x] \qquad (x \text{ does not occur in } t)$$

  and return $F[y/x]$.
- Separate the A_i into lower and upper bounds for x.
  If no lower and/or no upper bounds, return $\top$. Otherwise

  $$\text{return} \bigwedge_{(l<x)\in S} \bigwedge_{(x<u)\in S} l < u$$

  All returned formulas are implicitly put into NF.

Fourier-Motzkin Elimination

Examples

$$
\begin{array}{rl}
 & \exists x\, \exists y\; (3x + 5y < 7 \wedge 2x - 3y < 2) \\
\equiv_{\mathcal{R}_+} & \exists x\, \exists y\; ({}^2\!/\!_3 x - {}^2\!/\!_3 < y \;\wedge\; y < {}^7\!/\!_5 - {}^3\!/\!_5 x) \\
\equiv_{\mathcal{R}_+} & \exists x\; ({}^2\!/\!_3 x - {}^2\!/\!_3 < {}^7\!/\!_5 - {}^3\!/\!_5 x) \\
\equiv_{\mathcal{R}_+} & \exists x\; (x < {}^{31}\!/\!_{19}) \\
\equiv_{\mathcal{R}_+} & \top
\end{array}
$$

$$
\begin{array}{rl}
 & \exists x\, \forall y\; (3y \leq x \vee x \leq 2y) \\
\equiv_{\mathcal{R}_+} & \exists x\, \neg \exists y\; \neg(3y \leq x \vee x \leq 2y) \\
\equiv_{\mathcal{R}_+} & \exists x\, \neg \exists y\; \neg(x < 3y \wedge 2y < x) \\
\equiv_{\mathcal{R}_+} & \exists x\, \neg \exists y\; \left({}^1\!/\!_3 x < y \wedge y \leq {}^1\!/\!_2 x)\right) \\
\equiv_{\mathcal{R}_+} & \exists x\; \neg({}^1\!/\!_3 x < {}^1\!/\!_2 x) \\
\equiv_{\mathcal{R}_+} & \exists x\; ({}^1\!/\!_3 x \geq {}^1\!/\!_2 x) \\
\equiv_{\mathcal{R}_+} & \exists x\; (x \leq 0) \\
\equiv_{\mathcal{R}_+} & \top
\end{array}
$$

Can DNF (recall miniscoping) be avoided?

Ferrante and Rackoff's theorem

Theorem

Let F be quantifier-free and negation-free (not necessarily in DNF!) and assume all atoms that contain x are solved for x. Let

$$\begin{array}{rcl} L & = & \{l \mid (l < x) \in S_x\} \\ U & = & \{u \mid (x < u) \in S_x\} \end{array} \quad E \;=\; \{t \mid (x = t) \in S_x\}$$

where S_x is the set of atoms in F that contain x. Then:

$$R_+ \models \exists x\, F \;\leftrightarrow\; F[-\infty/x] \vee F[\infty/x] \vee \bigvee_{t \in E} F[t/x] \vee \bigvee_{l \in L} \bigvee_{u \in U} F[0.5(l+u)/x]$$

where $F[-\infty/x]$ ($F[\infty/x]$) is the result of applying this transformation to all solved atoms in F:

$$x < t \;\mapsto\; \top\ (\bot) \quad t < x \;\mapsto\; \bot\ (\top) \quad x = t \;\mapsto\; \bot\ (\bot)$$

Ferrante-Rackoff Procedure

$qe1(\exists x\ F)$:

1. Put F into NNF, eliminate all negations,
   put all atoms into normal form,
   solve those atoms for x that contain x.

2. Apply Ferrante and Rackoff's theorem.

Theorem

Eliminating all quantifiers with Ferrante and Rackoff's procedure from a formula of size n takes space $O(2^{cn})$ and time $O(2^{2^{dn}})$.

Quantifier Elimination
Presburger Arithmetic

See [Harrison] or [Enderton] under “Presburger”

Presburger Arithmetic

Definition
Let $\mathcal{Z}_+ := (\mathbb{Z}, +, 0, 1, \leq)$. Linear integer arithmetic is the theory $Z_+ = Th(\mathcal{Z}_+)$.

We allow additional function symbols as for linear arithmetic:

For every $c \in \mathcal{Z}$:

- c is a constant symbol
- $c\cdot$, multiplication with c, is a unary function symbol

Fact
Linear integer arithmetic does not have quantifier elimination

Proof. Show that no quantifier-free formula is $\mathcal{Z}_+$-equivalent to $\exists x\, x + x = y$.

Presburger Arithmetic

Definition

Let $\mathcal{P} := (\mathbb{Z}, +, 0, 1, \leq, 2\,|, 3\,|, \ldots)$, where $k\,|$ is a unary predicate symbol interpreted as "k divides ...".
Presburger Arithmetic is the theory $P := Th(\mathcal{P})$.

Definition

An atom A is in normal form if

$$
\begin{aligned}
A &= 0 \leq c_1 \cdot x_1 + \ldots + c_n \cdot x_n + c \quad \text{or} \\
A &= k \mid c_1 \cdot x_1 + \ldots + c_n \cdot x_n + c
\end{aligned}
$$

where $c_i \in \mathbb{Z} \setminus \{0\}$ and $k \geq 1$

Where necessary, atoms are put into normal form

Fact

Every atom is $\mathcal{P}$-equivalent to an atom in normal form.

