Observational Completeness on Abstract Interpretation

Gianluca Amato and Francesca Scozzari

Dipartimento di Scienze, Università di Chieti-Pescara
{amato,scozzari}@sci.unich.it

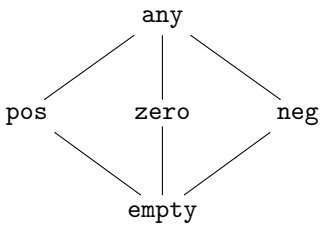
Abstract. In the theory of abstract interpretation, we introduce the observational completeness, which extends the common notion of completeness. A domain is complete when abstract computations are as precise as concrete computations. A domain is observationally complete for an observable π when abstract computations are as precise as concrete computations, if we only look at properties in π. We prove that continuity of state-transition functions ensures the existence of the least observationally complete domain. When state-transition functions are additive, the least observationally complete domain boils down to the complete shell.

1 Introduction

Abstract Interpretation. Abstract interpretation [3, 4] is a general theory for approximating the behavior of a discrete dynamic system. The idea is to replace the formal semantics of a system with an abstract semantics, computed over a domain of abstract objects. There are many different methods to describe the semantics of a system. Most of them are based on a partially ordered set (poset) $\langle C, \leq_C \rangle$ of states and a set F of monotone state-transition functions $f : C \to C$. The semantics $\mathcal{S}$ is defined as the (least) fixpoint of a semantic function $\mathcal{F}$ obtained as a composition of state-transition functions. The poset $\langle C, \leq_C \rangle$ is called *concrete domain* and $\mathcal{S} = \operatorname{lfp} \mathcal{F}$ is called the *concrete semantics*.

An abstract interpretation is specified by the poset $\langle A, \leq_A \rangle$ of *abstract objects*. The abstract objects describe the properties of the system we are interested in. The relationship between the concrete and abstract objects is formalized by a monotone concretization map $\gamma : A \to C$ which, given a property $a \in A$, yields the biggest concrete state $c \in C$ which enjoys the property a. Therefore, a property a is a correct approximation of a concrete state c when $c \leq_C \gamma(a)$.

For instance, consider the concrete domain $\wp(\mathbb{Z})$ with the standard ordering given by inclusion, and $Sign = \{\texttt{empty}, \texttt{pos}, \texttt{neg}, \texttt{zero}, \texttt{any}\}$ ordered as depicted on the right. The intuition is that `pos` represents the set of (strictly) positive integers, `zero` represents the singleton $\{0\}$, while `empty` represents the empty set of integers. This may be formalized



by defining γ as follows: $\gamma(\texttt{empty}) = \emptyset,\ \gamma(\texttt{pos}) = \{n \in \mathbb{Z} \mid n > 0\},\ \gamma(\texttt{neg}) = \{n \in \mathbb{Z} \mid n < 0\},\ \gamma(\texttt{zero}) = \{0\},\ \gamma(\texttt{any}) = \mathbb{Z}$.

Often, it is possible to define a monotone abstraction map $\alpha : C \to A$ which yields the largest properties a enjoyed by a concrete object c, such that $c \leq_C \gamma(a) \iff \alpha(c) \leq_A a$. In the previous example, the abstraction is given by $\alpha(c) = \{a \in A \mid c \leq_C \gamma(a)\}$. For instance, $\alpha(\{-1, -2\}) = \texttt{neg}$ while $\alpha(\{-1, 0\}) = \texttt{any}$.

An *abstract domain* is given by the poset A of the abstract objects and the pair of maps $\langle \alpha, \gamma \rangle$. However, since γ is uniquely determined by α and viceversa, in the following we specify an abstract domain just by giving α.

The goal of any abstract interpretation is to compute $\alpha(\mathcal{S})$, that is to find out the properties enjoyed by the semantics of the system. Instead of computing $\mathcal{S}$ (which is not computable) and then applying α, the idea is to replace, in the the definition of $\mathcal{F}$, every state-transition function f with an abstract counterpart $f^\# : A \to A$, which must be *correct*. We say that $f^\#$ is correct if, whenever a is a correct approximation of c, then $f^\#(a)$ is a correct approximation of $f(c)$. This is equivalent to say that any abstract computation $f^\#(\alpha(c))$ approximates the corresponding concrete computation $f(c)$, i.e.:

$$\alpha \circ f \leq_{A \to A} f^\# \circ \alpha \ , \tag{1}$$

where $\leq_{A \to A}$ is the pointwise extension of $\leq_A$. In particular, there is a *best correct abstraction* of f, denoted by f^α, which is $f^\alpha = \alpha \circ f \circ \gamma$. If we replace the state-transition functions in the definition of $\mathcal{F}$ with the corresponding best-correct abstractions, we obtain a new semantic function $\mathcal{F}^\#$ and a new abstract semantics $\mathcal{S}^\# = \operatorname{lfp} \mathcal{F}^\#$, and the theory of abstract interpretation ensures that

$$\alpha(\mathcal{S}) \leq_{\mathcal{A}} \mathcal{S}^\# \ . \tag{2}$$

As an example of abstract state-transition function, consider $inc : \wp(\mathbb{Z}) \to \wp(\mathbb{Z})$ such that $inc(X) = \{n + 1 \mid n \in X\}$. The best correct abstraction of inc is

$$inc^\alpha(a) = \begin{cases} \texttt{empty} & \text{if } a = \texttt{empty}, \\ \texttt{pos} & \text{if } a = \texttt{zero} \text{ or } a = \texttt{pos}, \\ \texttt{any} & \text{otherwise.} \end{cases}$$

