An Abstract Domain Extending Difference-Bound Matrices with Disequality Constraints

Mathias Péron and Nicolas Halbwachs

SYNCHRON'06 - december 1, 2006

Motivations for disequalities

Conviction

integer variables are used to address objects in many situations

- usefulness of the invariant $x \neq y$
  - alias phenomena: $A[x]$ and $A[y]$
  - other client analsis, optimization, independence analysis

Framework static verification, abstract interpretation theory

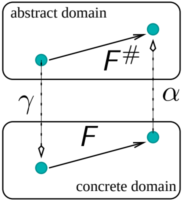
- allows conservative verification, computing an over-approximation of the fixpoint
- notion of *abstract domain*

Abstract Interpretation (1/2)

- Theory of Program's Dynamical Behavior Approximation

Problems complex values manipulation, iterative resolution of the fixpoint equation

- abstraction



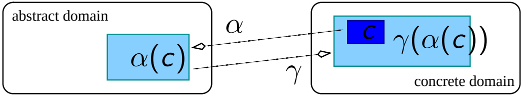
- conservative verification
- convergence

Abstract Interpretation (1/2)

- Theory of Program's Dynamical Behavior Approximation

Problems complex values manipulation, iterative resolution of the fixpoint equation

- abstraction
- conservative verification



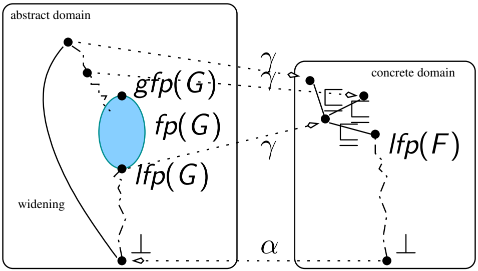
- convergence

Abstract Interpretation (1/2)

- Theory of Program's Dynamical Behavior Approximation

Problems complex values manipulation, iterative resolution of the fixpoint equation

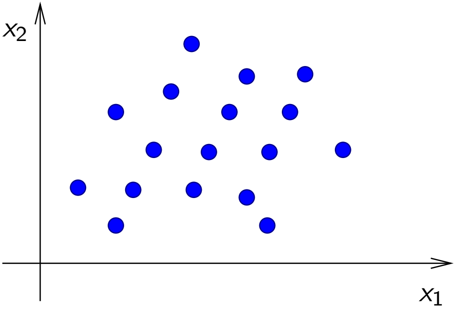
- abstraction
- conservative verification
- convergence



Abstract Interpretation (2/2)

- Classicals numerical abstract domains

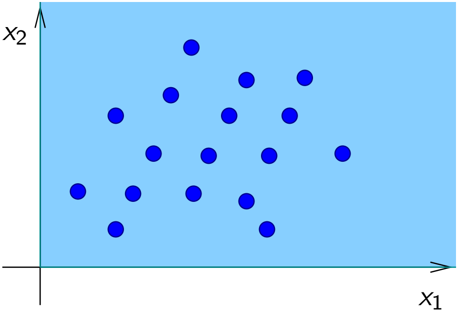
Abstraction of a set of states:



Abstract Interpretation (2/2)

- Classicals numerical abstract domains

Abstraction of a set of states: non-relational domains

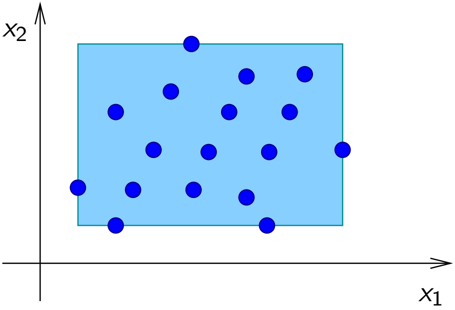


signs $0 \leq x_i$

Abstract Interpretation (2/2)

- Classicals numerical abstract domains

Abstraction of a set of states: non-relational domains

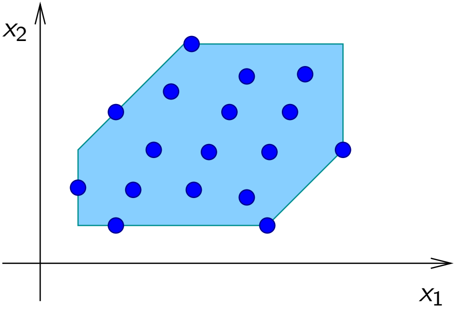


signs
intervals $lb \leq x_i \leq ub$

Abstract Interpretation (2/2)

