

A casual introduction to Abstract Interpretation

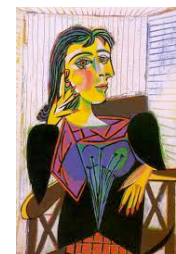
NYU, 29–30 March 2012

Patrick Cousot

cs.nyu.edu/~pcousot

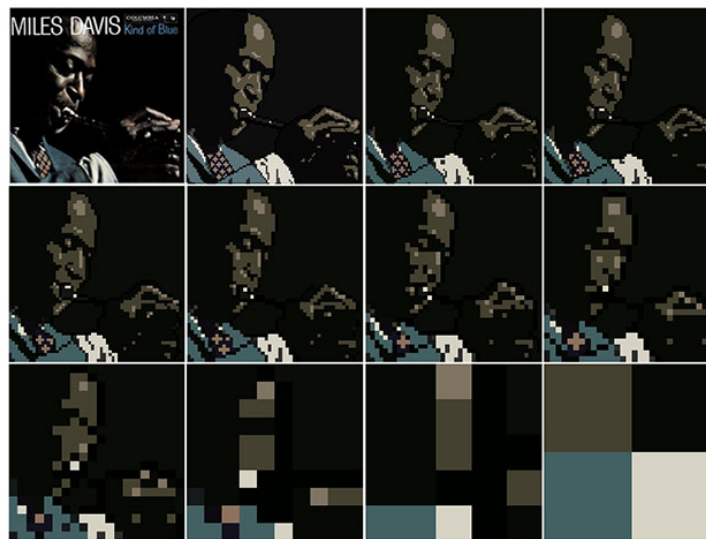
di.ens.fr/~cousot

Abstractions of Dora Maar by Picasso



Examples of Abstractions

Pixelation of a photo by Jay Maisel

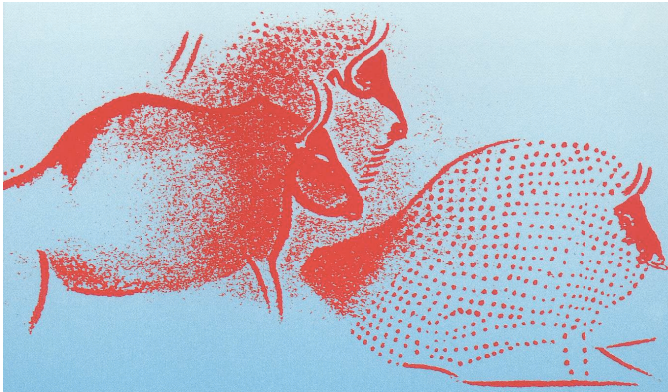


www.petapixel.com/2011/06/23/how-much-pixelation-is-needed-before-a-photo-becomes-transformed/

Image credit: Photograph by Jay Maisel

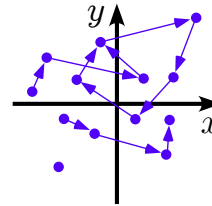
An old idea...

20 000 years old picture in a spanish cave:

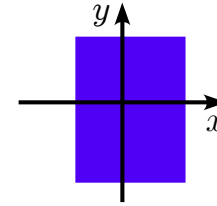


The concrete is not always well-known!

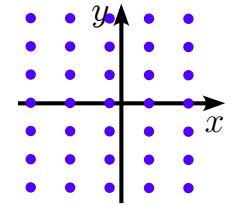
Numerical abstractions in Astrée



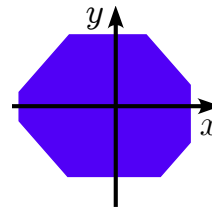
Collecting semantics:
partial traces



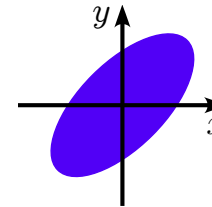
Intervals:
 $x \in [a, b]$



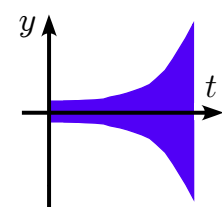
Simple congruences:
 $x \equiv a[b]$



Octagons:
 $\pm x \pm y \leq a$

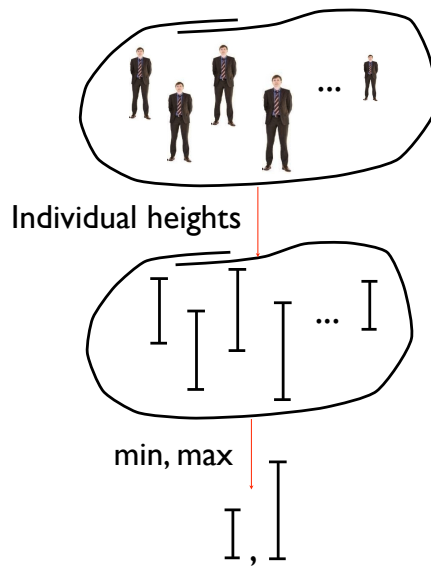
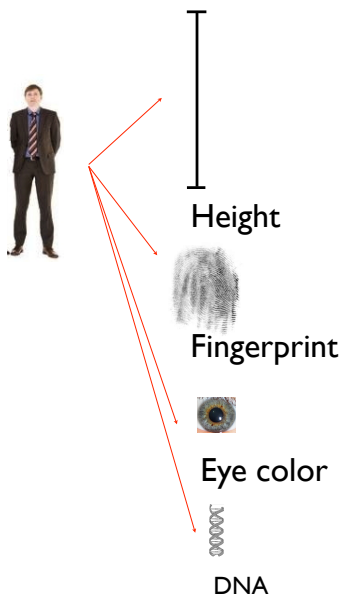


Ellipses:
 $x^2 + by^2 - axy \leq d$



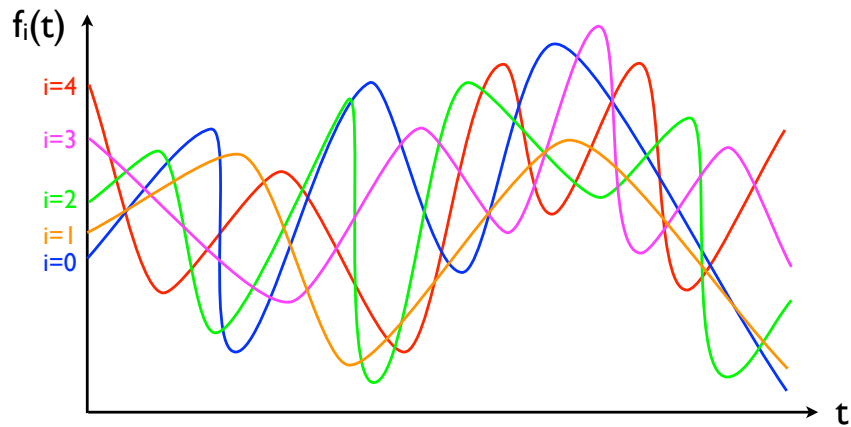
Exponentials:
 $-a^{bt} \leq y(t) \leq a^{bt}$

Abstractions of a man / crowd



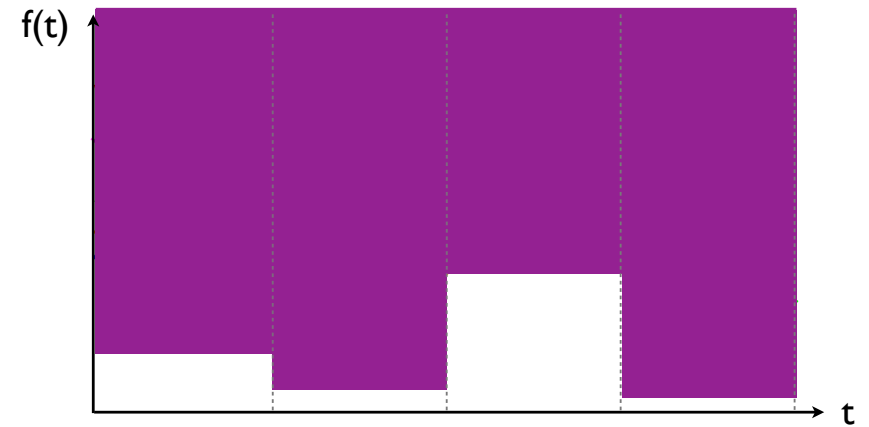
A slightly more detailed example

Set of functions

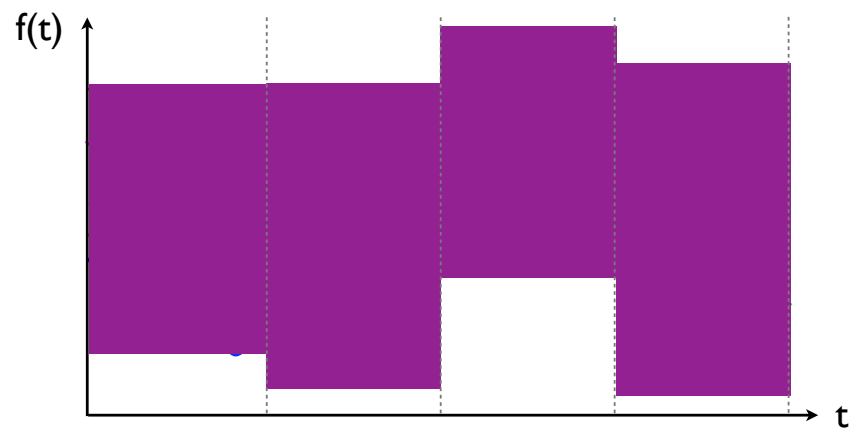


How to approximate $\{ f_1, f_2, f_3, f_4 \}$?

