

## FORWARD ANALYSIS FOR WSTS, PART I: COMPLETIONS

ALAIN FINKEL<sup>1</sup> AND JEAN GOUBAULT-LARRECQ<sup>1,2</sup>

<sup>1</sup> LSV, ENS Cachan, CNRS; 61 avenue du président Wilson, F-94230 Cachan

<sup>2</sup> INRIA Saclay Ile-de-France

*E-mail address:* {finkel, goubault}@lsv.ens-cachan.fr

**ABSTRACT.** Well-structured transition systems provide the right foundation to compute a finite basis of the set of predecessors of the upward closure of a state. The dual problem, to compute a finite representation of the set of successors of the downward closure of a state, is harder: Until now, the theoretical framework for manipulating downward-closed sets was missing. We answer this problem, using insights from domain theory (dcpos and ideal completions), from topology (sobrifications), and shed new light on the notion of adequate domains of limits.

### 1. Introduction

The theory of well-structured transition systems (WSTS) is 20 years old [9, 11, 2]. The most often used result of this theory [11] is the backward algorithm for computing a finite basis of the set  $\uparrow Pre^*(\uparrow s)$  of predecessors of the upward closure  $\uparrow s$  of a state  $s$ . The starting point of this paper is our desire to compute  $\downarrow Post^*(\downarrow s)$  in a similar way. We then need a theory to finitely (and effectively) represent downward-closed sets, much as upward-closed subsets can be represented by their finite sets of minimal elements. This will serve as a basis for constructing forward procedures.

The cover,  $\downarrow Post^*(\downarrow s)$ , contains more information than the set of predecessors  $\uparrow Pre^*(\uparrow s)$  because it characterizes a good approximation of the reachability set, while the set of predecessors describes the states from which the system may fail; the cover may also allow the computation of a finite-state abstraction of the system as a symbolic graph. Moreover, the backward algorithm needs a finite basis of the upward closed set of bad states, and its implementation is, in general, less efficient than a forward procedure: e.g., for lossy channel systems, although the backward procedure always terminates, only the non-terminating forward procedure is implemented in the tool TREX [1].

Except for some partial results [9, 7, 13], a general theory of downward-closed sets is missing. This may explain the scarcity of forward algorithms for WSTS. Quoting Abdulla *et al.* [3]: “Finally, we aim at developing generic methods for building downward closed languages, in a similar manner to the methods we have developed for building upward closed languages in [2]. This would give a general theory for forward analysis of infinite state systems, in the same way the work in [2] is for backward analysis.” Our contribution is to provide such a theory of downward-closed sets.

*Key words and phrases:* WSTS, forward analysis, completion, Karp-Miller procedure, domain theory, sober spaces, Noetherian spaces.



*Related Work.* Karp and Miller [16] proposed an algorithm that computes a finite representation of the downward closure of the reachability set of a Petri net. Finkel [9] introduced the WSTS framework and generalized the Karp-Miller procedure to a class of WSTS. This is done by constructing the completion of the set of states (by ideals, see Section 3) and in replacing the  $\omega$ -acceleration of an increasing sequence of states (in Petri nets) by its least upper bound (lub). However, there are no effective finite representations of downward closed sets in [9]. Emerson and Namjoshi [7] considered a variant of WSTS (using cpos, but still without a theory of effective finite representations of downward-closed subsets) for defining a Karp-Miller procedure to broadcast protocols—termination is then not guaranteed [8]. Abdulla *et al.* [1] proposed a forward procedure for lossy channel systems using downward-closed languages, coded as SREs. Ganty, Geeraerts, and others [13, 12] proposed a forward procedure for solving the coverability problem for WSTS equipped with an effective adequate domain of limits. This domain ensures that every downward closed set has a finite representation; but no insight is given how these domains can be found or constructed. They applied this to Petri nets and lossy channel systems. Abdulla *et al.* [3] proposed another symbolic framework for dealing with downward closed sets for timed Petri nets.

We shall see that these constructions are special cases of our completions (Section 3). We shall illustrate this in Section 4, and generalize to a comprehensive hierarchy of data types in Section 5. We briefly touch the question of computing approximations of the cover in Section 6, although we shall postpone most of it to future work. We conclude in Section 7.