- Classicals numerical abstract domains

Abstraction of a set of states: 2-relational domains

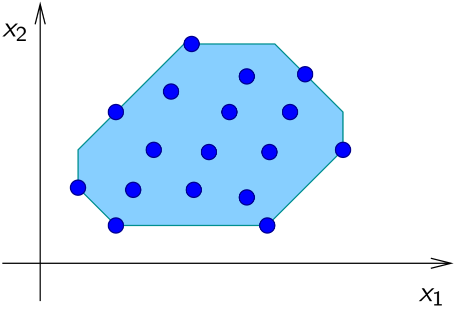


signs
intervals
zones (DBMs) $x_i - x_j \leq c$

Abstract Interpretation (2/2)

- Classicals numerical abstract domains

Abstraction of a set of states: 2-relational domains

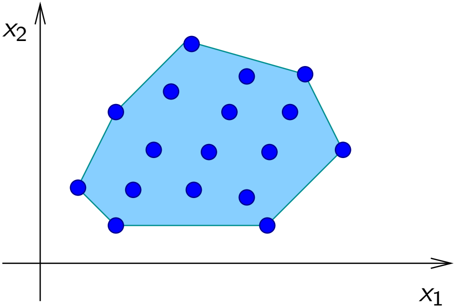


signs
intervals
zones (DBMs)
octagons $\pm x_i \pm x_j \leq c$

Abstract Interpretation (2/2)

- Classicals numerical abstract domains

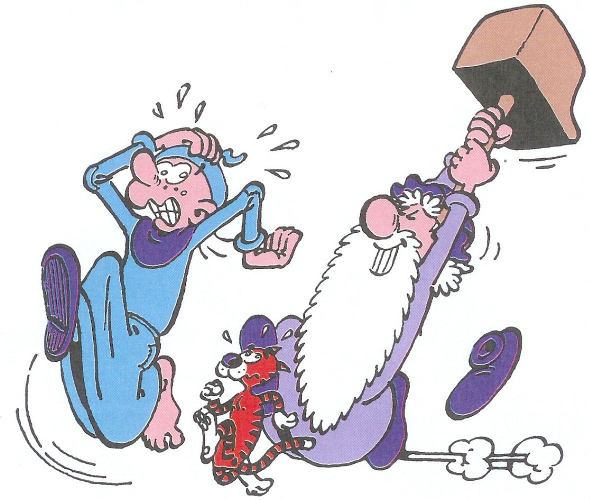
Abstraction of a set of states: n-relational domains



signs
intervals
zones (DBMs)
octagons
convex polyedra $\sum a_i x_i \leq c_i$

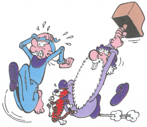
Disequations: Related Works

I'm a plagiarist !



Disequations: Related Works

I'm a plagiarist !



Consideration in abstract interpretation

- classical abstract domains are convex
- dynamic partitioning techniques

In other fields

finite unions of convex sets (MC), constraint propagation (CLP), etc

Which domain for disequalities ?

Goal

Extend an existing domain without increasing its complexity

Disequalities + equalities

a too poor analysis

- trivial deductions
  - $(x = y \wedge y = z) \Rightarrow x = z$
  - $(x = y \wedge x \neq z) \Rightarrow y \neq z$

Disequalities + ordering relations

enrich the deduction power

- non completely trivial deductions may be done

$(x \leq y \leq z \wedge x \neq y) \Rightarrow x \neq z$

Which domain for disequalities ?

DBM is a good candidate

$$c_1 \le x \le c_2$$
$$c_1 \le x - y \le c_2$$

Which domain for disequalities ?

DBM is a good candidate

$$c_1 \leq x \leq c_2 \qquad x \neq 0$$
$$c_1 \leq x - y \leq c_2 \quad x \neq y$$

- allow strict inequalities $x < y$
- respect our goal: $x - y \neq c$ impose unbounded representation

Which domain for disequalities ?

