Abstraction III

Property Abstraction

(based on Patrick Cousout's 2005 course "Abstract Interpretation")

Albert-Ludwigs-Universität Freiburg

Manuel Geffken

2014-07-01

Correspondance between concrete and abstract properties

Given a closure operator ρ on a poset $\langle L, \sqsubseteq \rangle$ (typically L is $\mathcal{P}(\Sigma)$), Morgado's theorem [1] states that for all $P, P' \in L$:

$$\rho(P) \sqsubseteq \rho(P') \Leftrightarrow P \sqsubseteq \rho(P')$$

that is, by definition of Galois connections ($1_L \stackrel{\text{def}}{=} \lambda x \in L.x$):

$$\langle L, \sqsubseteq \rangle \underset{\rho}{\overset{1_L}{\leftrightarrows}} \langle \rho(L), \sqsubseteq \rangle$$

[1] Proof of Morgado's theorem:
"$\Leftarrow$": $P \sqsubseteq \rho(P') \stackrel{\textit{increasing}}{\Longrightarrow} \rho(P) \sqsubseteq \rho(\rho(P')) \stackrel{\textit{idempotent}}{\Longrightarrow} \rho(P) \sqsubseteq \rho(P')$
"$\Rightarrow$": $\rho(P) \sqsubseteq \rho(P') \stackrel{\textit{extensive}}{\Longrightarrow} P \sqsubseteq \rho(P) \sqsubseteq \rho(P') \stackrel{\textit{trasitivity}}{\Longrightarrow} P \sqsubseteq \rho(P) \sqsubseteq \rho(P')$

Correspondance between concrete and representations of abstract properties

Proof.

We must prove
$\forall x \in L : \forall y \in \rho(L) : (\rho(x) \sqsubseteq y \iff (x \sqsubseteq 1_L(y))$. We have $y \in \rho(L)$ iff $\exists z \in L : \rho(z) = y$ so that this condition is equivalent to $\forall x, z \in (\rho(x) \sqsubseteq \rho(z)) \iff (x \sqsubseteq \rho(z))$ which directly follows from Morgado's theorem. Moreover ρ is surjective on $\rho(L)$. □

- Let $\langle A, \leq \rangle$ be an order-isomorphic representation of the abstract domain $\langle \rho(L), \sqsubseteq \rangle$. We have

$$\langle \rho(L), \sqsubseteq \rangle \underset{\epsilon}{\overset{\epsilon^{-1}}{\leftrightarrows}} \langle A, \leq \rangle$$

where ϵ^{-1} is the inverse of the bijection $\epsilon \in \rho(L) \to A$

By composition, we get:

$$\langle L, \sqsubseteq \rangle \underset{\epsilon \circ \rho}{\overset{1_L \circ \epsilon^{-1}}{\leftrightarrows}} \langle A, \leq \rangle$$

$$\begin{tikzcd}
\langle L, \sqsubseteq \rangle \arrow[r, shift right, "\rho"'] \arrow[rr, bend right, "\epsilon \circ \rho = \alpha"'] & \langle \rho(L), \sqsubseteq \rangle \arrow[l, shift right, "1_L"'] \arrow[r, shift right, "\epsilon"'] & \langle A, \leq \rangle \arrow[l, shift right, "\epsilon^{-1}"'] \arrow[ll, bend right, "1_L \circ \epsilon^{-1} = \gamma"']
\end{tikzcd}$$

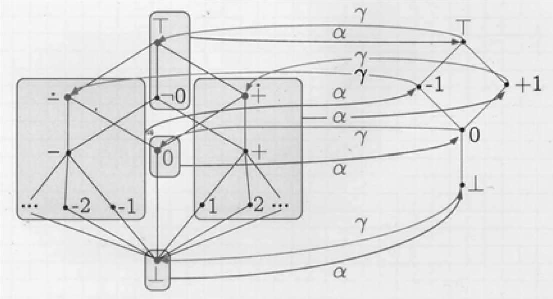
Specification of an abstract domain by a Galois surjection