## 2. Preliminaries

We shall borrow from theories of order, both from the theory of well quasi-orderings, as used classically in well-structured transition systems [2, 11], and from domain theory [5, 14]. We should warn the reader that this is one bulky section on preliminaries. We invite her to skip technical points first, returning to them on demand.

A *quasi-ordering*  $\leq$  is a reflexive and transitive relation on a set  $X$ . It is a (partial) *ordering* iff it is antisymmetric. A set  $X$  equipped with a partial ordering is a *poset*.

We write  $\geq$  the converse quasi-ordering,  $\approx$  the equivalence relation  $\leq \cap \geq$ ,  $<$  associated strict ordering ( $\leq \setminus \approx$ ), and  $>$  the converse ( $\geq \setminus \approx$ ) of  $<$ . The *upward closure*  $\uparrow E$  of a set  $E$  is  $\{y \in X \mid \exists x \in E \cdot x \leq y\}$ . The *downward closure*  $\downarrow E$  is  $\{y \in X \mid \exists x \in E \cdot y \leq x\}$ . A subset  $E$  of  $X$  is *upward closed* if and only if  $E = \uparrow E$ , i.e., any element greater than or equal to some element in  $E$  is again in  $E$ . *Downward closed* sets are defined similarly. When the ambient space  $X$  is not clear from context, we shall write  $\downarrow_X E$ ,  $\uparrow_X E$  instead of  $\downarrow E$ ,  $\uparrow E$ .

A quasi-ordering is *well-founded* iff it has no infinite strictly descending chain, i.e.,  $x_0 > x_1 > \dots > x_i > \dots$ . An *antichain* is a set of pairwise incomparable elements. A quasi-ordering is *well* if and only if it is well-founded and has no infinite antichain.

There are a number of equivalent definitions for well quasi-orderings (wqo). One is that, from any infinite sequence  $x_0, x_1, \dots, x_i, \dots$ , one can extract an infinite ascending chain  $x_{i_0} \leq x_{i_1} \leq \dots \leq x_{i_k} \leq \dots$ , with  $i_0 < i_1 < \dots < i_k < \dots$ . Another one is that any upward closed subset can be written  $\uparrow E$ , with  $E$  finite. Yet another, topological definition [15, Proposition 3.1] is to say that  $X$ , with its Alexandroff topology, is Noetherian. The *Alexandroff topology* on  $X$  is that whose opens are exactly the upward closed subsets. A subset  $K$  is compact if it satisfies the Heine-Borel property, i.e., every one may extract a finite subcover from any open cover of  $K$ . A topology is *Noetherian* iff every open subset is compact, iff any increasing chain of opens stabilizes [15, Proposition 3.2]. We shall cite results from the latter paper as the need evolves.

We shall be interested in rather particular topological spaces, whose topology arises from order. A *directed family* of  $X$  is any non-empty family  $(x_i)_{i \in I}$  such that, for all  $i, j \in I$ , there is a  $k \in I$  with  $x_i, x_j \leq x_k$ . The *Scott topology* on  $X$  has as opens all upward closed subsets  $U$  such that every directed family  $(x_i)_{i \in I}$  that has a least upper bound  $x$  in  $X$  intersects  $U$ , i.e.,  $x_i \in U$  for some  $i \in I$ . The Scott topology is coarser than the Alexandroff topology, i.e., every Scott-open is Alexandroff-open (upward closed); the converse fails in general. The Scott topology is particularly interesting on *dcpos*, i.e., posets  $X$  in which every directed family  $(x_i)_{i \in I}$  has a least upper bound  $\sup_{i \in I} x_i$ .

