DIM0436

31. Interpretação abstrata 2

20141120

Sumário

Reticulado

Definição

Um reticulado A é um *poset* tal que todo par $(a, b) \in A$ tem um supremo é um ínfimo.

Vocabulário particular

- A operação *join* de a e b ($a \wedge b = sup(\{a, b\})$) define o supremo de (a, b)
- A operação *meet* de a e b ($a \vee b = inf(\{a, b\})$) define o ínfimo de (a, b)

Exemplo

- Seja $A \neq \emptyset$, $(\mathcal{P}(A), \subseteq)$ é um reticulado
  - o supremo é a união dos conjuntos
  - o ínfimo é a interseção
- Qualquer conjunto totalmente ordenado define um reticulado

Axiomas dos reticulados

Seja $a, b, c \in (A, \vee, \wedge)$

- $a \vee b = b \vee a$
- $a \wedge b = b \wedge a$
- $a \vee (b \vee c) = (a \vee b) \vee c$
- $a \wedge (b \wedge c) = (a \wedge b) \wedge c$
- $a \vee (a \wedge b) = a$
- $a \wedge (a \vee b) = a$
- $a \vee a = a$
- $a \wedge a = a$

Reticulado completo

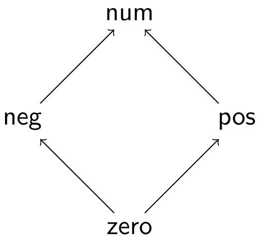
Definição

Um reticulado $(A, \vee, \wedge)$ é **completo** se $\forall B \subseteq A, \bigvee B$ e $\bigwedge B$ existem.

Teorema (Knaster-Tarski)

- *Seja* $(A, \vee, \wedge)$ *um reticulado completo é* $f : A \rightarrow A$ *uma função crescente.*
- *O conjunto de pontos fixos de* f *em* A *não é vazio e é um reticulado completo.*
- f *tem um menor é um maior ponto fixo em* A

Reticulado dos sinais



Relação de corretude

A relação R

$$R : V \times L \to \{\text{true}, \text{false}\}$$

- $R(v, l)$: o valor v é descrito pela propriedade l

Corretude

R é preservado durante computação

$$R(v_1, l_1) \wedge p \vdash v_1 \rightsquigarrow v_2 \wedge p \vdash l_1 \rhd l_2 \Rightarrow R(v_2, l_2)$$

Relação de corretude admissível

Seja $L = \{L, \sqsubseteq, \sqcap, \sqcup, \bot, \top\}$ um reticulado completo
Devemos ter:

1. $R(v, l_1) \wedge l_1 \sqsubseteq l_2 \Rightarrow R(v, l_2)$
2. $\forall l \in L' \sqsubseteq L : R(v, l) \Rightarrow vR(v, L')$

A condição 2 tem 2 consequências:

1. $R(v, \top)$
2. $R(v, l_1) \wedge R(v, l_2) \Rightarrow R(v, l_1 \sqcap l_2)$

Funções de representação

Ao invés de usar uma relação de corretude, pode-se usar uma função de representação

$$\beta : V \to L$$

- Intuitivamente, associa a um valor a **melhor** propriedade que o descreve.

$$\beta(v_1) \sqsubseteq \wedge p \vdash v_1 \rightsquigarrow v_2 \wedge p \vdash l_1 \triangleright l_2 \Rightarrow \beta(v_2) \sqsubseteq l_2$$

Equivalência dos critérios

Dada uma função de representação β, definimos uma relação de corretude R_β

$$R_\beta(v, l) \iff \beta(v) \sqsubseteq l$$

Dada uma relação de corretude R, definimos uma função de representação β_R

$$\beta_R(v)\{l | R(v, l)\}$$

Lema

1. *Dado $\beta : V \to L$, a relação \$$R_\beta$: $V \times$ $L \to$ {true, false } satisfaz condições 1 e 2 e \$$\beta_{R_\beta} = \beta$.*
2. *Dado $R : V \times L \to \{true, false\}$ satisfazendo 1 e 2, β_R é bem definido e $R_{\beta_R} = R$*