Completeness. Generally speaking, the inequalities (1) and (2) are strict. This means that computing in the abstract domain is (strictly) less precise than computing on the concrete one. For instance, $\alpha(inc(-1)) = \texttt{zero}$ but $inc^\alpha(\alpha(-1)) = \texttt{any}$. When $\alpha \circ f = f^\alpha \circ \alpha$ we say that the abstract domain is *complete* for the function f. Intuitively, when this happens, the best correct abstraction f^α perfectly mimics the concrete function f. For example, given $sq(X) = \{x^2 \mid x \in X\}$, the best correct abstraction is

$$sq^\alpha(a) = \begin{cases} \texttt{pos} & \text{if } a = \texttt{pos} \text{ or } a = \texttt{neg}, \\ a & \text{otherwise.} \end{cases}$$

It follows that $sq(\{-1,-2\}) = \{1,4\}$, and its abstraction is $\alpha(sq(\{-1,-2\})) = \texttt{pos}$, meaning that the square of any integer in $\{-1,-2\}$ is positive. The same result may be obtained by first abstracting $\{-1,-2\}$ and then computing sq^α, since $sq^\alpha(\alpha(\{-1,-2\})) = sq^\alpha(\texttt{neg}) = \texttt{pos}$. It is easy to show that, for any set of integers $X \in \wp(\mathbb{Z})$, it holds that $sq^\alpha(\alpha(X)) = \alpha(sq(X))$. Thus, the abstract domain *Sign* is complete for the function sq.

Completeness enjoys many good properties. If an abstract domain α is complete for f and g, then it holds that:

- α is complete for $f \circ g$ and $f^\alpha \circ g^\alpha = (f \circ g)^\alpha$;
- $\alpha(\mathrm{lfp}\, f) = \mathrm{lfp}(f^\alpha)$.

This implies that (2) is an equality, and therefore one does not lose precision by computing on the abstract domain.

When an abstract domain α is not as precise as the concrete one, that is, the abstract semantics $\mathcal{S}^\#$ does not coincide with $\alpha(\mathcal{S})$, then we need to refine the abstract domain α. This means to replace α by a new domain β and $\mathcal{S}^\#$ by a new abstract semantics $\mathcal{S}^\circ$, such that $\alpha(\mathcal{S})$ may be recovered by $\mathcal{S}^\circ$. Here, $\mathcal{S}^\circ$ is obtained by replacing all the state-transition functions f in $\mathcal{F}$ with f^β. Conceptually, the domain β is the *computational domain* and α is the *observational domain*, which contains all the properties we want to observe. The abstract objects in β which do not belong to α are only used to compute intermediate steps in order not to lose precision. Obviously, we want to keep β as small as possible.

In the literature of abstract interpretation, the standard way of refining α is to compute the least complete domain for F which includes α. This is called the *complete shell* of α, and may be constructively computed [8].

The Goal. In this paper, we show that the complete shell may not be the smallest abstract domain which enables us to recover the property $\alpha(\mathcal{S})$. This is because we are only interested in properties in α, and not in the new objects introduced by β. This observation suggests another notion of completeness, which we call *observational completeness*. A domain β is observationally complete for a function f and an observational domain α when every concrete computation may be approximated in β without losing precision on the properties in α. In order to formalize the observational completeness, we first need to introduce a new ordering between abstract domains. We say that an abstract domain β is *more precise than* an abstract domain β' for observing properties in α whenever the result of each computation on β observed on α is approximated by the result of the corresponding computation on β', observed on α. We show that, under suitable conditions, there exists the smallest observationally complete domain for a given set F of functions and an observational domain. We prove that any complete domain which contains α is also observationally complete for α, but the converse does not hold. We give the conditions under which the least observationally complete domain corresponds to the complete shell.

Plan of the Paper. The next section recalls some basic definitions and notations about abstract interpretation. In Sect. 3 we define the notion of observational completeness, in Sect. 4 we study the relationships between observational completeness and standard completeness. In Sect. 5 we briefly compare observational completeness to other notions of completeness in the literature, such as forwards completeness and fixpoint completeness.

2 Basic Notions of Abstract Interpretation

In the abstract interpretation theory, abstract domains can be equivalently specified either by Galois connections or by upper closure operators (ucos) [4]. When an abstract domain A is specified by a Galois connection, i.e., a pair of abstraction and concretization maps $\langle \alpha, \gamma \rangle$, then $\gamma \circ \alpha \in \mathit{uco}(C)$ is the corresponding uco on C. On the contrary, given an uco ρ, the corresponding Galois connection is $\langle \rho, \mathit{id} \rangle$. In the rest of the paper, we will use ucos, since they are more concise. Moroever, we assume that the concrete domain C is a complete lattice, which is a standard hypothesis in the abstract interpretation theory.