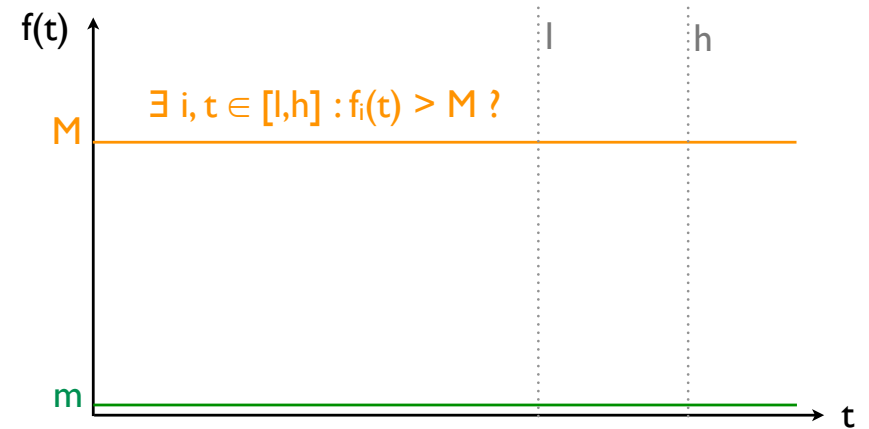
A less precise abstraction



Set of functions abstraction

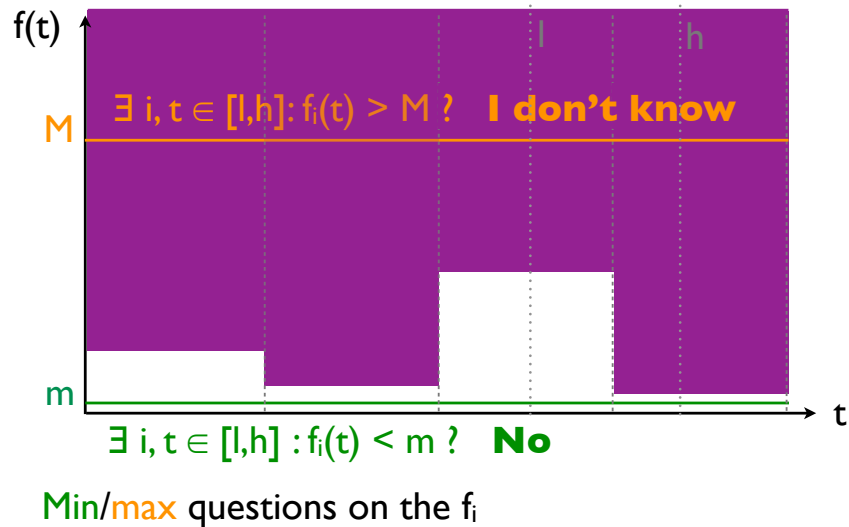


Concrete questions on the f_i

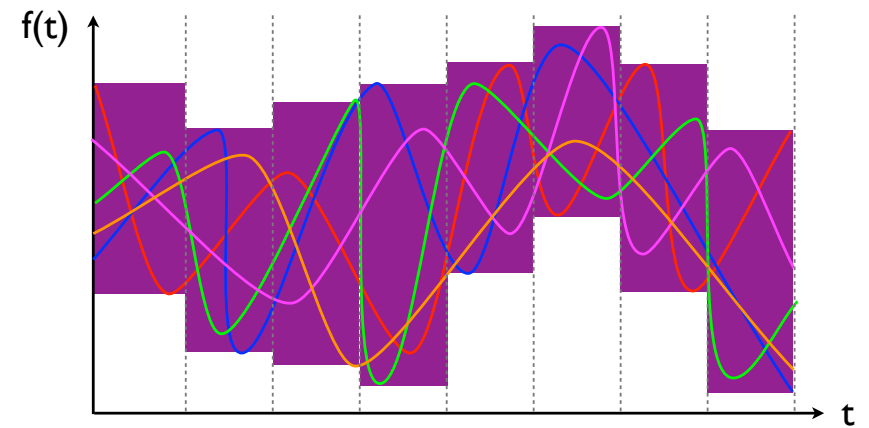


Min/max questions on the f_i

Concrete questions answered in the abstract

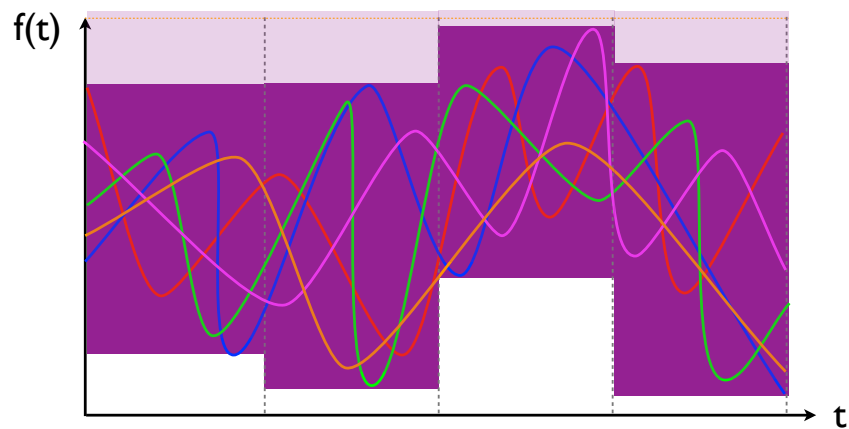


A more precise/refined abstraction

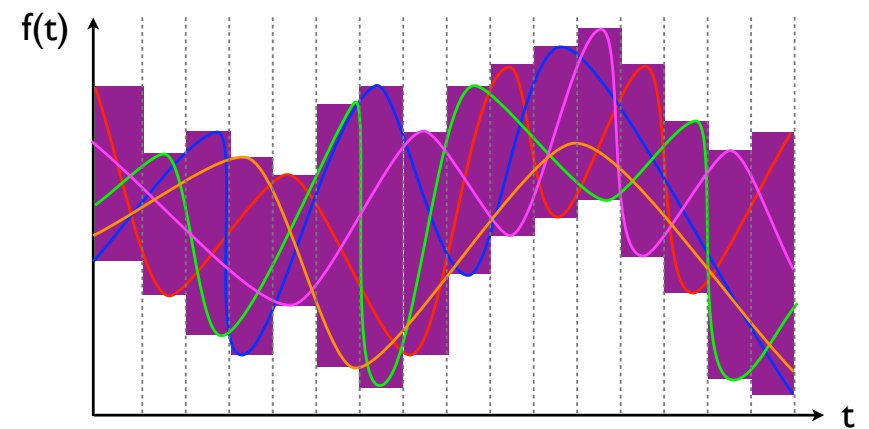


Soundness of the abstraction

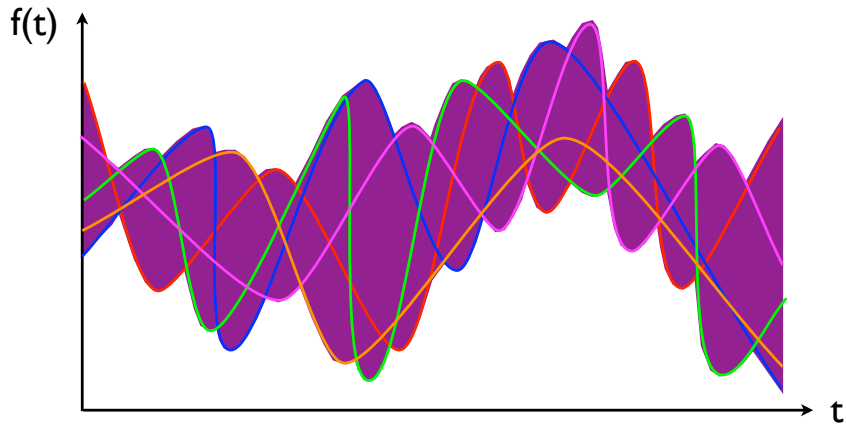
- No concrete case is ever forgotten:



An even more precise/refined abstraction

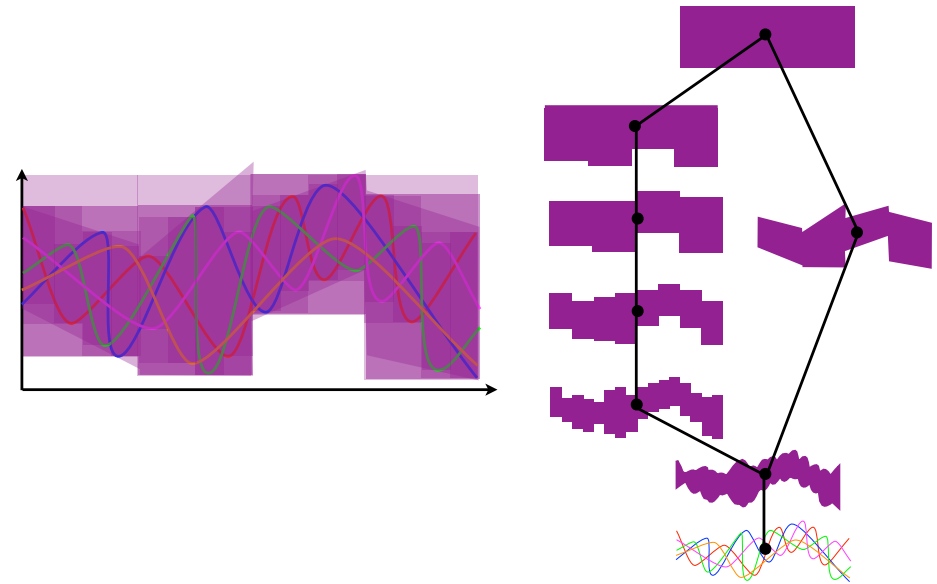


Passing to the limit

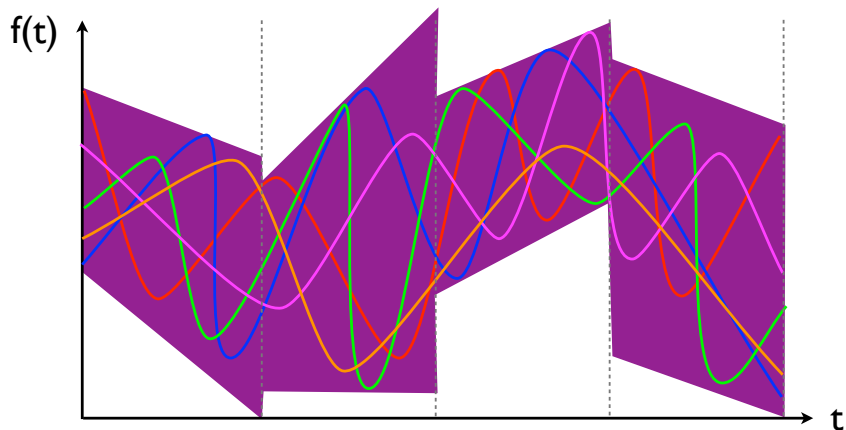


Sound and *complete* abstraction for min/max questions on the f_i

The hierarchy of abstractions



A non-comparable abstraction



Sound and *incomplete* abstraction for min/max questions on the f_i

Elements of Abstract Interpretation Theory Explained with ...

Patrick Cousot & Radhia Cousot, Static Determination of Dynamic Properties of Programs. In B. Robinet, editor, *Proceedings of the second international symposium on Programming*, Paris, France, pages 106–130, April 13-15 1976, Dunod, Paris.

Patrick Cousot, Radhia Cousot: Abstract Interpretation: A Unified Lattice Model for Static Analysis of Programs by Construction or Approximation of Fixpoints. *POPL* 1979: 269-282

Patrick Cousot, Radhia Cousot: Systematic Design of Program Analysis Frameworks. *POPL* 1979: 269-282

Patrick Cousot, Méthodes itératives de construction et d'approximation de points fixes d'opérateurs monotones sur un treillis, analyse sémantique des programmes. *Thèse Ès Sciences Mathématiques*, Université Joseph Fourier, Grenoble, France, 21 March 1978

Patrick Cousot, Semantic foundations of program analysis. In S.S. Muchnick & N.D. Jones, editors, *Program Flow Analysis: Theory and Applications*, Ch. 10, pages 303–342, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, U.S.A., 1981.

Elements of Abstract Interpretation Theory Explained with ... Flowers

Patrick Cousot & Radhia Cousot, Static Determination of Dynamic Properties of Programs. In B. Robinet, editor, *Proceedings of the second international symposium on Programming*, Paris, France, pages 106–130, April 13-15 1976, Dunod, Paris.

Patrick Cousot, Radhia Cousot: Abstract Interpretation: A Unified Lattice Model for Static Analysis of Programs by Construction or Approximation of Fixpoints. *POPL* 1977: 238-252

Patrick Cousot, Radhia Cousot: Systematic Design of Program Analysis Frameworks. *POPL* 1979: 269-282

Patrick Cousot, Méthodes itératives de construction et d'approximation de points fixes d'opérateurs monotones sur un treillis, analyse sémantique des programmes. *Thèse Ès Sciences Mathématiques*, Université Joseph Fourier, Grenoble, France, 21 March 1978

Patrick Cousot, Semantic foundations of program analysis. In S.S. Muchnick & N.D. Jones, editors, *Program Flow Analysis: Theory and Applications*, Ch. 10, pages 303–342, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, U.S.A., 1981.

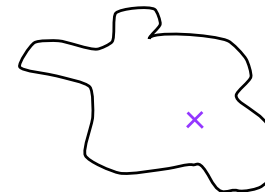
A mini graphical language

- Objects $o \in O$
- Operations on objects $O^n \longrightarrow O, n \geq 0$
- Logical operations on objects $O^n \longrightarrow \text{Booleans}, n \geq 0$

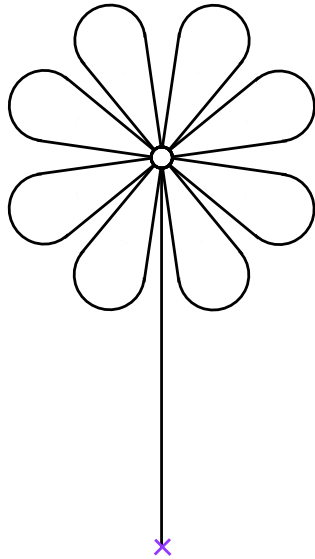
The concrete world

Objects

- An object $o \in O$ is defined by
 - An origin (a reference point \times)
 - A set of (infinitely small) black pixels (on a white background)
- Example I of object:

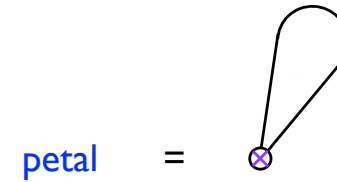


An example II of object: a flower



Constant objects

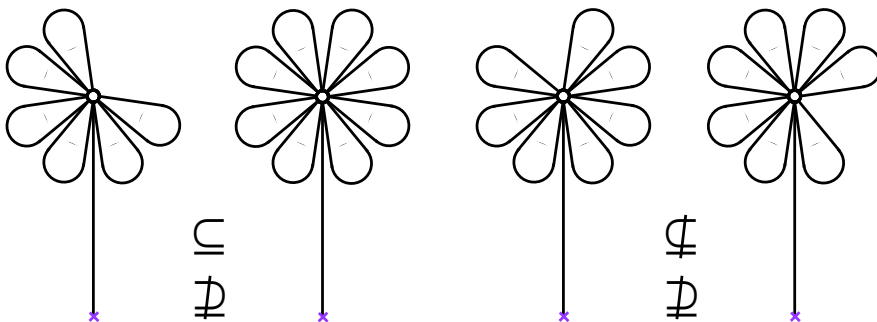
- A **petal** is an example of constant object



Logical operations on objects

- Inclusion \subseteq

- Examples:

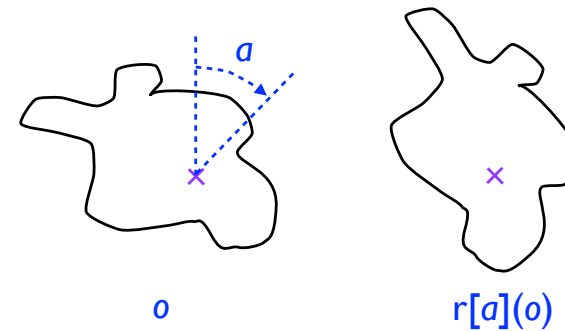


Operation on objects: rotation

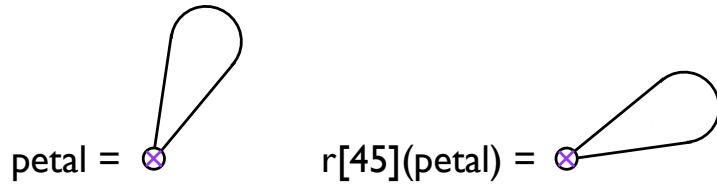
- rotation

$r[a](o)$

rotates the object o clockwise by angle a degrees around its origin



Example I of rotation

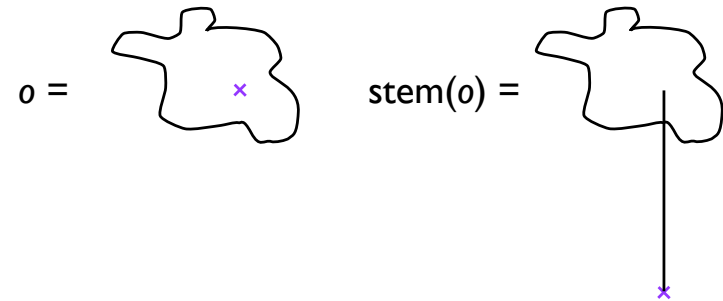


Operation on objects: add a stem

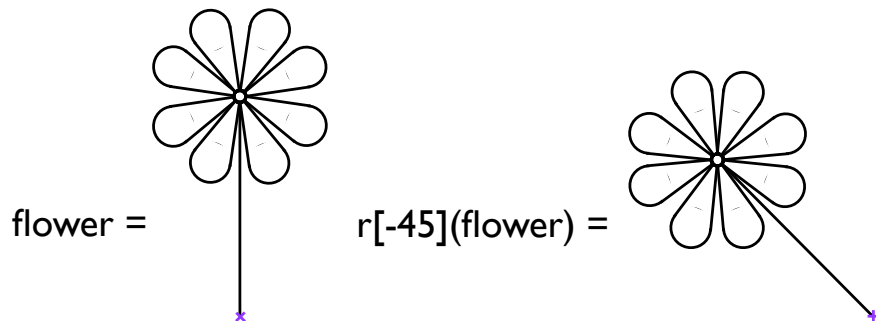
- Add a stem

$\text{stem}(o)$

adds a stem to object o (up to the origin of object o , with new origin at the root of the stem)



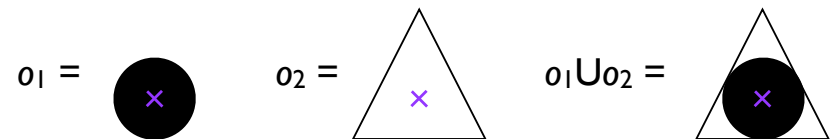
Example II of rotation



Operation on objects: union

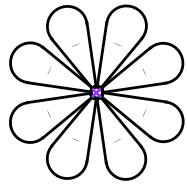
- The union $o_1 \cup o_2$ of objects o_1 and o_2 is the superposition of the pixels of o_1 and o_2 at their origins

- Example:

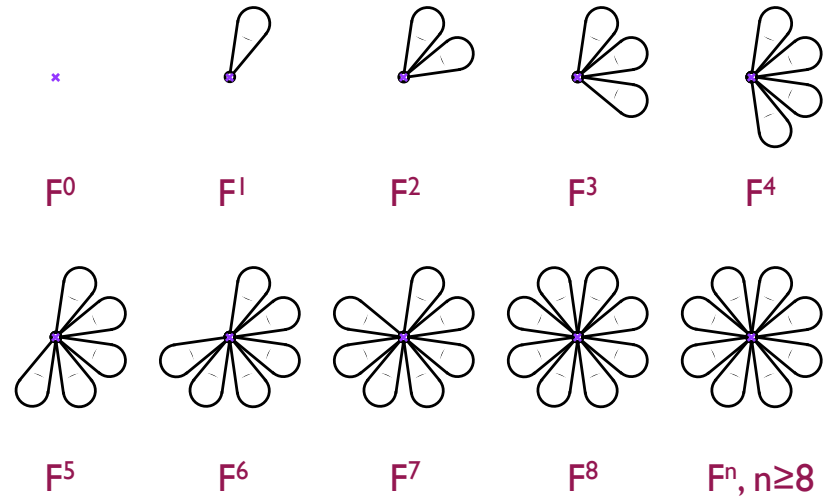


Example: corolla

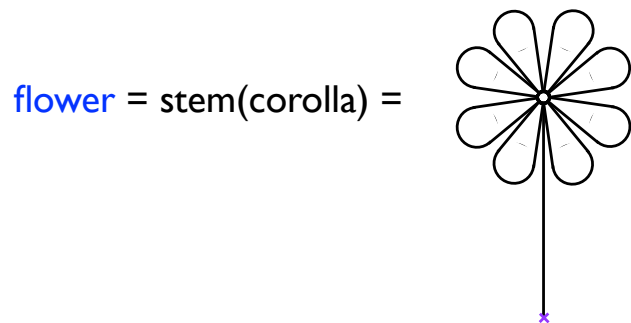
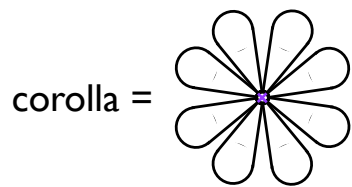
- $\text{corolla} = \text{petal} \cup r[45](\text{petal}) \cup r[90](\text{petal}) \cup r[135](\text{petal}) \cup r[180](\text{petal}) \cup r[225](\text{petal}) \cup r[270](\text{petal}) \cup r[315](\text{petal})$



Building a corolla iteratively



flower



Corolla transformer

- $F(X) = r[45]X \cup \text{petal}$

- Example:

- $X =$



- $r[45]X =$



- $r[45]X \cup \text{petal} =$



Iterates of a transformer to a fixpoint

- The iterates $F^n, n \geq 0$, of F from the empty set \emptyset are

$$F^0 = \emptyset$$

$$F^1 = F(F^0)$$

$$F^2 = F(F^1)$$

...

$$F^{n+1} = F(F^n)$$

...

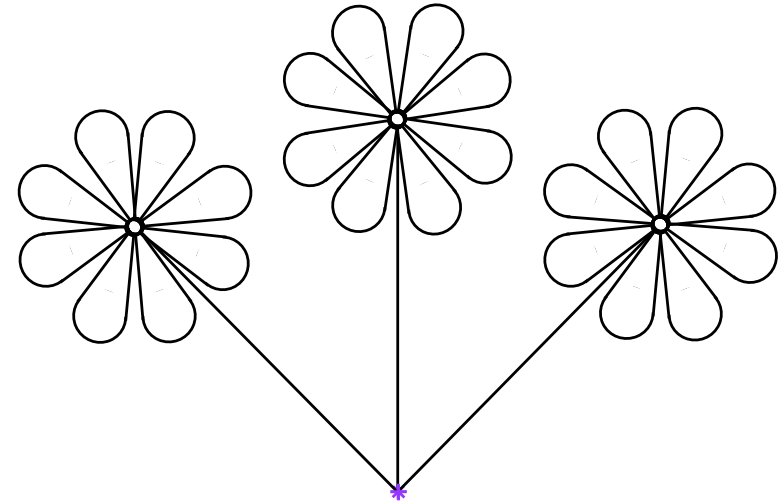
$$F^\omega = \bigcup_{n \geq 0} F^n = \text{lfp } F \quad (\text{assuming } F \text{ continuous})$$

- Least fixpoint: $F(\text{lfp } F) = \text{lfp } F$, and $F(x) = x$ implies $\text{lfp } F \subseteq x$

Patrick Cousot & Radhia Cousot. Constructive versions of Tarski's fixed point theorems. In *Pacific Journal of Mathematics*, Vol. 82, No. 1, 1979, pp. 43–57.

Concrete bouquet

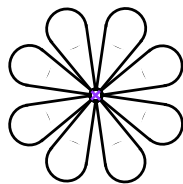
- bouquet = $r[-45](\text{flower}) \cup \text{flower} \cup r[45](\text{flower})$



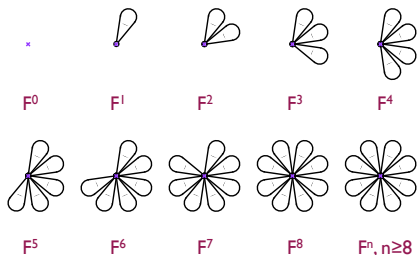
Fixpoint corolla

- $F(X) = r[45]X \cup \text{petal}$

- corolla = $\text{lfp } F =$



- Proof: the iterates are



The abstract world

Over-approximation

- An **over-approximation** of an object o is an object \bar{o} with
 - same origin
 - more pixels
- The **dual**⁽¹⁾ is an **under-approximation**, with less pixels

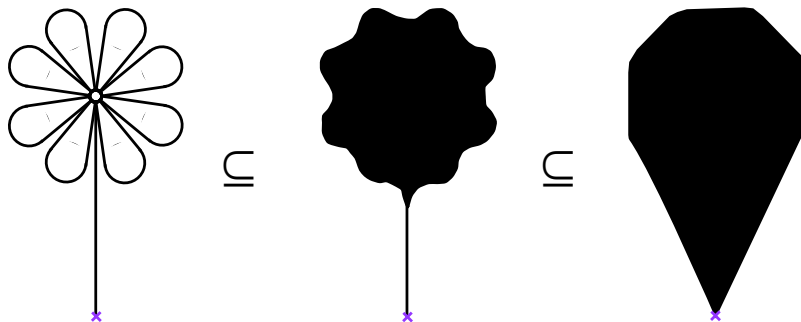
⁽¹⁾ Patrick Cousot, Méthodes itératives de construction et d'approximation de points fixes d'opérateurs monotones sur un treillis, analyse sémantique des programmes, *Thèse Ès Sciences Mathématiques*, Université Joseph Fourier, Grenoble, France, 21 March 1978
 Patrick Cousot, Semantic foundations of program analysis. In S.S. Muchnick & N.D. Jones, editors, *Program Flow Analysis: Theory and Applications*, Ch. 10, pages 303–342. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, U.S.A., 1981.