Uma generalização modesta

Programa

Um programa representa como transformar um valor v_1 em um valor v_2

$$p \vdash v_1 \rightsquigarrow v_2, v_1 \in V_1, v_2 \in V_2$$

Análise do programa

$$p \vdash l_1 \rhd l_2, l_1 \in L_1, l_2 \in L_2$$

- $\mathrm{l}_2 = \mathrm{f}_\mathrm{p}(\mathrm{l}_1)$

Corretude

$$R_1(v_1, l_1) \wedge p \vdash v_1 \rightsquigarrow v_2 \Rightarrow R_2(v_2, f_p(l_1))$$

Aproximação de pontos fixos

- A transformação do programa abstrato $p \vdash l_1 \rhd l_2$ é normalmente definido por uma função $f : L \to L$, monótona, tal que $f(l_1) = l_2$
- Para programas recursivos ou iterativos, deseja-se obter o **menor ponto fixo** (lfp) de f, como resultado de um processo **finito** iterativo
- Porem, a sequência iterative $(f^n(\bot))_n$ pode:
  - não estabilizar
  - não ter o seu supremo igual a $lfp(f)$

Pontos fixos

Seja $f : L \to L$ uma função monótona num reticulado completo $L = (L, \sqsubseteq, \sqcup, \sqcap, \perp, \top)$

- Um **ponto fixo** de f é um elemento $l \in L$, tal que $f(l) = l$
$$\text{Fix(f)} = \{l | f(l) = l\}$$
- f e **redutiva** em l sse $f(l) \sqsubseteq l$
$$\text{Red(f)} = \{l | f(l) \sqsubseteq l\}$$
- f e **extensiva** em l sse $f(l) \sqsupseteq l$
$$\text{Ext(f)} = \{l | f(l) \sqsupseteq l\}$$

$$lfp(f) = Fix(f) = Red(f) \in Fix(f) \subseteq Red(f)$$

$$f^n(\perp) \sqsubseteq \bigsqcup_n f^n(\perp) \sqsubseteq lfp(f) \sqsubseteq gfp(f) \sqsubseteq_n f^n(\top) \sqsubseteq f^n(\top)$$

Operador de *widening*

- Como $(f^n(\perp))_n$ pode não estabilizar, devemos considerar uma outra forma de aproximar $lfp(f)$
- A ideia é introduzir uma nova sequência $(f^n_\nabla)_n$ que vai estabilizar-se com um valor que é uma (sobre) aproximação correta do lfp.
- O operador ∇ é chamado de operador de *widening*: a precisão da aproximação e o custo de cálculo depende da escolha desse operador

Operador de supremo

Definição

Um operador $\check{\sqcup} : L \times L \to L$, $L = (L, \sqsubseteq)$ um reticulado completo é um operador de supremo se

$$l_1 \sqsubseteq (l_1 \check{\sqcup} l_2) \sqsupseteq l_2$$

Se $(l_n)_n$ for uma sequência e $\check{\sqcup}$ um operador de supremo, então $(l_n^{\check{\sqcup}})_n$ é uma cadeia crescente, e $l_n^{\check{\sqcup}} \sqsupseteq \bigsqcup\{l_0, l_1, \ldots, l_n\}$ para todo n.

Exemplo

- Considere de novo o reticulado

$$\overline{L} = \{\bot\} \cup \{[l,u] | l \in \mathbb{Z} \cup \{-\infty\} \wedge u \in \mathbb{Z} \cup \{+\infty\} \wedge l \leq u\}$$

- Considere o operador $\check{\sqcup}$ definido por

$$i_1 \check{\sqcup} i_2 = \begin{cases} i_1 \sqcup i_2 \text{ se } i_1 \sqsubseteq \mathcal{I} \vee i_2 \sqsubseteq i_1 \\ [-\infty, +\infty] \text{ senão} \end{cases}$$

Exercício

Seja $\mathcal{I} = [0,2]$

1. $[1,2]\check{\sqcup}[2,3] = ?$
2. $[2,3]\check{\sqcup}[1,2] = ?$

Operador de *widening*

Definição

Um operador $\nabla : L \times L \to L$ é um operador de widening sse:

- é um operador de supremo;
- para toda cadeia crescente $(l_n)_n$, a cadeia crescente $(l_n^\nabla)_n$ finalmente estabiliza-se.