The *way below* relation  $\ll$  on a poset  $X$  is defined by  $x \ll y$  iff, for every directed family  $(z_i)_{i \in I}$  that has a least upper bound  $z \geq y$ , then  $z_i \geq x$  for some  $i \in I$  already. Note that  $x \ll y$  implies  $x \leq y$ , and that  $x' \leq x \ll y \leq y'$  implies  $x' \ll y'$ . However,  $\ll$  is not reflexive or irreflexive in general. Write  $\uparrow E = \{y \in X \mid \exists x \in E \cdot x \ll y\}$ ,  $\downarrow E = \{y \in X \mid \exists x \in E \cdot y \ll x\}$ .  $X$  is *continuous* iff, for every  $x \in X$ ,  $\downarrow x$  is a directed family, and has  $x$  as least upper bound. One may be more precise: A *basis* is a subset  $B$  of  $X$  such that any element  $x \in X$  is the least upper bound of a directed family of elements way below  $x$  in  $B$ . Then  $X$  is continuous if and only if it has a basis, and in this case  $X$  itself is the largest basis. In a continuous dcpo,  $\uparrow x$  is Scott-open for all  $x$ , and every Scott-open set  $U$  is a union of such sets, viz.  $U = \bigcup_{x \in U} \uparrow x$  [5].

$X$  is *algebraic* iff every element  $x$  is the least upper bound of the set of finite elements below  $x$ —an element  $y$  is *finite* if and only if  $y \ll y$ . Every algebraic poset is continuous, and has a least basis, namely its set of finite elements.

$\mathbb{N}$ , with its natural ordering, is a wqo and an algebraic poset. All its elements are finite, so  $x \ll y$  iff  $x \leq y$ .  $\mathbb{N}$  is not a dcpo, since  $\mathbb{N}$  itself is a directed family without a least upper bound. Any finite product of continuous posets (resp., continuous dcpos) is again continuous, and the Scott-topology on the product coincides with the product topology. Any finite product of wqos is a wqo. In particular,  $\mathbb{N}^k$ , for any integer  $k$ , is a wqo and a continuous poset: this is the set of configurations of Petri nets.

It is clear how to complete  $\mathbb{N}$  to make it a cpo: let  $\mathbb{N}_\omega$  be  $\mathbb{N}$  with a new element  $\omega$  such that  $n \leq \omega$  for all  $n \in \mathbb{N}$ . Then  $\mathbb{N}_\omega$  is still a wqo, and a continuous cpo, with  $x \ll y$  if and only if  $x \in \mathbb{N}$  and  $x \leq y$ . In general, completing a wqo is necessary to extend coverability tree techniques [9, 13]. Geeraerts *et al.* (op. cit.) axiomatize the kind of completions they need in the form of so-called *adequate domains of limits*. We discuss them in Section 3. For now, let us note that the second author also proposed to use another notion of completion in another context, known as *sobriification* [15]. We need to recap what this is about.

A topological space  $X$  is always equipped with a *specialization quasi-ordering*, which we shall write  $\leq$  again:  $x \leq y$  if and only if any open subset containing  $x$  also contains  $y$ .  $X$  is  $T_0$  if and only if  $\leq$  is a partial ordering. Given any quasi-ordering  $\leq$  on a set  $X$ , both the Alexandroff and the Scott topologies admit  $\leq$  as specialization quasi-ordering. In fact, the Alexandroff topology is the finest (the one with the most opens) having this property. The coarsest is called the *upper topology*; its opens are arbitrary unions of complements of sets of the form  $\downarrow E$ ,  $E$  finite. The latter sets  $\downarrow E$ , with  $E$  finite, will play an important role, and we call them the *finitary closed* subsets. Note that finitary closed subsets are closed in the upper, Scott, and Alexandroff topologies, recalling that a subset is *closed* iff its complement is open. The *closure*  $cl(A)$  of a subset  $A$  of  $X$  is the smallest closed subset containing  $A$ . A closed subset  $F$  is *irreducible* if and only if  $F$  is non-empty, and whenever  $F \subseteq F_1 \cup F_2$  with  $F_1, F_2$  closed, then  $F \subseteq F_1$  or  $F \subseteq F_2$ . The finitary closed subset  $\downarrow x = cl(\{x\})$  ( $x \in X$ ) is always irreducible. A space  $X$  is *sober* iff every irreducible closed subset  $F$  is the closure of a unique point, i.e.,  $F = \downarrow x$  for some unique  $x$ . Any sober space is  $T_0$ , and any continuous cpo is sober in its Scott topology. Conversely, given a  $T_0$  space  $X$ , the space  $\mathcal{S}(X)$