DBM is a good candidate

$$c_1 \leq x \leq c_2 \qquad x \neq 0$$
$$c_1 \leq x - y \leq c_2 \quad x \neq y$$

- allow strict inequalities $x < y$
- respect our goal: $x - y \neq c$ impose unbounded representation

---

Outline

- Difference-Bound Matrices
- *disequalities* Difference-Bound Matrices
- Application to Program Analysis

Difference-Bound Matrices *(Dill 89)*

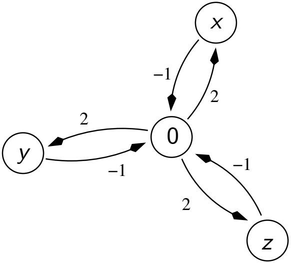
Var: finite set of variables $\{v_0\} \cup \{v_1, \ldots, v_{n-1}\}$
$\mathcal{V}$: variables domain, $\mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{R}$
$\overline{\mathcal{V}}$: extension of $\mathcal{V}$ with $+\infty$

Constraints ($c \in \mathcal{V}$)

$$constraint ::= v_i \leq c \mid v_i - v_j \leq c$$

Representation

$$1 \leq x, y, z \leq 2 \qquad \begin{array}{c} \begin{array}{cccc} \mathbf{0} & \mathbf{x} & \mathbf{y} & \mathbf{z} \end{array} \\ \begin{pmatrix} +\infty & -1 & -1 & -1 \\ 2 & +\infty & +\infty & +\infty \\ 2 & +\infty & +\infty & +\infty \\ 2 & +\infty & +\infty & +\infty \end{pmatrix} \end{array}$$



Emptiness Test, Closure

Satisfiability

- checking for the existence of negative cycles

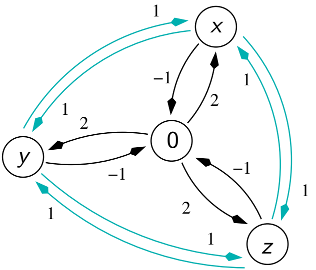
Closure *(for non-empty DBMs)*

- infering implicit constraints

shortest-path closure is well defined

e.g. Floyd-Warshall algorithm ($O(n^3)$)

$1 \leq x, y, z \leq 2$ $\quad \begin{array}{c} \begin{array}{cccc} \mathbf{0} & \mathbf{x} & \mathbf{y} & \mathbf{z} \end{array} \\ \left(\begin{array}{cccc} 0 & -1 & -1 & -1 \\ 2 & 0 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{array} \right) \end{array}$



Domain, Order, Normal Form

The shortest-path closure leads to a normal form

Domain

- $\mathcal{D}(M) = \{(s_1, \ldots, s_{n-1}) \in \mathcal{V}^{n-1} \mid \forall i, j \in [0..n-1] \quad s_j - s_i \leq M_{ij}^{\leq} \ \wedge \ s_0 = 0\}$

Order

- $M \trianglelefteq M' \iff \forall i, j \; M_{ij} \leq M'_{ij}$

property: $M \trianglelefteq M' \Rightarrow \mathcal{D}(M) \subseteq \mathcal{D}(M')$

Normal form *(for non-empty DBMs)*

- $\overline{M} = \inf_{\trianglelefteq} \{M' \mid \mathcal{D}(M') = \mathcal{D}(M)\}$

Complexity computing normal form, deciding emptiness, usual operations: $O(n^3)$

disequalities Difference-Bound Matrices *(VMCAI 07)*

Constraints ($c \in \mathcal{V}$)

$$constraint ::= v_i \leq c \mid v_i - v_j \leq c \mid v_i \neq 0 \mid v_i - v_j \neq 0$$

Representation

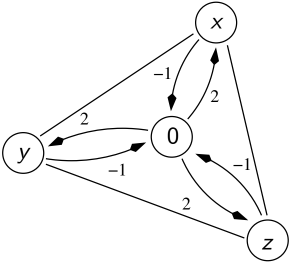
- dDBM: a pair of matrices $(M^{\leq}, M^{\neq})$

$M^{\leq}$ is a classical DBM $\qquad M^{\neq}$ is a symmetric boolean matrix

- disequal potential graph

$$\begin{cases} 1 \leq x, y, z \leq 2 \\ x \neq y \\ x \neq z \\ y \neq z \end{cases}$$

$$\begin{array}{cccc} \mathbf{0} & \mathbf{x} & \mathbf{y} & \mathbf{z} \end{array}$$
$$\begin{pmatrix} +\infty & -1 & -1 & -1 \\ 2 & +\infty & +\infty & +\infty \\ 2 & +\infty & +\infty & +\infty \\ 2 & +\infty & +\infty & +\infty \end{pmatrix}^{\leq}$$
$$\begin{pmatrix} F & F & F & F \\ F & F & T & T \\ F & T & F & T \\ F & T & T & F \end{pmatrix}^{\neq}$$