An uco ρ on the concrete domain C is a monotone, idempotent (i.e., $\rho(\rho(x)) = \rho(x)$) and extensive (i.e., $\rho(x) \geq x$) operator on C. Each uco ρ on C is uniquely determined by the set of its fixpoints, which is its image, i.e. $\rho(C) = \{x \in C \mid \rho(x) = x\}$, since $\rho = \lambda x. \bigwedge \{y \in C \mid y \in \rho(C), x \leq y\}$. Moreover, a subset $X \subseteq C$ is the set of fixpoints of an uco on C iff X is meet-closed, i.e. $X = \mathcal{M}(X) = \{\wedge Y \mid Y \subseteq X\}$. For any $X \subseteq C$, $\mathcal{M}(X)$ is called the Moore-closure of X. Often, we will identify closures with their sets of fixpoints. This does not give rise to ambiguity, since one can distinguish their use as functions or sets according to the context. It is well known that the set $\mathit{uco}(C)$ of all ucos on C, endowed with the pointwise ordering $\supseteq$, gives rise to a complete lattice. The top on $\mathit{uco}(C)$ is $\{\top_C\}$, the bottom is C, and the join operation is set intersection $\cap$. The ordering on $\mathit{uco}(C)$ corresponds to the standard order used to compare abstract domains: A_1 is more concrete than A_2 (or A_2 is more abstract than A_1) iff $A_1 \supseteq A_2$ in $\mathit{uco}(C)$.

An abstract domain $\rho \in \mathit{uco}(C)$ is complete for f iff $\rho \circ f = \rho \circ f \circ \rho$ holds. Giacobazzi et al. [8] give a constructive characterization of complete abstract domains, under the assumption of dealing with continuous concrete functions. A function $f : C \to C$ is (Scott) continuous if it preserves least upper bounds of chains in C, i.e., $f(\bigvee B) = \bigvee f(B)$ for any chain $B \subseteq C$. The idea is to build the greatest (i.e., most abstract) domain in $\mathit{uco}(C)$ which includes a given domain ρ and which is complete for a set $F \subseteq C \to C$ of continuous state-transition functions, i.e., for each function in F. In particular, [8] define a mapping $\mathcal{R}_F : \mathit{uco}(C) \mapsto \mathit{uco}(C)$ as follows:

$$\mathcal{R}_F(\rho) = \mathcal{M}\left(\bigcup_{f \in F, a \in \rho} \max(\{x \in C \mid f(x) \leq a\})\right) \ ,$$

where $\max(X)$ is the set of maximal elements in X. They prove that the most abstract domain which includes ρ and is complete for F is $\mathit{gfp}(\lambda\eta.\mathcal{M}(\rho \cup \mathcal{R}_F(\eta)))$. This domain is called the *complete shell* of ρ for F.

3 Observational Completeness

In abstract interpretation, it is common that, in order to observe a property π with a good deal of precision, we need to perform the computation in a richer domain $\rho \supseteq \pi$. In the following, we call π the *observational domain* and ρ the *computational domain*. In the rest of the paper, we assume given a complete lattice C (the concrete domain), a set F of monotone functions from C to C and an uco $\pi \subseteq C$ which represents the set of observable properties.

A common problem is to find a domain ρ such that if we perform any computation on ρ and we project over π, we obtain the same result of the concrete computation, projected over π. In order to formalize this notion, we first need to define the concept of computation (on both an abstract and a concrete domain).

Definition 1 (Computation). *A finite sequence $\xi = \langle f_1, \ldots, f_n \rangle$ of elements of F is called* computation. *Given a computation ξ, a domain α, and an element $c \in C$, we denote by $\xi^\alpha(c)$ the value $(\alpha \circ f_1 \circ \ldots \circ \alpha \circ f_n)(\alpha(c))$. As a special case, when ξ is the empty computation, we define $\xi^\alpha(c) = \alpha(c)$.*

Note that, if *id* is the identity abstraction, then $\xi^{id}(c) = (f_1 \circ \ldots \circ f_n)(c)$. We write $\xi(c)$ as a short form for $\xi^{id}(c)$.

We are now able to compare abstract domains in terms of precision of their computations. We say that a domain α is *more precise than* a domain β if it is the case that, the result of a computation on α projected over π is more precise (it is approximated by) the result of the corresponding computation on β projected over π.

Definition 2 (More Precise than). *We say that α is* more precise than *β for computing F observing π, and we write it as $\alpha \leq \beta$, when*

$$\pi\xi^\alpha(c) \leq \pi\xi^\beta(c)$$

for every computation ξ and $c \in C$.

Although the relation $\leq$ depends on F and π, we prefer to use just $\leq$ instead of a more precise notation such as $\leq_F^\pi$, in order to avoid a cumbersome notation. Since F and π are fixed, this does not cause ambiguities.

It is easy to check that $\leq$ is a preorder, which may be viewed as a generalization of the standard ordering between ucos: if $\alpha \supseteq \beta$ then $\alpha \leq \beta$. Our notion is more general than the standard ordering since it allows us to compare two different domains (α and β) w.r.t. their precision on a third domain (π), and does not require neither α nor β to be in any relation with π. Note that, if $\pi = id$, then $\alpha \leq \beta$ iff $\alpha \supseteq \beta$, since we also consider the empty computation.

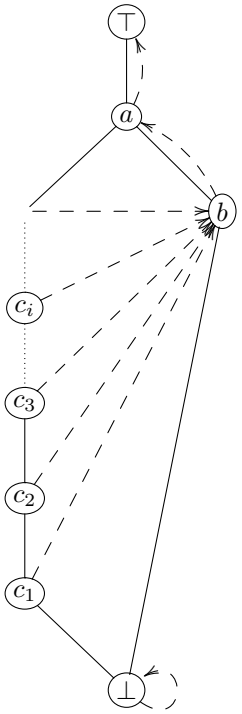
Our formal notion of precision suggests to define a corresponding notion of completeness. We say that a domain α is *observationally complete* for π if any computation on α projected over π, gives the same result of the corresponding concrete computation, projected over π. Here, the key notion is that any computation is always observed on π.