of all irreducible closed subsets of  $X$ , equipped with upper topology of the inclusion ordering  $\subseteq$ , is always sober, and the map  $\eta_S : x \mapsto \uparrow x$  is a topological embedding of  $X$  inside  $\mathcal{S}(X)$ .  $\mathcal{S}(X)$  is the *sobrification* of  $X$ , and can be thought as  $X$  together with all missing limits from  $X$ . Note in particular that a sober space is always a cpo in its specialization ordering [5, Proposition 7.2.13].

It is an enlightening exercise to check that  $\mathcal{S}(\mathbb{N})$  is  $\mathbb{N}_\omega$ . Also, the topology on  $\mathcal{S}(\mathbb{N})$  (the upper topology) coincides with that of  $\mathbb{N}_\omega$  (the Scott topology). In general,  $X$  is Noetherian if and only if  $\mathcal{S}(X)$  is Noetherian [15, Proposition 6.2], however the upper and Scott topologies do not always coincide [15, Section 7]. In case of ambiguity, given any poset  $X$ , we write  $X_a$  the space  $X$  with its Alexandroff topology.

Another important construction is the *Hoare powerdomain*  $\mathcal{H}(X)$  of  $X$ , whose elements are the closed subsets of  $X$ , ordered by inclusion. (We do allow the empty set.) We again equip it with the corresponding upper topology.

### 3. Completions of Wqos

One of the central problems of our study is the definition of a *completion* of a wqo  $X$ , with all missing limits added. Typically, the Karp-Miller construction [16] works not with  $\mathbb{N}^k$ , but with  $\mathbb{N}_\omega^k$ . We examine several ways to achieve this, and argue that they are the same, up to some details.

*ADLs, WADLs.* We start with Geeraerts *et al.*'s axiomatization of so-called *adequate domain of limits* for well-quasi-ordered sets  $X$  [13]. No explicit constructions for such adequate domains of limits is given, and they have to be found by trial and error. Our main result, below, is that there is a unique least adequate domain of limits: the *sobrification*  $\mathcal{S}(X_a)$  of  $X_a$ . (Recall that  $X_a$  is  $X$  with its Alexandroff topology.) This not only gives a concrete construction of such an adequate domain of limits, but also shows that we do not have much freedom in defining one.

An *adequate domain of limits* [13] (ADL) for a well-ordered set  $X$  is a triple  $(L, \preceq, \gamma)$  where  $L$  is a set disjoint from  $X$  (the set of *limits*); (L<sub>1</sub>) the map  $\gamma : L \cup X \rightarrow \mathbb{P}(X)$  is such that  $\gamma(z)$  is downward closed for all  $z \in L \cup X$ , and  $\gamma(x) = \downarrow_X x$  for all non-limit points  $x \in X$ ; (L<sub>2</sub>) there is a limit point  $\top \in L$  such that  $\gamma(\top) = X$ ; (L<sub>3</sub>)  $z \preceq z'$  if and only if  $\gamma(z) \subseteq \gamma(z')$ ; and (L<sub>4</sub>) for any downward closed subset  $D$  of  $X$ , there is a finite subset  $E \subseteq L \cup X$  such that  $\hat{\gamma}(E) = D$ . Here  $\hat{\gamma}(E) = \bigcup_{z \in E} \gamma(z)$ .

Requirement (L<sub>2</sub>) in [13] only serves to ensure that all closed subsets of  $L \cup X$  can be represented as  $\downarrow_{L \cup X} E$  for some finite subset  $E$ : the closed subset  $L \cup X$  itself is then exactly  $\downarrow_{L \cup X} \{\top\}$ . However, (L<sub>2</sub>) is unnecessary for this, since  $L \cup X$  already equals  $\downarrow_{L \cup X} E$  by (L<sub>3</sub>), where  $E$  is the finite subset of  $L \cup X$  such that  $\hat{\gamma}(E) = L \cup X$  as ensured by (L<sub>4</sub>). Accordingly, we drop requirement (L<sub>2</sub>):

**Definition 3.1** (WADL). Let  $X$  be a poset. A *weak adequate domain of limits* (WADL) on  $X$  is any triple  $(L, \preceq, \gamma)$  satisfying (L<sub>1</sub>), (L<sub>3</sub>), and (L<sub>4</sub>).