Domain, Order, Normal Form

Domain

- $\mathcal{D}(M) = \{(s_1, ..., s_{n-1}) \in \mathcal{V}^{n-1} \mid \forall i, j \in [0..n-1]$
  $s_j - s_i \leq M_{ij}^{\leq} \ \wedge\ M_{ij}^{\neq} \Rightarrow s_j - s_i \neq 0 \ \wedge\ s_0 = 0\}$

Order

- $M \trianglelefteq M' \iff \forall i, j\ M_{ij} \leq M'_{ij} \ \wedge\ {M'}_{ij}^{\neq} \Rightarrow M_{ij}^{\neq}$

property preserved: $M \trianglelefteq M' \Rightarrow \mathcal{D}(M) \subseteq \mathcal{D}(M')$

Normal form *(for non-empty dDBMs)*

- $\overline{M} = \inf_{\trianglelefteq} \{M' \mid \mathcal{D}(M') = \mathcal{D}(M)\}$

Dense Case

$\downarrow$

$O(n^3)$

Arithmetic Case

emptiness $\swarrow$ $\searrow$ normalization

NP-complete $\quad$ $O(n^5)$

Testing Emptiness

Independence of disequalities

Theorem (Lassez *et al.* 1992)

Let I be a system of linear inequalities, and D be a finite set of linear disequalities. Then the conjunction of I and D is feasible if and only if, for each single disequality $d \in D$, the conjunction of I and $\{d\}$ is feasible.

Emptiness test

- check if no variables given disequal by the dDBM are forced equal by the DBM component

a test runing in $O(n^2)$ on the normal form

Closure (1/5)

DBM component

- independence always hold, apply DBM closure

Constraint deduction rules

- (1) $v_i - v_j \leq c,\ c < 0 \ \Rightarrow\ v_i \neq v_j$
- (2) $v_i = v_j \ \wedge\ v_j \neq v_k \ \Rightarrow\ v_i \neq v_k$
- (3) $v_i \leq v_j \leq v_k \ \wedge\ v_j \neq v_k \ \Rightarrow\ v_i \neq v_k$

- rules (1) and (2) can easily be applied in $O(n^3)$

Closure algorithm

1. Apply the shortest-path closure on $M^{\leq}$;
2. Add implicit disequality constraints (rules (1) and (2)) to $M^{\neq}$

Closure (2/5)

Propagation of rule (3)

- done on a restriction/reduction of the disequal potential graph
  - restriction to zero-weighted arcs
  - reduction on nodes corresponding to equal variables

Closure algorithm

1. Apply the shortest-path closure on $M^{\leq}$;
2. Add implicit disequality constraints (rules (1) and (2)) to $M^{\neq}$;
3. Consider G the disequal potential graph of M where the set of directed edges is restricted to those with null weight ;
4. Compute $\mathcal{SCC}$, the set of strongly connected components of the directed graph of G ;
5. Consider $G^{\bullet}$ the mixed reduced graph of G constructed on $\mathcal{SCC}$;

Closure (3/5)

Propragation of rule (3)

- propagation of an irreflexive and symmetric relation along an order relation

let $G^\bullet = (V^\bullet, A^\bullet, E^\bullet)$

$$\left.\begin{array}{c}(v_1, v_2) \in A^\bullet, (v_2, v_3) \in A^\bullet \\ (v_1, v_2) \in E^\bullet \vee (v_2, v_3) \in E^\bullet\end{array}\right\} \Longrightarrow (v_1, v_3) \in E^\bullet$$

A kind of transitive closure

- Koubeck's algorithm is particulary interesting

worst-case complexity: $O((n^\bullet)^2 n_r^\bullet)$

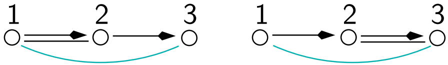
average complexity: $O((n^\bullet)^2 \log n^\bullet)$

Closure (3/5)

Propragation of rule (3)