Definition 3 (Observational Completeness). *We say that a domain α is* observationally complete *(for F and π) if α is more precise than the concrete domain, i.e., $\alpha \leq id$.*

Among all the observationally complete domains, we are interested in the least (most abstract) one w.r.t. set inclusion. In general, the least observationally complete domain does not exist, as the following example shows.

Example 1. Let us consider the diagram on the right, where the nodes are the elements of the domain $C = \{\top, \bot, a, b, c_1, c_2, \ldots, c_i, \ldots\}$, solid and dotted edges represent the ordering on C and dashed arrows represent a function $f : C \to C$.

Let $\pi = \{\top, a\}$, $\rho_1 = \{\top, a, b, \bot\} \cup \{c_i \mid i \text{ is even}\}$ and $\rho_2 = \{\top, a, b, \bot\} \cup \{c_i \mid i \text{ is odd}\}$. It is easy to check that both ρ_1 and ρ_2 are observationally complete. However, $\rho = \rho_1 \cap \rho_2 = \{\top, a, b, \bot\}$ is not observationally complete, since, for the computation $\xi = \langle f \rangle$, we have that $\pi(\xi(c_1)) = a$ while $\pi(\xi^{\rho}(c_1)) = \pi(\rho(f(\rho(c_1)))) = \pi(\rho(f(a)) = \top$. □



As a key result we show that, if all the functions in F are continuous, the least observationally complete domain exists.

Theorem 1. *If F is a set of continuous functions, than*

$$\sigma = \mathcal{M}\Big(\bigcup\{\max\{x \in C \mid \xi(x) \leq a\} \mid \xi \text{ computation and } a \in \pi\}\Big)$$

is the least observationally complete domain (for F and π).

In order to exploit this notion of observational completeness for approximating the formal semantics $\mathcal{S}$, we need to show that an observationally complete domain σ preserves the least fixpoint of any composition of functions from F. This result implies that, we can safely approximate the concrete semantic function $\mathcal{F}$ with the abstract semantic function on σ without losing precision on π.

Theorem 2 (Fixpoint Preservation). *Let α be observationally complete for F and π. Then α preserves the least fixpoint of any composition of functions from F, when observing π. In formulas, we have that:*

$$\forall f_1, \ldots f_n \in F,\ \pi(\mathrm{lfp}(f_1 \circ \ldots \circ f_n)) = \pi(\mathrm{lfp}(\alpha \circ f_1 \circ \alpha \circ f_2 \circ \ldots \alpha \circ f_n \circ \alpha)) \ .$$

The previous theorem allows us to say that $\pi(\mathrm{lfp}(\mathcal{F})) = \pi(\mathrm{lfp}(\mathcal{F}^{\#}))$ for any observationally complete domain. In other words, if we only want to observe π, an observationally complete domain does not lose precision in the fixpoint computation involving any composition of functions from F.

4 Observational Completeness and Complete shell

In this section we study the relationships between observational completeness and the standard notion of completeness, in particular between the least observationally complete domain and the complete shell.

It is immediate to show that if α is complete for F and $\alpha \supseteq \pi$, then α is observationally complete for F and π. More generally, α is observationally complete for F and α. We wonder whether:

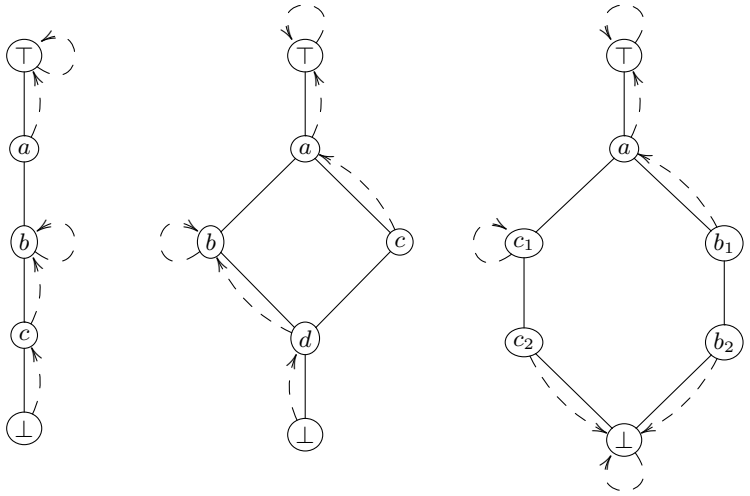
a) every observationally complete domain for F and π is complete for F;
b) the least observationally complete domain for F and π is the complete shell of π for F.

With respect to the first question, note that if α is observationally complete for F and π, every $\beta \supseteq \alpha$ is still observationally complete. This does not hold for completeness: α may be complete for F, although some $\beta \supseteq \alpha$ may not. This is because in the observational completeness the observable properties remain fixed when we refine the initial domain, while for standard completeness the notion of observational domain coincides with the computational domain.

Example 2. Consider the concrete domain $C = \{\top, a, b, c, \bot\}$ depicted in Fig. 1a. The domain $\alpha = \{\top, a, b\}$ is complete for the function depicted in the diagram, hence it is also observationally complete for $\pi = \{\top, a\}$. However, the domain $\{\top, a, b, c\}$ is observationally complete for π but it is not complete. □

Since completeness implies observational completeness, we may argue that the least observationally complete domain for π and F coincides with the complete shell of π. In the general case this is not true and the least observationally complete domain is more abstract than the complete shell. The next examples illustrate this case.