- Inversely, we can consider a Galois surjection

$$\langle L, \sqsubseteq \rangle \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \langle A, \leq \rangle$$

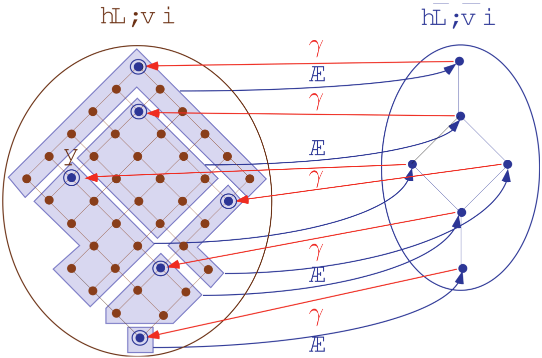
- Then $\rho = \gamma \circ \alpha$ is a closure operator and $\langle A, \leq \rangle$ is order-isomorphic to $\langle \rho(L), \sqsubseteq \rangle$
- We have an order-isomorphic representation of the abstract domain $\langle \rho(L), \sqsubseteq \rangle$, which is a Moore family.

Specification of an abstract domain by a Galois surjection, example



Because α is surjective, γ is injective and order is preserved, each element in the Moore family $\{\bot, 0, \dot{-}, \dot{+}, \top\}$ has a unique isomorphic representation $\{\bot, 0, -1, +1, \top\}$. This would not be the case when α is not surjective.

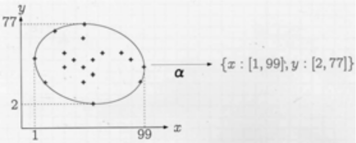
Galois Connection $\langle L, \sqsubseteq\rangle \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \langle \overline{L}, \overline{\sqsubseteq}\rangle$



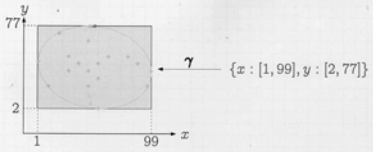
- ◉: Moore family of best approximations;
- ▭: concrete values with the same abstraction.

A graphical illustration of the specification of an abstraction by a Galois surjection

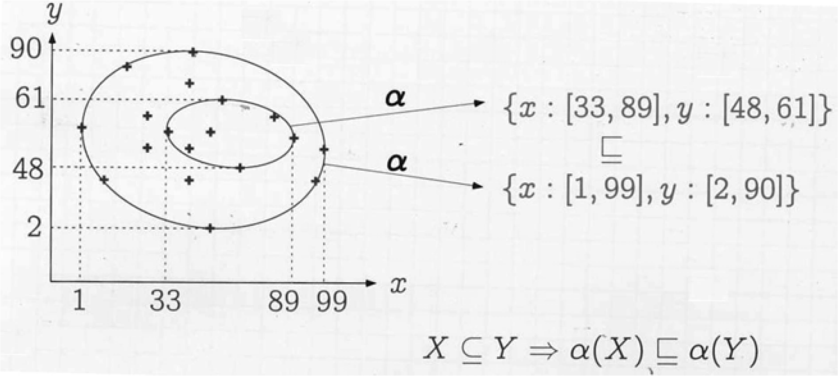
- Abstraction of a set of points in $\mathbb{R}^2$ by an interval:



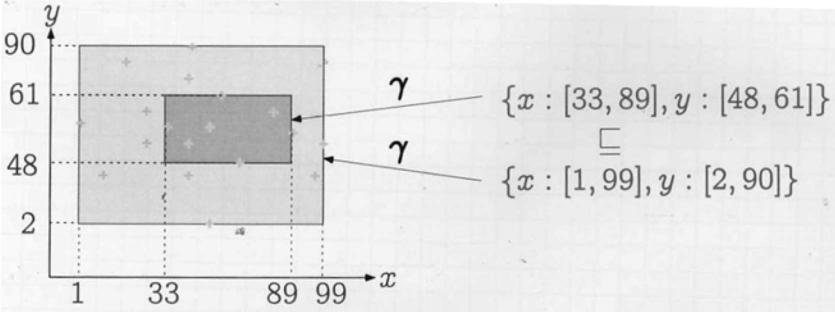
- Concretization:



- The abstraction α is monotone:

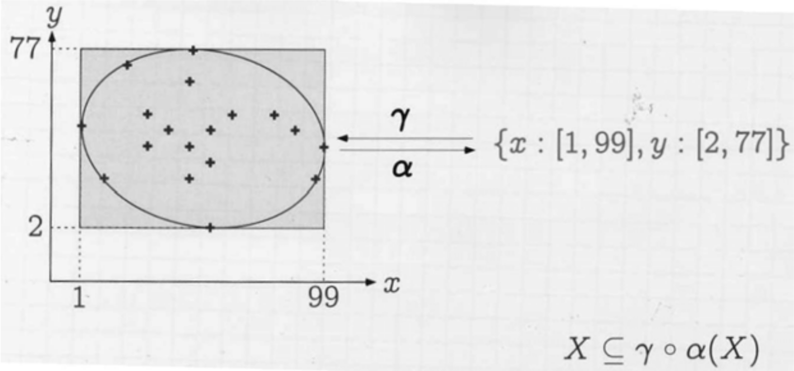


- The concretization γ is monotone:

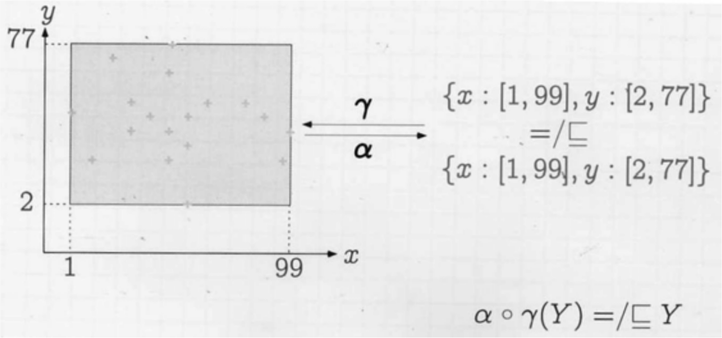


$$X \sqsubseteq Y \Rightarrow \gamma(X) \subseteq \gamma(Y)$$

- $\gamma \circ \alpha$ is extensive (indeed an upper closure operator):



- The composition $\alpha \circ \gamma$ is:
  - The identity for Galois surjections
  - Reductive (ideed a lower-closure operator) for Galois connections [2]



[2] providing the least abstract properties with similar expressive power that is same concretization.

- The intuition of $\sqsubseteq$ is that $\overline{P} \sqsubseteq \overline{P}'$ implies $\gamma(\overline{P}) \subseteq \gamma(\overline{P}')$ so that $\overline{P}$ is more precise than $\overline{P}'$ when expressed in the concrete.
- So $\alpha \circ \gamma(\overline{P}) \sqsubseteq \overline{P}$ means that concretization can loose no information, since if the concrete property P is overapproximated by $\overline{P}$ then

$$
\begin{aligned}
& P \subseteq \gamma(\overline{P}) \\
\iff & P \subseteq \gamma(\alpha \circ \gamma(\overline{P}))
\end{aligned}
$$

so that using $\overline{P}$ or $\alpha \circ \gamma(\overline{P})$ is exactly the same in the concrete, as far as precision is concerned.

Why are abstract domains complete lattices in the presence of best abstractions?