Uso

$$\mathrm{f}_\nabla^n = \begin{cases} \perp & \text{se } n = 0 \\ f_\nabla^{n-1} & \text{se } f(f_\nabla^{n-1}) \sqsubseteq f_\nabla^{n-1} \\ f_\nabla^{n-1} \nabla f(f_\nabla^{n-1}) & \text{senão} \end{cases}$$

Exemplo

Consideramos de novo os intervalos. Seja K um conjunto finito de inteiros. A ideia é que deveríamos ter

$$[z_1, z_2] \nabla [z_3, z_4] \approx [LB(z_1, z_3), UB(z_2, z_4]$$

$$\text{LB(x, y)} = \begin{cases} x \text{ se } x \leq y \\ k \text{ se } y < x \wedge k = max\{k \in K | k \leq y\} \\ -\infty \text{ se } y < x \wedge \forall k \in K : y < k \end{cases}$$

$$\text{UB(x, y)} = \begin{cases} x \text{ se } y \leq x \\ k \text{ se } x < y \wedge k = min\{k \in K | y \leq k\} \\ +\infty \text{ se } x < y \wedge \forall k \in K : k < y \end{cases}$$

$$\text{i}_1 \nabla i_2 = \begin{cases} \perp \text{ se } i_1 = i_2 = \perp \\ [LB(inf(i_1), inf(i_2)), UB(sup(i_1), sup(i_2))] \text{ senão} \end{cases}$$

- Suponha que
  - $K = \{3, 5\}$
  - $(\text{int}_\text{n})_\text{n}$ = [0,1], [0,2], [0,3], [0,4], [0,5], [0,6], [0, 7],...

Problema do laço

```
i := 1;
while i <= 100 do i := i + 1;
```

- Na semântica concreta `i` vale 101 após a execução do laço, e vale entre 1 e 100 dentro do laço
- Matematicamente, o menor ponto fixo
$$\mathrm{lfp}(F) = \mathrm{lfp}_{emptyset}(F) = \{i \in \mathbb{Z} | 1 \leq i \leq 100\}$$
do operador
$$F \in L \mapsto L = \lambda X.(\{1\} \cup \{i+1 | i \in X\}) \cap \{i \in \mathbb{Z} | i \leq 100\}$$

no reticulado completo $L = wp(\mathbb{Z})(\subseteq, \emptyset, \cap, \cup)$

- Uma sobre-aproximação correta é o invariante de laço $A = \{i \in \mathbb{Z} | i \geq 0\}$

Operador de *narrowing*

Definição

Um operador $\triangle : L \times L \to L$ é um operador de widening sse:

- é um operador de supremo;
- para toda cadeia crescente $(l_n)_n$, a cadeia crescente $(l_n^\nabla)_n$ finalmente estabiliza-se.

Conexão de Galois

Definição

Sejam $(P, \leq)$ e $(Q, \sqsubseteq)$ conjuntos parcialmente ordenados
Um para (α, γ) de mapeamentos:

- $\alpha : P \mapsto Q$
- $\gamma : Q \mapsto P$

é uma **conexão de Galois** sse

$$\forall x \in P, \forall y \in Q : \alpha(x) \sqsubset y \iff x < \gamma(y)$$

notado

$$(P, \leq) \overset{\alpha}{\underset{\gamma}{\rightleftarrows}} (Q, \sqsubseteq)$$

Abstração e concretização

Abstração

Uma função de **abstração** α é um mapeamento de um objeto concreto o para uma aproximação do domínio de interpretação $\alpha(0)$.

Concretização

Uma função de **concretização** γ é um mapeamento de um objeto abstrato $\overline{o}$ para uma um objeto concreto $\gamma(\overline{o})$.