Example 3. Consider the concrete domain $C = \{\top, a, b, c, d, \bot\}$ depicted in Fig. 1b. Assume $\pi = \{\top, a\}$. If we build the complete shell of π for F, in the first step we include the element b and c, since they are the maximal $x \in C$ such that $f(x) \leq a$, and the element d since it is the meet of b and c. At the second step, we also include $\bot$, which is the greatest element $x \in C$ such that $f(x) \leq c$. Note that, in each step, we consider all the elements generated in the previous steps, forgetting the observational domain π which started the process. However, it is trivial to check that the domain $\alpha = C \setminus \{\bot\}$ has the same precision of C when observing π, i.e. $\pi f^i(x) = \pi(\alpha f \alpha)^i(x)$ for every $x \in C$ and $i \in \mathbb{N}$. When $x \in \alpha$, the stronger property $f^i(x) = (\alpha f \alpha)^i(x)$ holds. When $x = \bot$, it is not true that $f^i(\bot) = (\alpha f \alpha)^i(\bot)$: for example it does not hold for $i = 1$. However, for each i, $\pi f^i(\bot) = a = \pi(\alpha f \alpha)^i(\bot)$, hence α is observationally complete. □



(a) Example 2 (b) Example 3 (c) Example 4

Fig. 1: Counterexamples

Example 4. Consider the concrete domain $C = \{\top, a, b_1, b_2, c_1, c_2, \bot\}$ depicted in Fig. 1c. If $\pi = \{\top, a\}$, the complete shell is the entire domain C. However, c_2 is useless when observing π, since the least observationally complete domain is $C \setminus \{c_2\}$. □

4.1 The Case of Additive Functions

We will show that, when all the functions in F are (completely) additive, the least observationally complete domain and the complete shell coincide. However, in order to prove this result, we need to give an alternative construction for the complete shell, more similar to the construction for the least observationally complete domain. In more details, we replace the standard refinement operator $\mathcal{R}_F : uco(C) \mapsto uco(C)$ given in Sect. 2 with a new operator $\hat{\mathcal{R}}_F : \wp(C) \mapsto \wp(C)$ simply obtained by removing the Moore closure from the definition of $\mathcal{R}_F$. Therefore, we define

$$\hat{\mathcal{R}}_F(X) = \bigcup_{f \in F, a \in X} \max\{x \in C \mid f(x) \leq a\} \ .$$

Note that $\hat{\mathcal{R}}_F(X)$ may not be an uco even if X is an uco, and that $\mathcal{R}_F(X) = \mathcal{M}\left(\hat{\mathcal{R}}_F(X)\right)$. We recall that, given a function $G : \wp(C) \to \wp(C)$, we have that $G^\omega(X) = \cup_{i \in \mathbb{N}} G^i(X)$ where $G^0(X) = X$ and $G^{i+1}(X) = G(G^i(X))$.

Theorem 3. *For every set F of continuous maps, the complete shell of π for F is given by*

$$S = \mathcal{M}\left(\hat{\mathcal{R}}_F^\omega(\pi)\right) \ .$$

This new construction, which is the key result to prove Theorem 4, is interesting in itself, since it sheds a new light on the construction of the complete shell. First of all, it shows that it is not necessary to compute the Moore closure at each step of the refinement, but it suffices to compute it at the end. Secondly, it shows that we need at most ω steps of refinement to reach the fixpoint.

We recall that a function $f : C \to C$ is (completely) additive if it preserves arbitrary least upper bounds, i.e., $f(\bigvee B) = \bigvee f(B)$ for any $B \subseteq C$.

Theorem 4. *If F is a set of completely additive functions, the complete shell S of π for F is the smallest observationally complete domain σ for F and π.*

It is worth noting that, even if F is a set of additive functions, this theorem does not imply that observational completeness and completeness are the same thing: in Example 2 the function f is additive, yet there is an observationally complete domain which is is not complete.

5 Conclusions and Related Work

Different kinds of completeness have been proposed in the literature. The first notion of completeness appears in Cousot and Cousot [4]. In the same paper, the notion of *fixpoint completeness* is formalized. A domain α is fixpoint complete for a function f when it preserves the least fixpoint of f, in formulas $\alpha(\mathrm{lfp}\, f) = \mathrm{lfp}(\alpha \circ f \circ \alpha)$. Cousot and Cousot have shown that complete domains are also fixpoint complete. A detailed study on completeness and fixpoint completeness can be found in Giacobazzi et al. [8], where the authors solve the problem of synthesizing complete abstract domains.

Cousot and Cousot [2] introduced a different notion of completeness called *exactness*. The same notion has been renamed as *forward completeness* (F-completeness) by Giacobazzi and Quintarelli [7] who apply the completeness results on model checking. Moreover, to distinguish between standard completeness and F-completeness, Giacobazzi and Quintarelli renamed the former as *backward completeness* (B-completeness). A domain α is F-complete for a function f when $f \circ \alpha = \alpha \circ f \circ \alpha$. Intuitively, this means that the result of any abstract computation coincides with the result of the corresponding concrete one, when the starting object is an abstract object.

Our notion of observational completeness differs from all the previous notions (B-completeness, fixpoint completeness, F-completeness). The main point is that, in our notion, we have two concepts of observational and computational domain and, most importantly, the observational domain is kept fixed when refining. We believe that, in any static analysis or semantics definition, the observable property does not change when looking for better domains. On the contrary, B-completeness is self-referential, since the observational domain changes when refining the domain. More precisely, given a domain π, the complete shell of π for f is the least abstract domain β containing π which is observationally complete for β (and thus it is observationally complete for π). Moreover, the self-referentiality of completeness yields some counter-intuitive behaviors. For

instance, if α is complete and β contains α, it may well happen that β is not complete even if, according to our intuition, β is "richer" than α. This does not happen for observational completeness, where supersets of observationally complete domains are still observationally complete (see Example 2).