Abstraction

- An **abstraction** of an object o is a mathematical/computer representation of an over-approximation of this object o
- The abstraction is sometimes **exact** else is a **strict over-approximation**
 - Examples abstraction by plain squares

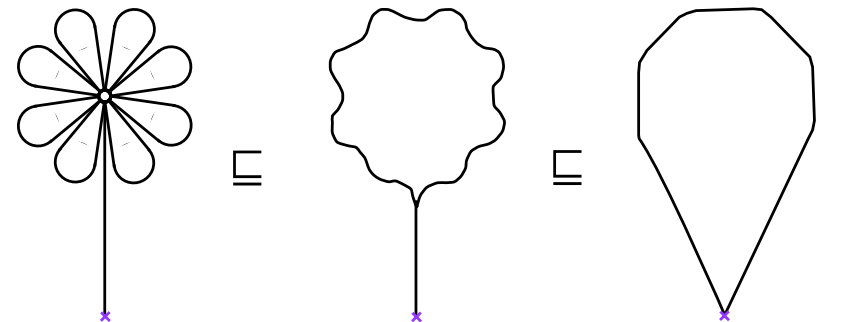


Examples of over-approximations of flowers



Examples of abstractions of flowers

- Encode a concrete over-approximation by its **outline**



concrete

abstract

more abstract

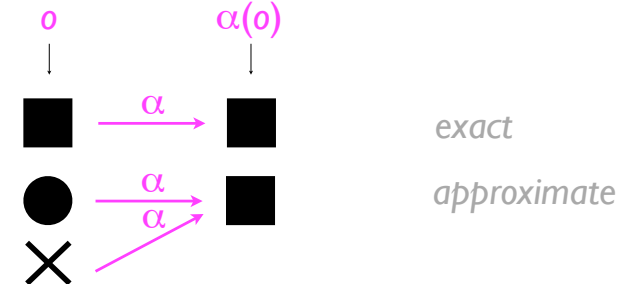
A Touch of Abstract Interpretation Theory

Patrick Cousot, Radhia Cousot: Abstract Interpretation: A Unified Lattice Model for Static Analysis of Programs by Construction or Approximation of Fixpoints. POPL 1977: 238-252
 Patrick Cousot, Radhia Cousot: Systematic Design of Program Analysis Frameworks. POPL 1979: 269-282

Abstraction function

- The abstraction function $\alpha \in O \rightarrow \bar{O}$ maps concrete objects $o \in O$ to their approximation by an abstract object $\alpha(o) \in \bar{O}$

- Example I of abstraction function by plain squares:

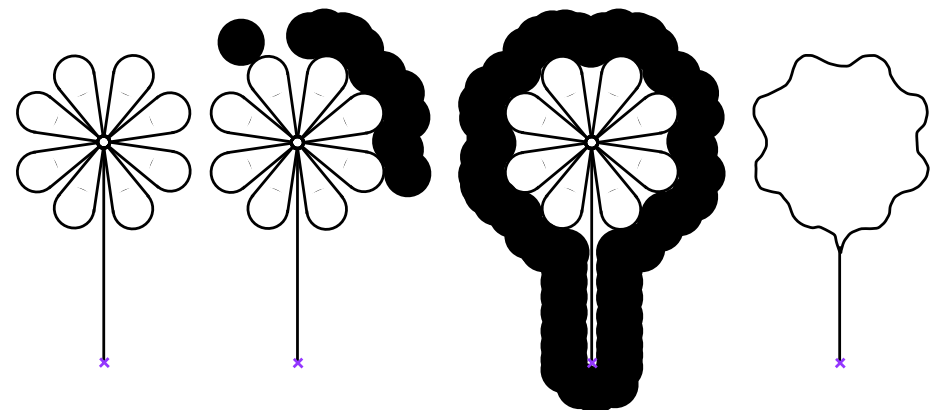


Patrick Cousot, Radhia Cousot: Systematic Design of Program Analysis Frameworks. POPL 1979: 269-282
 Patrick Cousot, Radhia Cousot, Laurent Mauborgne: The Reduced Product of Abstract Domains and the Combination of Decision Procedures. FOSSACS 2011: 456-472

Abstract domain

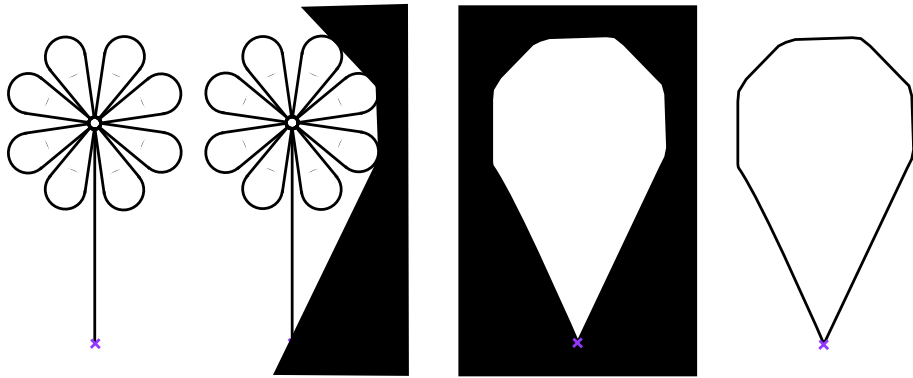
- An abstract domain is
 - a set of abstract objects \bar{O} (abstracting concrete objects)
 - a set of abstract operations (abstracting the concrete operations) $\bar{O}^n \rightarrow \bar{O}, n \geq 0$
 - a set of logical abstract operations $\bar{O}^n \rightarrow \text{Booleans}, n \geq 0$

Example II of abstraction function



Outlining brush: ●

Example III of abstraction function



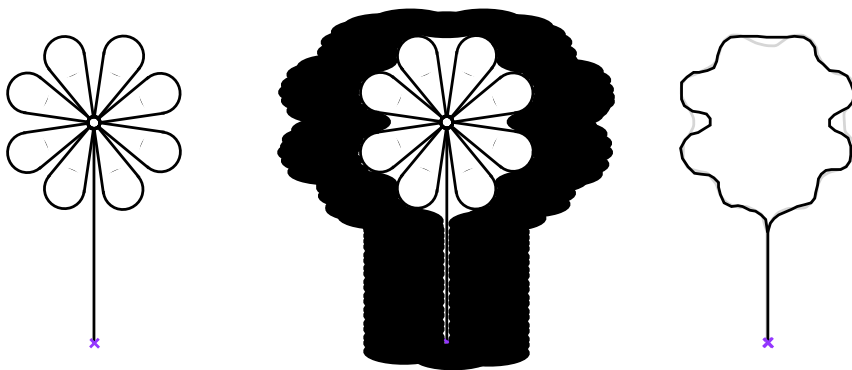
Outlining brush of infinite diameter

Concretization function

- A concretization function $\gamma \in \bar{O} \rightarrow O$ maps an abstract object $\bar{o} \in \bar{O}$ to the concrete objects $\gamma(\bar{o}) \in O$ that it represents/approximates
- $\gamma(\bar{o})$ is the concrete meaning/semantics of \bar{o}

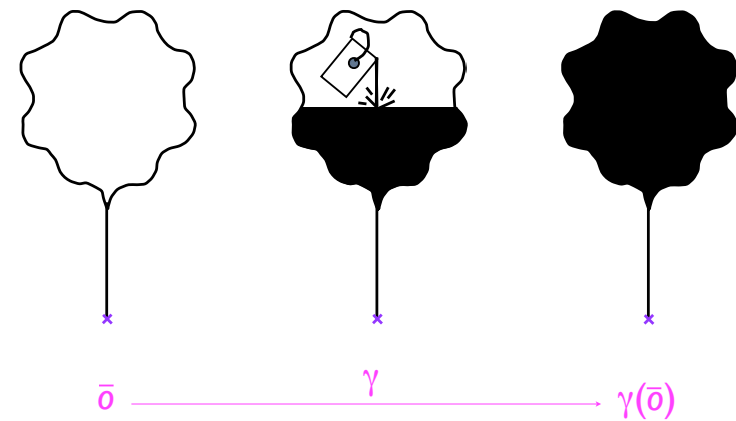
The hierarchy of abstractions

- Larger brush diameter: more abstract
- Different brush shapes: may be non-comparable abstractions