- propagation of an irreflexive and symmetric relation along an order relation

let $G^{\bullet} = (V^{\bullet}, A^{\bullet}, E^{\bullet})$



A kind of transitive closure

- Koubeck's algorithm is particulary interesting

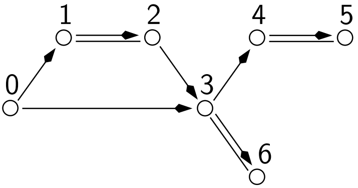
worst-case complexity: $O((n^{\bullet})^2 n_r^{\bullet})$

average complexity: $O((n^{\bullet})^2 \log n^{\bullet})$

Closure (4/5)

Adapting Koubeck's algorithm

- the result of reachable nodes is partitionning into 2 sets
  - set of nodes reachable by some path traversing an arc doubled by an edge
  - set of other reachable nodes

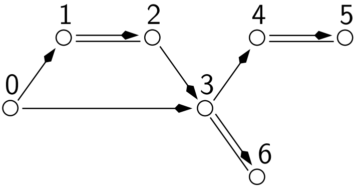


$v \in V^\bullet$	$\Phi(v) = \Phi_1(v), \Phi_2(v)$
6	$(\emptyset\ ,\ \{6\})$
5	$(\emptyset\ ,\ \{5\})$
4	,
3	,
2	,
1	,
0	,

Closure (4/5)

Adapting Koubeck's algorithm

- the result of reachable nodes is partitionning into 2 sets
  - set of nodes reachable by some path traversing an arc doubled by an edge
  - set of other reachable nodes

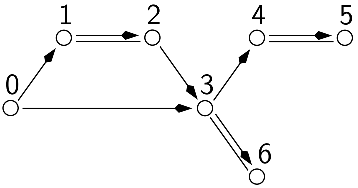


$v \in V^\bullet$	$\Phi(v) = \Phi_1(v), \Phi_2(v)$
6	$(\emptyset\ ,\ \{6\})$
5	$(\emptyset\ ,\ \{5\})$
4	$(\{5\}\ ,\ \{4\})$
3	,
2	,
1	,
0	,

Closure (4/5)

Adapting Koubeck's algorithm

- the result of reachable nodes is partitionning into 2 sets
  - set of nodes reachable by some path traversing an arc doubled by an edge
  - set of other reachable nodes

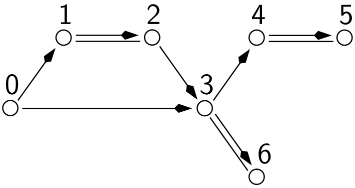


$v \in V^\bullet$	$\Phi(v) = \Phi_1(v), \Phi_2(v)$
6	$(\emptyset \ , \ \{6\})$
5	$(\emptyset \ , \ \{5\})$
4	$(\{5\} \ , \ \{4\})$
3	$(\{5,6\} \ , \ \{3,4\})$
2	,
1	,
0	,

Closure (4/5)

Adapting Koubeck's algorithm

- the result of reachable nodes is partitionning into 2 sets
  - set of nodes reachable by some path traversing an arc doubled by an edge
  - set of other reachable nodes

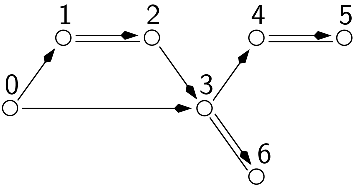


$v \in V^\bullet$	$\Phi(v) = \Phi_1(v), \Phi_2(v)$
6	$(\emptyset\ ,\ \{6\})$
5	$(\emptyset\ ,\ \{5\})$
4	$(\{5\}\ ,\ \{4\})$
3	$(\{5,6\}\ ,\ \{3,4\})$
2	$(\{5,6\}\ ,\ \{2,3,4\})$
1	,
0	,

Closure (4/5)

Adapting Koubeck's algorithm

- the result of reachable nodes is partitionning into 2 sets
  - set of nodes reachable by some path traversing an arc doubled by an edge
  - set of other reachable nodes

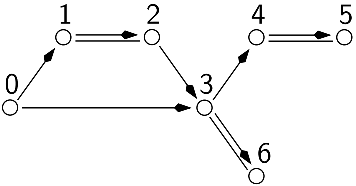


$v \in V^\bullet$	$\Phi(v) = \Phi_1(v), \Phi_2(v)$
6	$(\emptyset \ , \ \{6\})$
5	$(\emptyset \ , \ \{5\})$
4	$(\{5\} \ , \ \{4\})$
3	$(\{5,6\} \ , \ \{3,4\})$
2	$(\{5,6\} \ , \ \{2,3,4\})$
1	$(\{2,3,4,5,6\} \ , \ \{1\})$
0	,