The notion of F-completeness does not fix any observable property. This kind of completeness is useful when we are interested in a subset of the concrete domain closed for the application of any state-transition function.

Finally, fixpoint completeness does not take into consideration intermediates steps during the abstract computation. In fact, it is only required that the abstract least fixpoint (computed on the abstract domain) coincides with the abstraction of the concrete least fixpoint. Giacobazzi et al. [8] show that, even under strong hypotheses, the existence of the least fixpoint complete domain containing π cannot be ensured. They show that, even if the concrete domain is a complete Boolean algebra or a finite chain, and the concrete function f is both additive and co-additive, the least fixpoint complete domain containing π does not necessarily exist. The counterexamples suggest that finding reasonable conditions for the existence of least fixpoint complete domains is not viable.

Other notions of completeness have been proposed for dealing with logic (see, for instance, Cousot and Cousot [5], Schmidt [9], Dams et al. [6]). In general, completeness problems on fragments of temporal logic are considered (covering, preservation, strong preservation). All these notions are very different from the other forms of completeness, since they consider only fixed logical operators, and, in general, one is not interested in least fixpoint preservation.

As a future work, we think that observational completeness could be generalized, in order to be relative to an abstract domain, instead of the concrete one. We say that a domain α is observationally complete for π and F relatively to the domain β, when the result of any abstract computation, observed over π, is more precise than the corresponding abstract computation on the domain β, observed over π. Here, the novelty is that the domains α, β and π do not need to be in any relation. Thus, the least observationally complete domain for π and F relatively to β could be incomparable with β.

Observational completeness naturally arises once we fix the preorder $\leq$ on domains, which formalizes the intuitive notion of precision. The novelty with respect to the standard treatment of completeness is that we have two orderings on ucos: standard inclusion $\supseteq$ and precision $\leq$. The latter is used to define what an observationally complete domain is, while the former selects, among observationally complete domains, the preferred one. In the complete shell construction the two orderings coincide. In principle, we could change the standard inclusion ordering, obtaining a different notion of "least" observationally complete domain. For instance, we could compare two abstract domains on the base of their cardinality or of a suitable notion of "complexity" of their abstract objects.

References

1. G. Birkhoff. *Lattice Theory*, volume XXV of *AMS Colloquium Publications*. American Mathematical Society, third edition, 1967.

2. P. Cousot. Types as abstract interpretations, invited paper. In *POPL '97: Proceedings of the 24th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*, pages 316–331. ACM Press, New York, NY, USA, Jan. 1997.
3. P. Cousot and R. Cousot. Abstract interpretation: A unified lattice model for static analysis of programs by construction or approximation of fixpoints. In *POPL '77: Proceedings of the 4th ACM SIGACT-SIGPLAN symposium on Principles of programming languages*, pages 238–252. ACM Press, New York, NY, USA, Jan. 1977.
4. P. Cousot and R. Cousot. Systematic design of program analysis frameworks. In *POPL '79: Proceedings of the 6th ACM SIGACT-SIGPLAN symposium on Principles of programming languages*, pages 269–282. ACM Press, New York, NY, USA, Jan. 1979.
5. P. Cousot and R. Cousot. Temporal abstract interpretation. In *POPL '00: Proceedings of the 27th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*, pages 12–25. ACM Press, New York, NY, USA, Jan. 2000.
6. D. Dams, R. Gerth, and O. Grumberg. Abstract interpretation of reactive systems. *ACM Transactions on Programming Languages and Systems*, 19(2):253–291, 1997.
7. R. Giacobazzi and E. Quintarelli. Incompleteness, counterexamples, and refinements in abstract model-checking. In P. Cousot, editor, *Static Analysis, 8th International Symposium, SAS 2001 Paris, France, July 1618, 2001 Proceedings*, volume 2126 of *Lecture Notes in Computer Science*, pages 356–373. Springer, Berlin Heidelberg, 2001.
8. R. Giacobazzi, F. Ranzato, and F. Scozzari. Making abstract interpretations complete. *Journal of the ACM*, 47(2):361–416, 2000.
9. D. A. Schmidt. Comparing completeness properties of static analyses and their logics. In N. Kobayashi, editor, *Programming Languages and Systems, 4th Asian Symposium, APLAS 2006, Sydney, Australia, November 8-10, 2006. Proceedings*, volume 4279 of *Lecture Notes in Computer Science*, pages 183–199. Springer, Berlin Heidelberg, 2006.

A Appendix

In some proofs we will make use of the Hausdorff's maximality principle [1]. We recall that a chain Y in a poset P is maximal (with respect to set inclusion) whenever for any other chain Y' in P, $Y \subseteq Y'$ implies $Y = Y'$. The Hausdorff's maximal principle says that every chain in a poset P can be extended to a maximal chain in P.

Theorem 1. *If F is a set of continuous functions, than*

$$\sigma = \mathcal{M}\Big(\bigcup\{\max\{x \in C \mid \xi(x) \leq a\} \mid \xi \text{ computation and } a \in \pi\}\Big)$$

is the least observationally complete domain (for F and π).