Oval outlining brush: 

Example of concretization

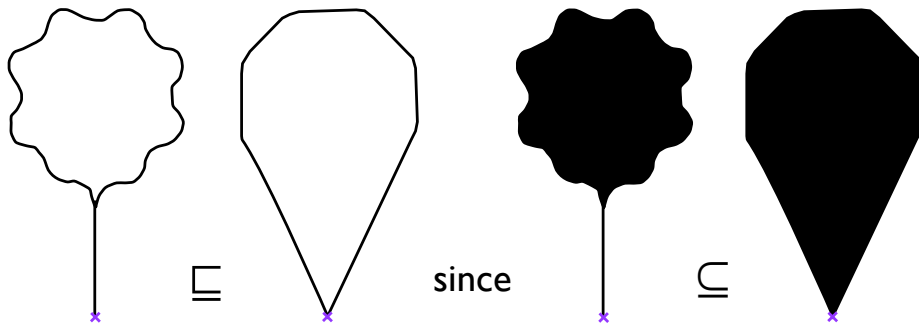


Abstract logical operation: abstract inclusion

- The **abstract flower inclusion** is defined as

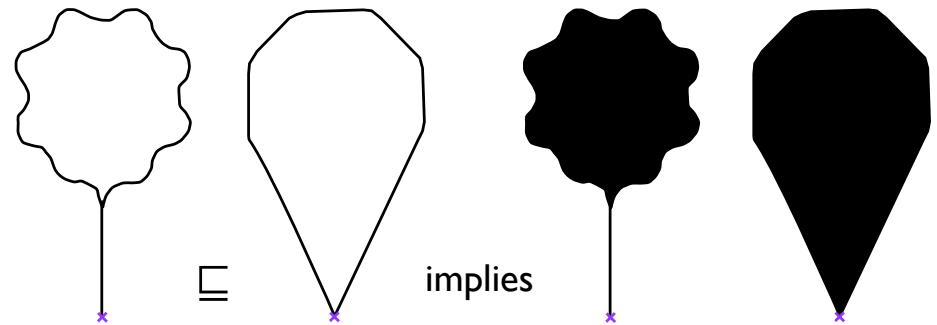
$$\bar{o}_1 \sqsubseteq \bar{o}_2 \text{ if and only if } \gamma(\bar{o}_1) \subseteq \gamma(\bar{o}_2)$$

- Example:



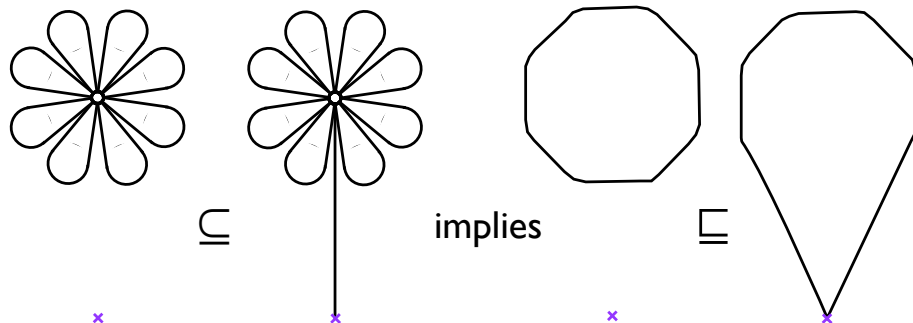
Galois connection 2/4

- γ is increasing
- Proof: by definition of \sqsubseteq , $\bar{o}_1 \sqsubseteq \bar{o}_2$ implies $\gamma(\bar{o}_1) \subseteq \gamma(\bar{o}_2)$



Galois connection 1/4

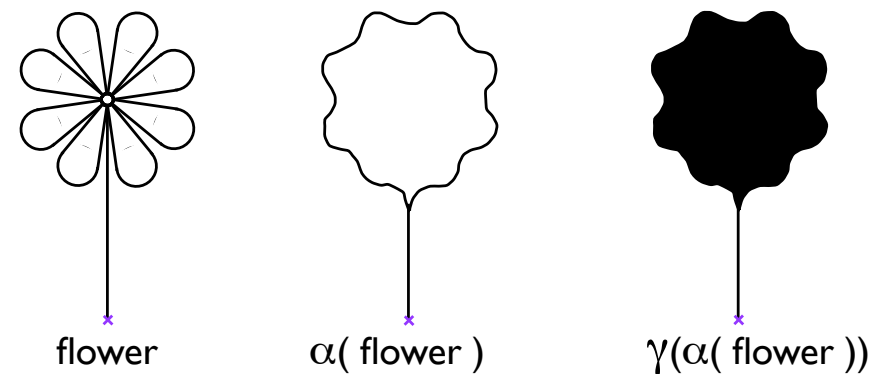
- α is increasing



The larger the concrete, the larger the abstract

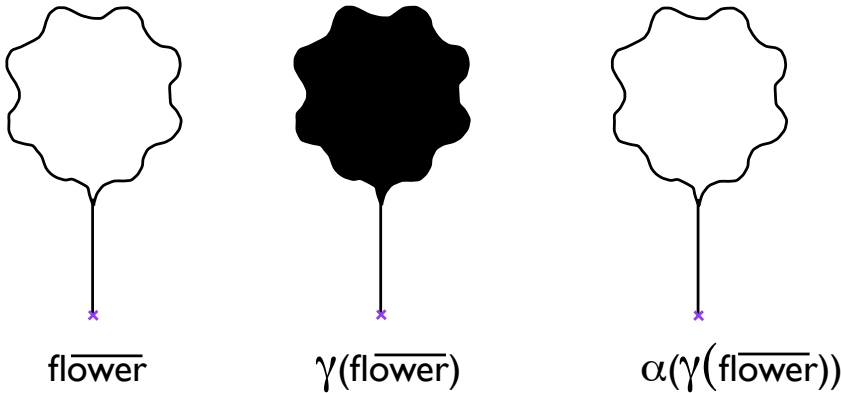
Galois connection 3/4

- For all concrete objects $x \in O$, $x \subseteq \gamma \circ \alpha(x)$
- Intuition: **soudness** (over-approximation)



Galois connection 4/4

- For all abstract objects $y \in \bar{O}$, $\alpha \circ \gamma(y) = y$
- Intuition: α returns the most precise abstraction



Example of biological abstraction

- Let *Species* be the set of all chemical species ($C, c_1, c'_1, \dots, c_k, c'_k, \dots \in \text{Species}$).
- Let *Local_view* be the set of all local views
- Let $\alpha \in \wp(\text{Species}) \rightarrow \wp(\text{Local_view})$ be the function that maps any set of complexes into the set of their local views.
- $$\alpha(\{R(Y1 \sim u, l!1), E(r!1)\}) = \{R(Y1 \sim u, l!r.E); E(r!l.R)\}$$
- The function α defines a Galois connexion: $\wp(\text{Species}) \xrightleftharpoons[\alpha]{\gamma} \wp(\text{Local_view})$
- (The function γ maps a set of local views into the set of complexes that can be built with these local views).