Exemplos

Exemplo

Sejam P e Q dois conjuntos e $b : P \to Q$ uma função bijetora, com inversa b^{-1}.

$$(P, =) \underset{b^{-1}}{\overset{b}{\rightleftarrows}} (Q, =)$$

- $(P, =)$ é P ordenado por igualdade

Exemplo

Sejam C, A conjuntos e $f : C \mapsto A$. Defina

- $\alpha(X) \hat{=} \{f(x) | x \in X\}$
- $\gamma(X) \hat{=} \{x | f(x) \in Y\}$

então

$$(\wp(C), \subseteq) \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} (\wp(A), \subseteq)$$

Conexão de Galois: propriedade

Teorema

Numa conexão de Galois

$$(P, \leq) \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} (Q, \sqsubseteq)$$

uma adjunção determina a outra

- $\alpha(x) = \sqcap\{y | x \geq \gamma y\}$
- $\gamma(x) = \bigvee\{x | \alpha(x) \sqsubseteq y\}$

Conexão de Galois: formalmente

$$(D, \subseteq) \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} (\bar{D}, \sqsubseteq)$$

sse

1. α **é monótono**

$$\forall x, y \in D : x \subseteq y \Rightarrow \alpha(x) \sqsubseteq \alpha(y)$$

2. γ **é monótono**

$$\forall \bar{x}, \bar{y} \in \bar{D} : \bar{x} \sqsubseteq \bar{y} \Rightarrow \gamma(\bar{x}) \subseteq \gamma(\bar{y})$$

3. $1_P \leq \gamma \circ \alpha$

$$\forall x \in D : x \subseteq \gamma(\alpha(x))$$

4. $\alpha \circ \gamma \sqsubseteq 1_Q$

$$\forall \bar{y} \in \bar{D} : \alpha(\gamma(\bar{y})) \sqsubseteq \bar{y}$$

sse

$$\forall x \in D, \bar{y} \in \bar{D} : \alpha(x) \sqsubseteq y \iff x \subseteq \gamma(y)$$

Exemplo (intervalos)

$\wp(\mathbb{Z})$ é aproximada usando o reticulado dos intervalos:

$$\bar{L} = \{\bot\} \cup \{[l, u] | l \in \mathbb{Z} \cup \{-\infty\} \wedge u \in \mathbb{Z} \cup \{+\infty\} \wedge l \leq u\}$$

ordenado por $\sqsubseteq$:

- $\bot \sqsubseteq [l, u] \stackrel{\text{def}}{=} \text{true}$
- $[l_0, u_0] \subseteq [l_1, u_1] \stackrel{\text{def}}{=} l_1 \leq l_0 \leq u_0 \leq u_1$

Essa aproximação é formalizada pela conexão de Galois:

$$\begin{array}{ll} \gamma(\bot) = \emptyset & \alpha(\emptyset) = \bot \\ \gamma([l, u]) = \{x \in \mathbb{Z} | l \leq x \leq u\} & \alpha(X) = [min\ X, max\ X] \end{array}$$

O conjunto $\{1, 2, 5\} \in \wp(\mathbb{Z})$ é corretamente aproximado por $[1, 5] \in \bar{L}$.

Exercício

Dada a conexão de Galois do slide anterior com

$$\begin{array}{ll} \gamma(\bot) = \emptyset & \alpha(\emptyset) = \bot \\ \gamma([l, u]) = \{x \in \mathbb{Z} | l \leq x \leq u\} & \alpha(X) = [minX, maxX] \end{array}$$

Qual é o resultado de:

- $\gamma([0, 3])$
- $\gamma([0, +\infty)$
- $\alpha(\{0, 1, 3\})$
- $\alpha(\{2 * z | z > 0\})$

Abstração: reticulado dos sinais

Definição

- $\gamma : \mathit{Sign} \to \mathcal{P}(\mathbb{Z})\backslash\{\emptyset\}$

$$
\begin{array}{lcl}
\gamma(\text{zero}) & = & \{0\} \\
\gamma(\text{pos}) & = & \{x \mid x > 0\} \\
\gamma(\text{neg}) & = & \{x \mid x < 0\} \\
\gamma(\text{num}) & = & \mathbb{Z}
\end{array}
$$