Proof. First of all, we show that σ is observationally complete. We prove, by induction on the length of ξ, that $\xi(x) \leq a$ implies $\xi^\sigma(x) \leq a$ for each computation ξ and $a \in \pi$. If $|\xi| = 0$, then $\xi(x) = x$ and $\xi^\sigma(x) = \sigma(x)$. Note that $\sigma \supseteq \pi$ since if $a \in \pi$ then $a = \bigvee\{x \in C \mid x \leq a\}$. Hence $x \leq a$ implies $\sigma(x) \leq a$. Now assume

$|\xi| = i+1$. If $\xi(x) \leq a$, consider the poset $C' = \{c \in C \mid \xi(c) \leq a\}$ and the chain $\{x\} \subseteq C'$. By Hausdorff's maximality principle there exists a maximal chain $Y \supseteq \{x\}$ which is contained in C'. Let $y = \bigvee Y$. By continuity of ξ, we have that $\xi(y) \leq a$. Since Y is a maximal chain in C', then $y \in \max C'$. Moreover, by definition of σ we have that $y \in \sigma$. It follows that $\xi(\sigma(x)) \leq \xi(y) \leq a$. If $\xi = \xi_1 \cdot f$, by inductive hypothesis $\xi^\sigma(x) = \xi_1^\sigma(f(\sigma(x))) \leq a$ since $\xi_1(f(\sigma(x))) \leq a$.

Now we show that σ is the least observationally complete domain. Assume, without loss of generality, that $v \in \max\{x \in C \mid \xi(x) \leq a\}$ for some computation ξ and $a \in \pi$ and let ρ be a domain such that $v \notin \rho$. Then, $\xi(\rho(v)) \not\leq a$ since $\rho(v) > v$ and by definition of v. Hence, also $\xi^\rho(v) \not\leq a$, which means ρ is not observationally complete. □

Theorem 2 (Fixpoint Preservation). *Let α be observationally complete for F and π. Then α preserves the least fixpoint of any composition of functions from F, when observing on π. In formulas, we have that:*

$$\forall f_1, \ldots f_n \in F,\ \pi(\mathrm{lfp}(f_1 \circ \ldots \circ f_n)) = \pi(\mathrm{lfp}(\alpha \circ f_1 \circ \alpha \circ f_2 \circ \ldots \alpha \circ f_n \circ \alpha)) \ .$$

Proof. It clearly holds that

$$\pi(\mathrm{lfp}(f_1 \circ \ldots \circ f_n)) \leq \pi(\mathrm{lfp}(\alpha \circ f_1 \circ \alpha \circ f_2 \circ \ldots \alpha \circ f_n \circ \alpha)) \ ,$$

since α is extensive. We now show the other direction. We prove that, for any $c \in C$ and ordinal ϵ, it holds that

$$\pi\left(\textstyle\bigvee_{i\in\epsilon}(\alpha \circ f_1 \circ \alpha \circ f_2 \circ \ldots \alpha \circ f_n \circ \alpha)^i(c)\right) \leq \pi\left(\textstyle\bigvee_{i\in\epsilon}(f_1 \circ f_2 \circ \ldots \circ f_n)^i(c)\right) \ .$$

Since π is an upper closure operator, it is complete for arbitrary lubs. It follows that:

$$\pi\left(\textstyle\bigvee_{i\in\epsilon}(\alpha \circ f_1 \circ \alpha \circ f_2 \circ \ldots \alpha \circ f_n \circ \alpha)^i(c)\right) = \pi\left(\textstyle\bigvee_{i\in\epsilon}\pi(\alpha \circ f_1 \circ \alpha \circ f_2 \circ \ldots \alpha \circ f_n \circ \alpha)^i(c)\right) \ .$$

Since α is observationally complete for F and π, then

$$\pi\left(\textstyle\bigvee_{i\in\epsilon}\pi(\alpha \circ f_1 \circ \alpha \circ f_2 \circ \ldots \alpha \circ f_n \circ \alpha)^i(c)\right) = \pi\left(\textstyle\bigvee_{i\in\epsilon}\pi(f_1 \circ f_2 \circ \ldots \circ f_n)^i(c)\right) \ ,$$

which is equivalent to $\pi\left(\bigvee_{i\in\epsilon}(f_1 \circ f_2 \circ \ldots \circ f_n)^i(c)\right)$. □

Lemma 1. *For every set F of continuous functions and every $X \subseteq C$, we have*

$$\mathcal{M}\left(\hat{\mathcal{R}}_F(\mathcal{M}(X))\right) = \mathcal{M}\left(\hat{\mathcal{R}}_F(X)\right) \ .$$

Proof. It is immediate by monotonicity of $\hat{\mathcal{R}}_F$ and $\mathcal{M}(\_)$ that $\mathcal{M}\left(\hat{\mathcal{R}}_F(\mathcal{M}(X))\right) \supseteq \mathcal{M}\left(\hat{\mathcal{R}}_F(X)\right)$. For the converse inequality, since $\mathcal{M}(\_)$ is an upper closure

operator on $\wp(C)$, it is enough to prove that $\hat{\mathcal{R}}_F(\mathcal{M}(X)) \subseteq \mathcal{M}\left(\hat{\mathcal{R}}_F(X)\right)$. Given $y \in \hat{\mathcal{R}}_F(\mathcal{M}(X))$, we have $y \in \max\{x \in C \mid f(x) \leq a\}$ and $a = \bigwedge_{i\in I} a_i$ where $f \in F$ and $\{a_i\}_{i\in I} \subseteq X$.