Jérôme Feret, Reachability Analysis of Biological Signalling Pathways by Abstract Interpretation. In Proceedings of the International Conference of Computational Methods in Sciences and Engineering (ICCMSE 2007), Corfu, Greece, 25-30 september, T.E. Simos(Ed.), 2007, American Institute of Physics conference proceedings 963(2), pp 619-622.
Vincent Danos, Jérôme Feret, Walter Fontana, Jean Krivine: Abstract Interpretation of Cellular Signalling Networks. VMCAI 2008: 83-97

Galois connection: all in one

- Notation:

$$\langle O, \sqsubseteq \rangle \xrightleftharpoons[\alpha]{\gamma} \langle \bar{O}, \sqsubseteq \rangle$$

- Equivalent definition

$$\forall o \in O, \bar{o} \in \bar{O}: \alpha(o) \sqsubseteq \bar{o} \quad \text{iff} \quad o \sqsubseteq \gamma(\bar{o})$$

and

\Rightarrow soundness
 \Leftarrow best abstraction

α surjective

(otherwise $\alpha \circ \gamma(y) \sqsubseteq y$)

Specification of abstract operations

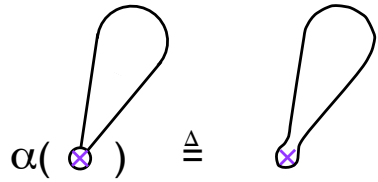
- $\text{cte} \triangleq \alpha(\text{cte})$ constant
- $\text{op}_1(x) \triangleq \alpha(\text{op}_1(\gamma(x)))$ unary
- $\text{op}_2(x, y) \triangleq \alpha(\text{op}_2(\gamma(x), \gamma(y)))$ binary
-
- $\text{op}_n(x_1, \dots, x_n) \triangleq \alpha(\text{op}_n(\gamma(x_1), \dots, \gamma(x_n)))$ n-ary

- Can be less precise

$$\alpha(\text{op}_n(\gamma(x_1), \dots, \gamma(x_n))) \sqsubseteq \text{op}_n(x_1, \dots, x_n)$$

Abstract constants

- Abstract petal



A commutation theorem on rotation

- $\alpha(r[a](y)) = \bar{r}[a](\alpha(y)) \quad \forall y \in \bar{O}$

- Proof:

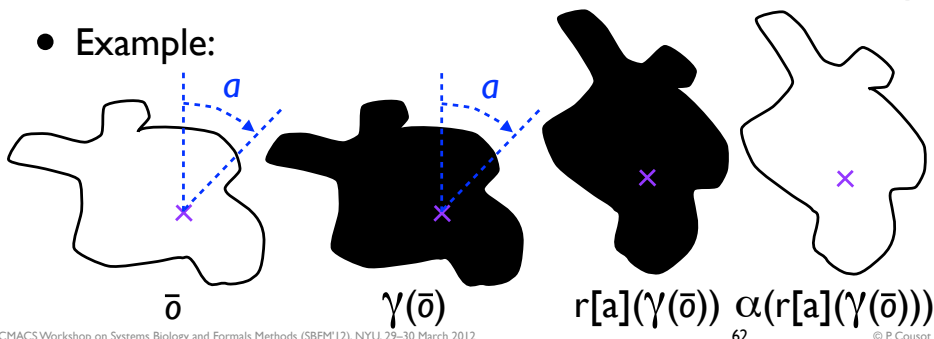
$$\begin{aligned} & \alpha(r[a](y)) && \alpha \circ \gamma \text{ is the identity} \\ & = \alpha(\gamma(\alpha(r[a](y)))) && \alpha \text{ preserves rotation} \\ & = \alpha(\gamma(r[a](\alpha(y)))) && \gamma \text{ preserves rotation} \\ & = \alpha(r[a](\gamma(\alpha(y)))) && \text{definition abstract rotation} \\ & \stackrel{\Delta}{=} \bar{r}[a](\alpha(y)) \end{aligned}$$

Abstract rotation

- Abstract rotation

$$\begin{aligned} \bar{r}[a](\bar{o}) & \stackrel{\Delta}{=} \alpha(r[a](\gamma(\bar{o}))) && \text{definition} \\ & = \alpha(\gamma(r[a](\bar{o}))) && \text{rotation preserves shape} \\ & = r[a](\bar{o}) && \text{identity} \end{aligned}$$

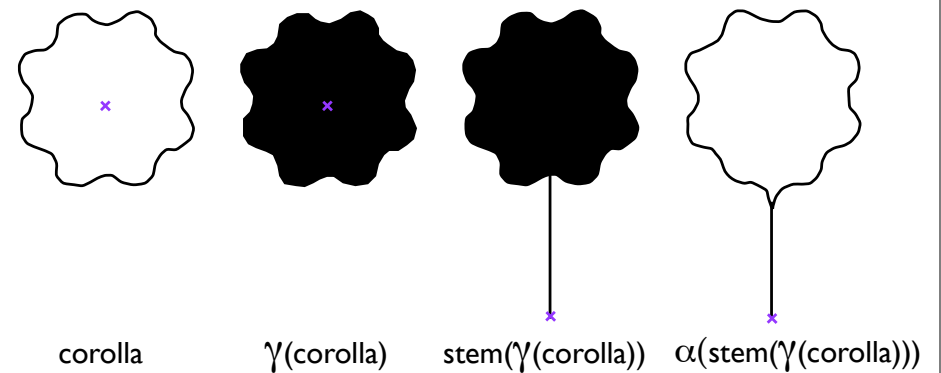
- Example:



Abstract stems

- $\overline{\text{stem}}(y) \stackrel{\Delta}{=} \alpha(\text{stem}(\gamma(y)))$

- Example:



Abstract union

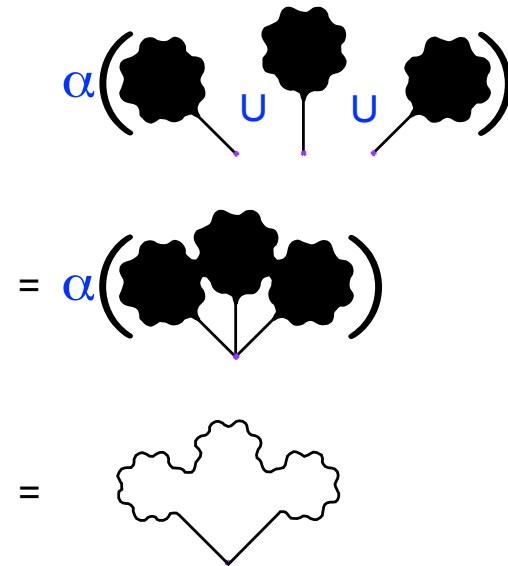
- $x \sqcup y \triangleq \alpha(\gamma(x) \cup \gamma(y))$

- Join abstraction theorem:

$$\alpha(x) \sqcup \alpha(y) = \alpha(x \cup y)$$

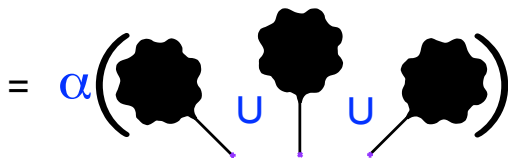
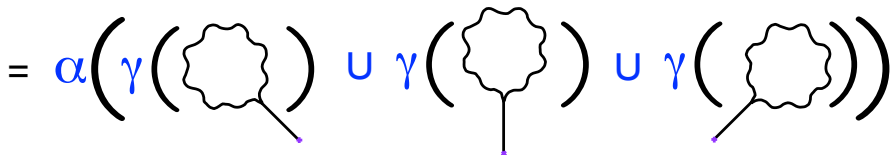
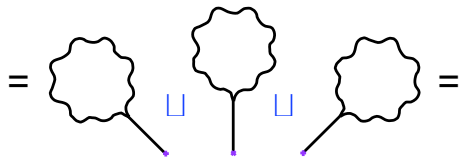
Galois connection

Abstract bouquet (cont'd)



Abstract bouquet (cont'd)

- $\overline{\text{bouquet}} = \overline{r[-45](\text{flower})} \sqcup \overline{\text{flower}} \sqcup \overline{r[45](\text{flower})}$



A theorem on abstract bouquets

- $\overline{\text{bouquet}}$

$$= \overline{r[-45](\text{flower})} \sqcup \overline{\text{flower}} \sqcup \overline{r[45](\text{flower})}$$

flower = $\alpha(\text{flower})$

$$= \overline{r[-45](\alpha(\text{flower}))} \sqcup \overline{\alpha(\text{flower})} \sqcup \overline{r[45](\alpha(\text{flower}))}$$

rotation commutation theorem

$$= \alpha(\overline{r[-45](\text{flower})}) \sqcup \overline{\alpha(\text{flower})} \sqcup \alpha(\overline{r[45](\text{flower})})$$

join abstraction theorem

$$= \alpha(\overline{r[-45](\text{flower})}) \sqcup \overline{\alpha(\text{flower})} \sqcup \alpha(\overline{r[45](\text{flower})})$$

definition concrete bouquet

$$= \overline{\alpha(\text{bouquet})}$$

Abstract corolla transformer

- Corolla transformer commutation theorem:

- $\alpha(F(x))$

$$= \alpha(\text{petal} \cup r[45](x)) \quad \text{definition F}$$

$$= \alpha(\text{petal}) \sqcup \alpha(r[45](x)) \quad \text{join abstraction theorem}$$

$$= \overline{\text{petal}} \sqcup \alpha(r[45](x)) \quad \text{definition abstract petal}$$

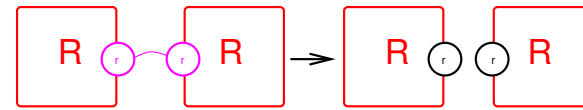
$$= \overline{\text{petal}} \sqcup \bar{r}[45](\alpha(x)) \quad \text{rotation commut. theorem}$$

$$= \bar{F}(\alpha(x))$$

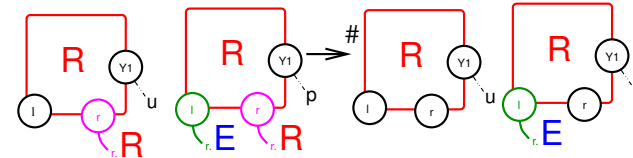
$$\text{by defining } \bar{F}(y) = \overline{\text{petal}} \sqcup \bar{r}[45](y)$$