Concretização: reticulado dos sinais

Definição

- $\alpha : \mathcal{P}(\mathbb{Z})\backslash\{\emptyset\} \to \mathit{Sign}$

$$
\begin{array}{lcll}
\alpha(\{0\}) & = & \textsf{zero} & \\
\alpha(X) & = & \textsf{pos} & \textsf{se } \forall x \in X > 0 \\
\alpha(X) & = & \textsf{neg} & \textsf{se } \forall x \in X < 0 \\
\alpha(X) & = & \textsf{num} & \textsf{senão}
\end{array}
$$

Exemplo

$$
\begin{array}{rcl}
\alpha(\{2,3,1\}) & = & \textsf{pos} \\
\alpha(\{-1,-2,-3\}) & = & \textsf{neg} \\
\alpha(\{1,2,-4\}) & = & \textsf{num}
\end{array}
$$

Conexão de Galois a partir de funções de representação

Uma função de representação $\beta : V \to L$ gera uma conexão de Galois

$$(\wp(V), \alpha, \gamma, L)$$

com

1. $\alpha(V') = \sqcup\{\beta(v) | v \in V'\}$
2. $\gamma(l) = \{v \in V | \beta(v) \sqsubseteq l\}$
3. com $V' \subseteq V$, $l \in L$

Isso define uma adjunção

$$\begin{aligned}
\alpha(V') \subseteq l \quad &\Leftrightarrow \quad \sqcup\{\beta(v) | v \in V'\} \subseteq l \\
&\Leftrightarrow \quad \forall v \in V' : \beta(v) \subseteq l \\
&\Leftrightarrow \quad V' \subseteq \gamma(l)
\end{aligned}$$

Conexão de Galois a partir de função de extração

Uma **função de extração** $\nu : V \rightarrow D$ e um mapeamento dos valores de V para a suas melhores descrições D.
Isso gera uma função de representação $\beta_\nu : V \rightarrow \wp(D)$

$$\beta_\nu(v) = \{\nu(v)\}$$

A conexão de Galois associada é

$$(\wp(V), \alpha_\nu, \gamma_\nu, \wp(D)$$

1. $\alpha_\nu(V') = \cup\{\beta_\nu(v) | v \in V'\} = \{\nu(v) | v \in V'\}$
2. $\gamma_\nu(D') = \{v \in V | \beta_\nu(v) \subseteq D'\} = \{v | \nu(v) \subseteq Dkj'\}$

Exemplo

Sejam dois reticulados completos $(\wp(\mathbb{Z}, \subseteq)$ e $(\wp(Sign), \subseteq)$ com $Sign = \{-, 0, +\}$
A função de extração $sign : \mathbb{Z} \to Sign$ define o sinal dos inteiros

$$sign(z) = \begin{cases} - \text{ se } z < 0 \\ 0 \text{ se } z = 0 \\ + \text{ se } z > 0 \end{cases}$$

Isso da uma conexão de Galois

$$(\wp(\mathbb{Z}), \alpha_s, \gamma_s, \wp(Sign))$$

com

- $\alpha_s(Z) = \{sign(z) | z \in Z\}$
- $\gamma_s(S) = \{z \in \mathbb{Z} | sign(z) \in S\}$

$Z \subseteq \mathbb{Z}, S \subseteq Sign$

Ordem e operações abstratas

- $\alpha \sqsubseteq y$ é definida como $\gamma(x) \subseteq \gamma(y)$
- $x \sqcup y = \alpha(\gamma(x) \cup \gamma(y))$

Outras propriedades

Teorema

Temos $(P, \leq) \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} (Q, \sqsubseteq)$ *sse* $(Q, \sqsupset) \underset{\alpha}{\overset{\gamma}{\rightleftarrows}} (Q, >)$

O ***dual*** *duma conexão de Galois* (α, γ) *é* (γ, α)

Teorema

A composição de conexões de Galois é também uma conexão de Galois.