For each $i \in I$, consider the set $Y_i = \max\{x \in C \mid f(x) \leq a_i\} \subseteq \hat{\mathcal{R}}_F(X)$. Since $f(y) \leq a \leq a_i$ and f is continuous, there is an $y_i \in Y_i$ such that $y_i \geq y$. We may find y_i as the least upper bound of a maximal chain in Y_i containing y, which exists by Hausdorff's maximality principle. It is enough to prove that $y = \bigwedge_{i\in I} y_i$. By definition of the y_i's, we have $y \leq \bigwedge_{i\in I} y_i$. Moreover, $f(\bigwedge_{i\in I} y_i) \leq f(y_i) \leq a_i$ hence $f(\bigwedge_{i\in I} y_i) \leq a$. Since y is a maximal element such that $f(y) \leq a$, this means that $y = \bigwedge_{i\in I} y_i$. □

Theorem 3. *For every set F of continuous maps, the complete shell S of π for F is given by*

$$S = \mathcal{M}\left(\hat{\mathcal{R}}_F^{\omega}(\pi)\right) .$$

Proof. It can be easily proved that

$$S = \mathcal{M}\left(G^{\kappa}(\{\top\})\right)$$

for some ordinal κ, where $G : uco(C) \mapsto uco(C)$ is the map

$$G(\rho) = \mathcal{M}\left(\pi \cup \mathcal{R}_F(\rho)\right) = \mathcal{M}\left(\pi \cup \hat{\mathcal{R}}_F(\rho)\right) .$$

It is enough to prove that $\mathcal{M}\left(\hat{\mathcal{R}}_F^{\omega}(\pi)\right)$ is a subset of S and a fixpoint of G. In order to prove $\mathcal{M}\left(\hat{\mathcal{R}}_F^{\omega}(\pi)\right) \subseteq S$ it is enough to show that $\mathcal{R}_F^i(\pi) \subseteq G^{i+1}(\{\top\})$ for every $i < \omega$. The proof is by induction over i. For $i = 0$, we have $\hat{\mathcal{R}}_F^0(\pi) = \pi \subseteq G(\{\top\})$. If $i = j + 1$, $\hat{\mathcal{R}}_F^i(\pi) = \hat{\mathcal{R}}_F(\hat{\mathcal{R}}_F^j(\pi)) \subseteq \hat{\mathcal{R}}_F(G^i(\{\top\})) \subseteq G(G^i(\{\top\})) = G^{i+1}(\{\top\})$. Now we prove that $\mathcal{M}\left(\hat{\mathcal{R}}_F^{\omega}(\pi)\right)$ is a fixpoint of G. We have that

$$\begin{aligned}
G\left(\mathcal{M}\left(\hat{\mathcal{R}}_F^{\omega}(\pi)\right)\right) & \\
&= \mathcal{M}\left(\pi \cup \hat{\mathcal{R}}_F\left(\mathcal{M}\left(\hat{\mathcal{R}}_F^{\omega}(\pi)\right)\right)\right) \\
&= \mathcal{M}\left(\pi \cup \mathcal{M}\left(\hat{\mathcal{R}}_F\left(\mathcal{M}\left(\hat{\mathcal{R}}_F^{\omega}(\pi)\right)\right)\right)\right) \\
&= \mathcal{M}\left(\pi \cup \mathcal{M}\left(\hat{\mathcal{R}}_F(\hat{\mathcal{R}}_F^{\omega}(\pi))\right)\right) \\
&= \mathcal{M}\left(\pi \cup \bigcup\{\hat{\mathcal{R}}_F^{i+1}(\pi) \mid i < \omega\}\right) \\
&= \mathcal{M}\left(\hat{\mathcal{R}}_F^{\omega}(\pi)\right) .
\end{aligned}$$

This concludes the proof. □

Theorem 4. *If F is a set of completely additive functions, the complete shell S of π for F is the least observationally complete domain σ for F and π.*

Proof. We know that S is observationally complete, since it is complete and contains π. It is enough to prove that if $a \in \hat{\mathcal{R}}^{\omega}_F(\pi)$ and $a \notin \rho$, then ρ is not observationally complete for π and F. Assume $a \in \hat{\mathcal{R}}^{i}_F(\pi)$ and there is no $j < i$ such that $a \in \hat{\mathcal{R}}^{i}_F(\pi)$. It means there exist a computation $\xi = \langle f_1, \dots, f_i \rangle$ of maps in F and a sequence $a_0, \dots, a_i$ of objects in C such that $a = a_i$, $a_0 \in \pi$ and $a_j \in \max\{x \in C \mid f_j(x) \leq a_{j-1}\}$ for any $j \in [1, \dots, i]$. It immediate to check that $\xi(a) \leq a_0$. We prove that $\xi^{\rho}(a) \not\leq a_0$.

Note that, if f is completely additive, then $\max\{x \in C \mid f(x) \leq y\}$ is a singleton for any $y \in C$. Therefore, if $\max\{x \in C \mid f(x) \leq y\} = \{z\}$ and $x \not\leq z$, then $f(x) \not\leq y$. In our proof, this means that, for each $j \in [1, \dots, i]$, $x \not\leq a_j$ implies $f_j(x) \not\leq a_{j-1}$. Since $\rho(f(x)) \geq f(x)$, this also implies $\rho(f_j(a_j)) \not\leq a_{j-1}$. Since $a \notin \rho$, then $\rho(a) > a$, i.e. $\rho(a) \not\leq a$, hence $\xi^{\rho}(a) \not\leq a_0$. □