Example of biological transformer

- Concrete rule:



- Abstract rule:



Abstract transformer

- An abstract transformer \bar{F} is

- *Sound* iff

$$\forall P \in \mathcal{P} : \alpha \circ F(P) \sqsubseteq \bar{F} \circ \alpha(P)$$

- *Complete* iff

$$\forall P \in \mathcal{P} : \alpha \circ F(P) = \bar{F} \circ \alpha(P)$$

Fixpoint abstraction

- For an increasing and sound abstract transformer, we have a *fixpoint approximation*

$$\alpha(\text{lfp}^{\leq} F) \sqsubseteq \text{lfp}^{\sqsubseteq} \bar{F}$$

- For an increasing, sound, and complete abstract transformer, we have an *exact fixpoint abstraction*

$$\alpha(\text{lfp}^{\leq} F) = \text{lfp}^{\sqsubseteq} \bar{F}$$

Abstract corolla

- $\overline{\text{corolla}} = \alpha(\text{corolla}) = \alpha(\text{lfp}^{\subseteq} F) = \text{lfp}^{\subseteq} \bar{F}$

since $F(x) = \text{petal} \cup r[45](x)$

and $\bar{F}(y) = \overline{\text{petal}} \sqcup \bar{r}[45](y)$

do commute: $\alpha(F(x)) = \bar{F}(\alpha(x))$

Example of biological fixpoint

- Concrete reachability transformer:

$$\mathbb{F} : \begin{cases} \wp(\text{Species}) & \rightarrow \wp(\text{Species}) \\ X & \mapsto X \cup \left\{ c'_j \mid \begin{array}{l} \exists R_k \in \mathcal{R}, c_1, \dots, c_m \in X, \\ c_1, \dots, c_m \rightarrow_{R_k} c'_1, \dots, c'_n \end{array} \right\} \end{cases}$$

- Reachable species from Species_0

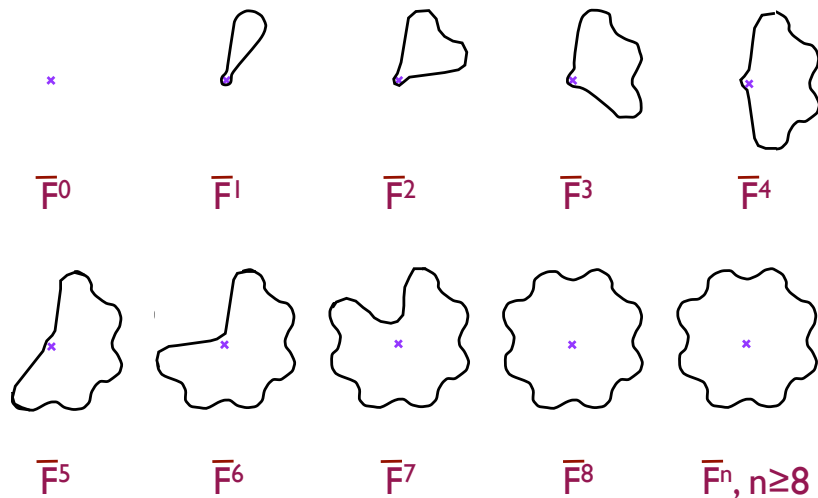
$$\text{lfp}^{\subseteq} \lambda X. \text{Species}_0 \cup \mathbb{F}(X)$$

- Abstract reachability transformer:

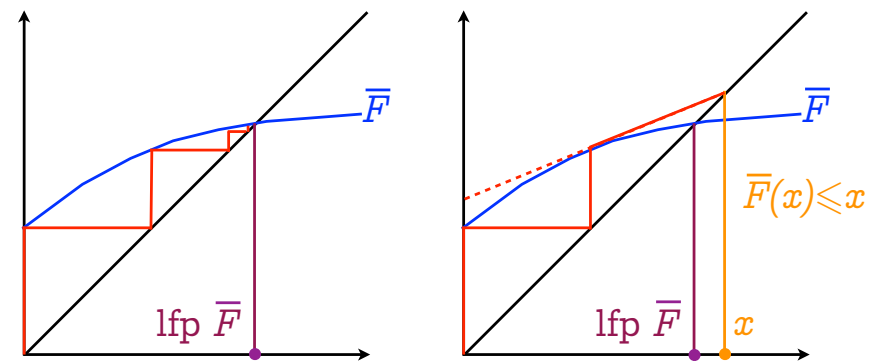
$$\mathbb{F}^{\sharp} : \begin{cases} \wp(\text{Local_view}) & \rightarrow \wp(\text{Local_view}) \\ X & \mapsto X \cup \left\{ l'_j \mid \begin{array}{l} \exists R_k \in \mathcal{R}, l_1, \dots, l_m \in X, \\ l_1, \dots, l_m \rightarrow_{R_k}^{\sharp} l'_1, \dots, l'_n \end{array} \right\} \end{cases}$$

Jérôme Feret, Reachability Analysis of Biological Signalling Pathways by Abstract Interpretation. In Proceedings of the International Conference of Computational Methods in Sciences and Engineering (ICCMSE 2007), Corfu, Greece, 25-30 september, T.E. Simos(Ed.), 2007, American Institute of Physics conference proceedings 963(2), pp 619-622.
Vincent Danos, Jérôme Feret, Walter Fontana, Jean Krivine: Abstract Interpretation of Cellular Signalling Networks. VMCAI 2008: 83-97

Iterates for the abstract corolla



Convergence acceleration with widening



Infinite iteration

Accelerated iteration with widening
(e.g. with a widening based on the derivative as in Newton-Raphson method)

Abstraction of the graphical language

- Any graphical program can be abstracted by replacing the concrete objects/operations by abstract ones
- The soundness follows by induction on the syntax of programs

Software

- **Ait**: static analysis of the worst-case execution time of control/command software (www.absint.com/ait/)
- **Astrée**: proof of absence of runtime errors in embedded synchronous real time control/command software (www.absint.com/astree/), **AstréeA** for asynchronous programs (www.astreea.ens.fr/)
- **C Global Surveyor**, NASA, static analyzer for flight software of NASA missions (www.cmu.edu/silicon-valley/faculty-staff/venet-arnaud.html)
- **Checkmate**: static analyzer of multi-threaded Java programs (www.pietro.ferrara.name/checkmate/)
- **CodeContracts Static Checker**, Microsoft (msdn.microsoft.com/en-us/devlabs/dd491992.aspx)
- **Fluctuat**: static analysis of the precision of numerical computations (www-list.cea.fr/labs/gb/LSI/fluctuat/index.html)

Applications of Abstract Interpretation in Computer Science

See **Software Horror Stories** (www.cs.tau.ac.il/~nachumd/horror.html)

Software

- **Infer**: Static analyzer for C/C++ (monoidics.com/)
- **Julia**: static analyzer for Java and Android programs (www.juliasoft.com/juliasoft-android-java-verification.aspx?Id=201177234649)
- **Predator**: static analyzer of C dynamic data structures using separation logic (www.fit.vutbr.cz/research/groups/verifit/tools/predator/)
- **Terminator**: termination proof (www.cs.ucl.ac.uk/staff/p.ohearn/Invader/Invader/Invader_Home.html)

• etc

Libraries:

- **Apron** numerical domains library (apron.cri.enscm.fr/library/)
- **Parma Polyhedral Library** (bugseng.com/products/ppl/)
- etc

Hardware

- (Generalized) symbolic trajectory evaluation (Intel)

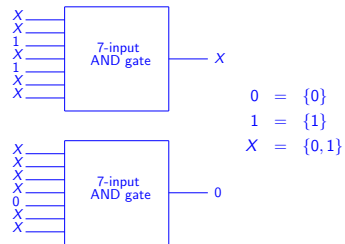
Intel's Successes with Formal Methods

John Harrison
Intel Corporation
15 March 2012

Tsinghua software day, March 15, 2012, Tsinghua University, Beijing, China

Example of ternary simulation

If some inputs are undefined, the output often is too, but not always:



Jin Yang and Carl-Johan H. Seger, *Generalized Symbolic Trajectory Evaluation – Abstraction in Action*, Formal Methods in Computer-Aided Design, Lecture Notes in Computer Science, 2002, Volume 2517/2002, 70–87.

Jin Yang, Seger, C.-J.H., *Introduction to generalized symbolic trajectory evaluation*, IEEE Transactions on Very Large Scale Integration (VLSI) Systems 11(3), June 2003, 345–353.

Conclusion

System biology

- See [SBFM'2012](#) !

Conclusion

If the simulation/analysis/checking of your model does not scale up, consider using (sound (and complete)) abstractions