Closure (4/5)

Adapting Koubeck's algorithm

- the result of reachable nodes is partitionning into 2 sets
  - set of nodes reachable by some path traversing an arc doubled by an edge
  - set of other reachable nodes



$v \in V^\bullet$	$\Phi(v) = \Phi_1(v), \Phi_2(v)$
6	$(\emptyset\ ,\ \{6\})$
5	$(\emptyset\ ,\ \{5\})$
4	$(\{5\}\ ,\ \{4\})$
3	$(\{5,6\}\ ,\ \{3,4\})$
2	$(\{5,6\}\ ,\ \{2,3,4\})$
1	$(\{2,3,4,5,6\}\ ,\ \{1\})$
0	$(\{2,3,4,5,6\}\ ,\ \{0,1\})$

Closure (5/5)

Final stage

- report the new disequalities in initial *d*DBM

Closure algorithm

1. Apply the shortest-path closure on $M^{\leq}$;
2. Add implicit disequality constraints (rules (1) and (2)) to $M^{\neq}$;
3. Consider G the disequal potential graph of M where the set of directed edges is restricted to those with null weight ;
4. Compute $\mathcal{SCC}$, the set of strongly connected components of the directed graph of G ;
5. Consider $G^{\bullet}$ the mixed reduced graph of G constructed on $\mathcal{SCC}$;
6. Compute $\mathcal{O}$, a topological order on the directed acyclic graph of $G^{\bullet}$;
7. Apply the disequality propagation algorithm (rule (3)) on $G^{\bullet}$ with respect to $\mathcal{O}$;
8. Add induced disequality constraints into $M^{\neq}$

note: new disequalities are not subject to rule (2)

Complexity $O(n^3)$

Testing Emptiness

NP-completeness

Theorem (Hunt 1980)

The satisfiability problem of a set of potential constraints in presence of disequations is NP-complete

brute force technique

consider for each disequality cases $x - y \leq -1$ and $x - y \geq 1$

- leads to 2^d problems of DBM emptiness

Inert disequalities *(Seater et al 02)*

disequalities wich either eliminates alone all solutions or cannot participate in the absence of solution

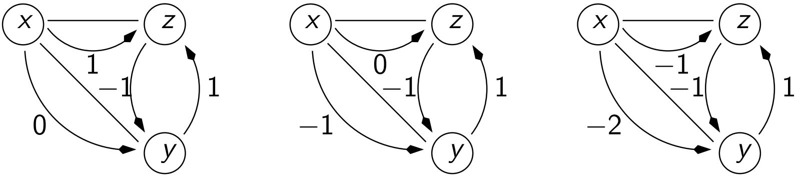
- *e.g.* variables not bounded are inert

Closure (1/2)

Narrowing of the bounds

$(x - y \leq 0 \ \wedge \ x \neq y) \Rightarrow (x - y \leq -1)$

- an iterative process



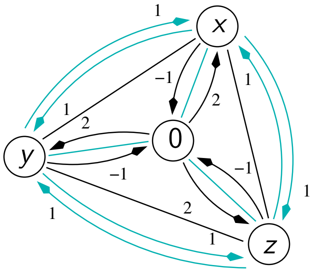
Closure algorithm

repeat
  Apply steps **1** and **2** of dense closure;
  Narrow ;
until *to_narrow* $= \emptyset$;

note: rule (3) taken into account by iteration of narrowing and FW

Complexity $O(n^5) \ldots (O(n^4))$?

Closure (2/2)



Lattice of dDBMs, Analysis Results

Lattice defined

with classical lattice operators + a widenning

Other operators

existential quantification and projection

post-condition of an assignement $(x = y, w \neq 0)\ x \leftarrow x + w$

abstraction of conditions

Implementation

- based on the general fixpoint computation developed by Bertrand Jeannet

only toys examples have been succesfully analyzed

My expectations

Conclusions

- a new numerical abstract domain dealing with both potential constraints and disequalities
  - complexity is $O(n^3)$ when variables take values in a dense set
  - in the arithmetic case, apart the emptiness problems which is exponential, operations are in $O(n^5)$

Future work

- integrate the new domain in an exisiting analyzer to deal with large examples
- implementation in the APRON interface
- extend this work to octagons (expressing $x \neq -y$)
- propose a domain expressing disequalities of the form $x - y \neq c$