- $(P, \leq) \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} (Q, \sqsubseteq)$
- $(Q, \sqsubseteq) \underset{\delta}{\overset{\beta}{\rightleftarrows}} (R, \preceq)$

então $(P, \leq) \underset{\gamma \circ \delta}{\overset{\beta \circ \alpha}{\rightleftarrows}} (R, \preceq)$

Galois surjeção / injeção

Teorema

Se $(P, \leq) \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} (Q, \sqsubseteq)$ *então:*

$$\alpha \text{ é sobrejetora } \iff \gamma \text{ é bijetora } \iff \alpha \circ \gamma = 1_Q$$

Teorema

Por dualidade, se $(P, \leq) \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} (Q, \sqsubseteq)$ *então:*

$$\gamma \text{ é sobrejetora } \iff \alpha \text{ é bijetora } \iff \gamma \circ \alpha = 1_P$$

Inserção de Galois (*to be continued ...*)

$$(D, \subseteq) \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} (\bar{D}, \sqsubseteq)$$

sse

1. **α é monótono**

$$\forall x, y \in D : x \subseteq y \Rightarrow \alpha(x) \sqsubseteq \alpha(y)$$

2. **γ é monótono**

$$\forall \bar{x}, \bar{y} \in \bar{D} : \bar{x} \sqsubseteq \bar{y} \Rightarrow \gamma(\bar{x}) \subseteq \gamma(\bar{y})$$

3. $1_P \leq \gamma \circ \alpha$

$$\forall x \in D : x \subseteq \gamma(\alpha(x))$$

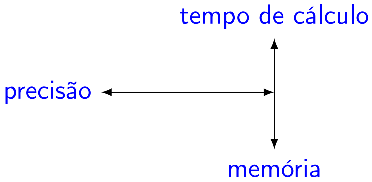
4. $\alpha \circ \gamma = 1_Q$

$$\forall \bar{y} \in \bar{D} : \alpha(\gamma(\bar{y})) = \bar{y}$$

Como escolher o seu domínio abstrato ?

Na prática, escolher o domínio abstrato é **fundamental**

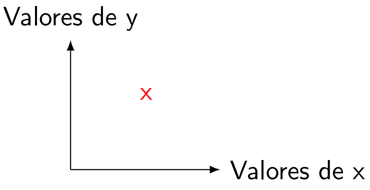
- deve ser suficientemente **preciso**
- em particular, deve permitir expressar a propriedade desejada
- deve ser calculável com o custo **tempo/memória razoavel**
  - i.e. horas de cálculo, Gb de RAM para casos reais



- **Domínio não relacional**: nenhuma relação entre elemento é conservada. Pouco preciso mas pouco custoso
- **Domínio relacional**: relações entre elenentos do domínio. Mais préciso mas custa

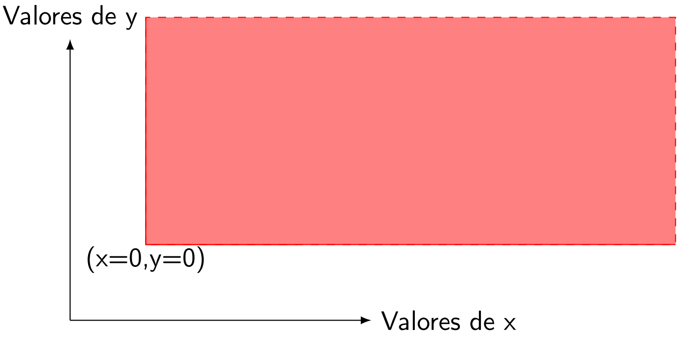
Domínio das constantes

- $x = z$ $(z \in \mathbb{Z})$
- domínio não relacional
- se o valor exato não é conhecido, perde a informação inteira