Elimination of "$\neg$"

Fact

P has a computable function aneg that maps every negated atom to a quantifier-free and negation-free $\mathcal{P}$-equivalent formula.

$$\mathcal{P} \models \neg(s \leq t) \leftrightarrow t+1 \leq s$$
$$\mathcal{P} \models \neg(k \mid t) \leftrightarrow k \mid t+1 \vee k \mid t+2 \vee \cdots \vee k \mid t+(k-1)$$

$$\begin{aligned} aneg(\neg(s \leq t)) &= \quad t+1 \leq s \\ aneg(\neg(k \mid t)) &= \quad k \mid t+1 \vee k \mid t+2 \vee \cdots \vee k \mid t+(k-1) \end{aligned}$$

Quantifier Elimination for $\mathcal{P}$

We define a QEP for formulas of the form $\exists x\ (A_1 \wedge \cdots \wedge A_n)$, where all A_i are atoms in NF and x occurs in every A_i.

$qe1ca(\exists x\ (A_1 \wedge \cdots \wedge A_n))$ proceeds in two steps.

Let $F = A_1 \wedge \cdots \wedge A_n$.

Step 1: Set all coeffs of x in F to 1 or -1:

1. Set all coeffs of x in F to the lcm m of all coeffs of x
2. Set all coeffs of x to 1 or -1 and add $\wedge\ m \mid x$

Step 2: Separate into upper and lower bounds, with some new reasoning because of the $k \mid t$ atoms.

Quantifier Elimination for $\mathcal{P}$: Step 1

The result of Step 1 is

$$R := \mathit{coeff}\,1(A_1) \wedge \cdots \wedge \mathit{coeff}\,1(A_l) \ \wedge m \mid x$$

where

- m is the (positive) lcm of all coefficients of x in F (e.g. lcm $\{-6, 9\} = 18$), and
- $$\begin{array}{rcl} \mathit{coeff}\,1(0 \leq \Sigma_{i=1}^n c_i x_i + c) & = & (0 \leq \Sigma_{i=1}^n c'_i x_i + c') \quad \text{and} \\ \mathit{coeff}\,1(d \mid \Sigma_{i=1}^n c_i x_i + c) & = & (d' \mid \Sigma_{i=1}^n c'_i x_i + c') \end{array}$$
  where, assuming x is the k-th variable x_k and letting $m' = m/|c_k|$ we set:

$$\begin{array}{rclcrcl} c'_i & = & m' \cdot c_i \ \mathit{if} \ i \neq k & \quad & c'_k & = & \mathit{if} \ c_k > 0 \ \mathit{then} \ 1 \ \mathit{else} \ -1 \\ c' & = & m' \cdot c & & d' & = & m' \cdot d \end{array}$$

Lemma $\mathcal{P} \models (\exists x\ F) \leftrightarrow (\exists x\ R)$

Quantifier Elimination for $\mathcal{P}$: Step 2

The result of Step 2 is the quantifier-free formula

$$R' := \quad \textit{if } \ L = \emptyset \quad \text{then} \ \bigvee_{i=0}^{m-1} \bigwedge D[i/x] \quad \text{else} \ \bigvee_{i=0}^{m-1} \bigvee_{l \in L} R[l + i/x]$$

where

$$\begin{array}{ll} L & := \{-t \mid (0 \leq x + t) \in R\} \\ U & := \{t \mid (0 \leq -x + t) \in R\} \\ D & := \{(d \mid t) \in R\} \\ m & := \text{(positive) lcm of } \{d \mid (d \mid t) \in D \text{ for some } t\} \end{array}$$

The correctness of tis step follows from the

Lemma (Periodicity Lemma)

If $A \in D$, i.e. $A = (d \mid x + t)$ and $x \notin fv(t)$, and $i \equiv j \pmod d$ then $\mathcal{P} \models A[i/x] \leftrightarrow A[j/x]$.

Example

We eliminate the existential quantifier of

$$\exists x\,(3x - y + 1 > 0) \wedge (2x - 6 < z) \wedge (4 \mid 5x + 1)$$

We set all coefficients of x to 1. With $\mathrm{lcm}\{3, 2, 5\} = 30$ we get:

$$\exists x\,(30x - 10y + 10 > 0) \wedge (30x - 90 < 15z) \wedge (24 \mid 30x + 6)$$

We rescale $x := 30x$ adding the atom $(30 \mid x)$ and split the inequalities in lower and upper bounds:

$$\exists x\,(10y - 10 < x) \wedge (x < 90 + 15z) \wedge (24 \mid x + 6) \wedge (30 \mid x)$$

Example

Periodicity: if $(24 \mid x+6) \wedge (30 \mid x)$ holds for some x, then it also holds for every $x + k \cdot \mathrm{lcm}\{24, 30\} = x + k \cdot 120$ where $k \in \mathbb{Z}$.

Therefore: $(10y - 10 < x) \wedge (x < 90 + 15z)$ is satisfied by x iff it is satisfied by $10y - 10 + i$ for some $1 \leq i \leq 120$.

So $\exists x\,(10y - 10 < x < 90 + 15z) \wedge (24 \mid x+6) \wedge (30 \mid x)$ is $\mathcal{P}$-equivalent to

$$\bigvee_{i=1}^{120} \left(\begin{array}{c} 10y - 10 < 10y - 10 + i < 90 + 15z \\ \wedge \\ (24 \mid 10y - 10 + i + 6) \wedge (30 \mid 10y - 10 + i) \end{array} \right)$$

and so $\mathcal{P}$-equivalent to:

$$\bigvee_{i=1}^{120} \left(\begin{array}{c} (10y + i < 100 + 15z) \\ \wedge \\ (24 \mid 10w - 4 + i) \wedge (30 \mid 10w - 10 + i) \end{array} \right)$$