- The abstractions start from the complete lattice of concrete properties $\langle \mathcal{P}(\Sigma), \subseteq, \emptyset, \Sigma, \cup, \cap, \neg \rangle$ where objects in Σ represent program computations and the elements of $\mathcal{P}(\Sigma)$ represent properties of these program computations
- We have defined abstract domains with best approximations in three equivalent different ways
  - As a Moore family;
  - As a closure operator (which fixpoints form the abstract domain);
  - As the image of the concrete domain by a Galois surjection.

- In all cases, it follows that the abstract domain is a complete lattice:
  - A Moore family of a complete lattice is a complete lattice;
  - The image of a complete lattice by an upper closure operator is a complete lattice;
  - The image of a complete lattice by the surjective abstraction of a Galois connection is a complete lattice.
- In general this property does <u>not</u> hold in absence of a best abstraction or if arbitrary points are added to the abstract domain as shown next.

Relaxing the condition on the uniqueness of the representation of abstract properties: Galois connections

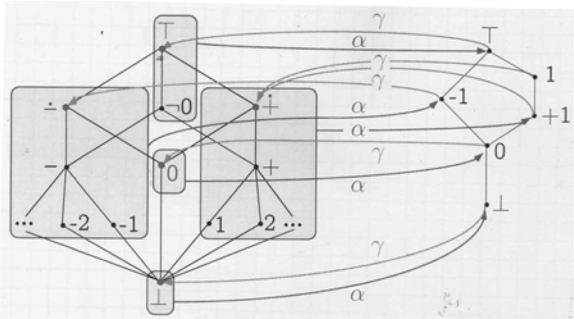
- Assume the correspondence between concrete and abstract properties is a non-surjective (α is not surjective) Galois connection:

$$\langle L, \sqsubseteq \rangle \underset{\alpha}{\overset{\gamma}{\underset{\longrightarrow\!\!\!\!\rightarrow}{\longleftarrow}}} \langle A, \leq \rangle$$

- γ is not injective, which means that at least two different abstract properties $\overline{P}_1$ and $\overline{P}_2$ have exactly the same concretization:

$$\overline{P}_1 \neq \overline{P}_2 \wedge \gamma(\overline{P}_1) = \gamma(\overline{P}_2)$$

Example of non-surjective Galois connection based abstraction



Here "1" and "+1" are two different encodings of the same concrete property $\dot{+}$ (i.e; positive or zero).

Reduction

- With non-surjective Galois connections $\langle L, \sqsubseteq \rangle \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \langle A, \leq \rangle$ there are at least two different representations in the abstract of at least one concrete property
- This may happen when abstract computer representations of the same concrete property are not unique (e.g. sets represented by ordered trees)
- Reduction is always mathematically possible, by considering $\langle L, \sqsubseteq \rangle \underset{\alpha_\equiv}{\overset{\gamma_\equiv}{\leftrightarrows}} \langle A_\equiv, \leq_\equiv \rangle$ where $\overline{P} \equiv \overline{P}' \Leftrightarrow \gamma(\overline{P}) = \gamma(\overline{P}')$, $\alpha_\equiv(P) = [\alpha(P)]_\equiv$, $\gamma([\overline{P}]_\equiv) = \gamma(\overline{P})$ and $[\overline{P}]_\equiv \leq_\equiv [\overline{P}']_\equiv \Leftrightarrow \overline{P} \leq \overline{P}'$

- Example:
  - Abstract properties are intervals $[a, b]$ meaning $\gamma([a, b]) \stackrel{def}{=} \{x \mid \text{minint} \leq a \leq x \leq b \leq \text{maxint}\}$
  - The empty set is represented by any $[a, b]$ with $b < a$. This This can be left as is or normalized as e.g. $[\text{maxint}, \text{minint}]$
  - The supremum is represented by any $[a, b]$ with $a \leq \text{minint}$ and $\text{maxint} \leq b$. This can be left as is or better normalized as e.g. $[\text{minint}, \text{maxint}]$
  - Sometimes it is better to have a "normal form", but this reduction may also be sometimes algorithmically very expensive

The interval complete lattice with "normal form" for the empty set and the supremum