Domínio dos sinais

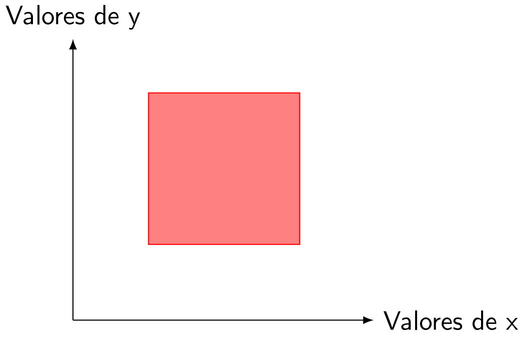
- $x op 0$, avec op $\in \{\geq, >, \leq, <, =, \neq\}$
- domínio não relacional
- conservação dos valores possíveis



Domínio dos intervalos

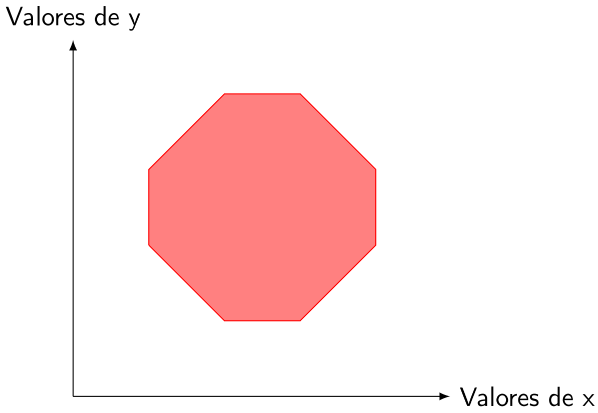
*

- $x \in [i_0, i_1]$
- domínio não relacional
- conservação de um intervalo agrupando todos os valores possíveis



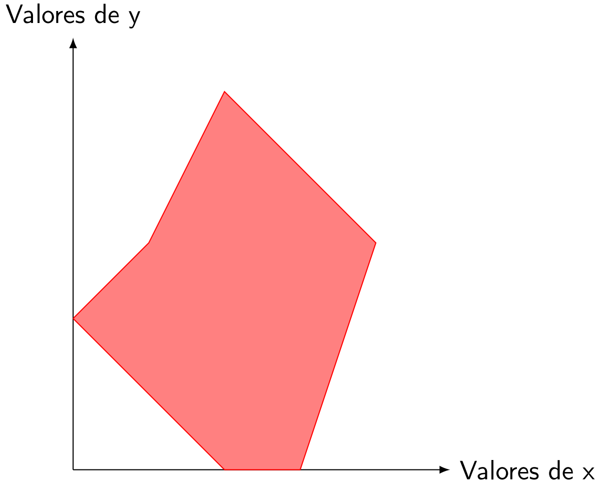
Domínio dos octogonos

- $\pm x \pm y \leq c$
- domínio relacional
- conservação de relações lineares simples entre elementos



Domínio dos poliedros

- $kx + ly \leq c$
- domínio relacional
- relações lineares complexas entre elementos



Análise de valor

Descrição

- Análise por interpretação abstrata de programas **sequenciais**
- Cálculo dos domínios de variação das variáveis do programa
- Inferência da ausência de erros de execução

Resumo

1. Formalização
2. Pontos fixos
3. Conexões de Galois
4. Domínios abstratos
5. Análise de valor de Frama-C
6. Apresentações

Referências

Patrick Cousot and Radhia Cousot, *Abstract interpretation: A unified lattice model for static analysis of programs by construction or approximation of fixpoints*, Proceedings of the 4th ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages (New York, NY, USA), POPL '77, ACM, 1977, pp. 238–252.

Flemming Nielson, Hanne Riis Nielson, and Chris Hankin, *Principles of program analysis (2. corr. print)*, Springer, 2005.

Perguntas ?

http://dimap.ufrn.br/~richard/dim0436

Dicas

- http://research.microsoft.com/en-us/um/people/simonpj/papers/giving-a-talk/giving-a-talk.htm
- http://matt.might.net/articles/academic-presentation-tips/
- http://www2.cs.uregina.ca/~pwlfong/CS499/reading-paper.pdf (Partes da compreensão e da síntese)
- Você apresenta para seus companheiros/colegas e não só para o professor!
- Duração: 15 min ($\approx$ 7 slides)