**Proposition 3.2.** Let  $X$  be a poset. Given a WADL  $(L, \preceq, \gamma)$  on  $X$ ,  $\gamma$  defines an order-isomorphism from  $(L \cup X, \preceq)$  to some subset of  $\mathcal{H}(X_a)$  containing  $\mathcal{S}(X_a)$ .

Conversely, assume  $X$  wqo, and let  $Y$  be any subset of  $\mathcal{H}(X_a)$  containing  $\mathcal{S}(X_a)$ . Then  $(Y \setminus \eta_S(X_a), \preceq, \gamma)$  is a weak adequate domain of limits, where  $\gamma$  maps each  $x \in X$  to  $\downarrow_X x$  and each  $F \in Y \setminus \eta_S(X_a)$  to itself;  $\preceq$  is defined by requirement (L<sub>3</sub>).

*Proof.* The Alexandroff-closed subsets of  $X$  are just its downward-closed subsets. So  $\gamma(z)$  is in  $\mathcal{H}(X_a)$  for all  $z$ , by (L<sub>1</sub>). Let  $Y$  be the image of  $\gamma$ . By (L<sub>3</sub>),  $\gamma$  defines an order-isomorphism of  $L \cup X$  onto  $Y$ . It remains to show that  $Y$  must contain  $\mathcal{S}(X_a)$ . Let  $F$  be any irreducible closed

subset of  $X_a$ . By  $(L_4)$ , there is a finite subset  $E \subseteq L \cup X$  such that  $F = \bigcup_{x \in E} \gamma(x)$ . Since  $F$  is irreducible, there must be a single  $x \in E$  such that  $F = \gamma(x)$ . So  $F$  is in  $Y$ .

Conversely, let  $X$  be wqo,  $L = Y \setminus \eta_S(X_a)$ , and  $\gamma, \preceq$  be as in the Lemma. Properties  $(L_1)$  and  $(L_3)$  hold by definition. For  $(L_4)$ , note that  $X_a$  is a Noetherian space, hence  $\mathcal{S}(X_a)$  is, too [15, Proposition 6.2]. However, by [15, Corollary 6.5], every closed subset of a sober Noetherian space is finitary. In particular, take any downward closed subset  $D$  of  $X$ . This is closed in  $X_a$ , hence its image  $\eta_S(D)$  by the topological embedding  $\eta_S$  is closed in  $\eta_S(X_a)$ , i.e., is of the form  $\eta_S(X_a) \cap F$  for some closed subset  $F$  of  $\mathcal{S}(X_a)$ . Also,  $D = \eta_S^{-1}(F)$ . Since  $\mathcal{S}(X_a)$  is both sober and Noetherian,  $F$  is finitary, hence is the downward-closure  $\downarrow_{\mathcal{S}(X_a)} E'$  of some finite subset  $E'$  in  $\mathcal{S}(X_a)$ . Let  $E$  be the set consisting of the (limit) elements in  $E' \cap L$ , and of the (non-limit) elements  $x \in X$  such that  $\downarrow_X x \in E'$ . We obtain  $\hat{\gamma}(E) = \bigcup_{z \in E'} z$ . On the other hand,  $D = \eta_S^{-1}(F) = \{x \in X \mid \downarrow x \in \downarrow_{\mathcal{S}(X_a)} E'\} = \{x \in X \mid \exists z \in E' \cdot \downarrow x \subseteq z\} = \bigcup_{z \in E'} z = \hat{\gamma}(E)$ . So  $(L_4)$  holds. ■

I.e., up to the coding function  $\gamma$ , there is a unique *minimal* WADL on any given wqo  $X$ : its sobrification  $\mathcal{S}(X_a)$ . There is also a unique largest one: its Hoare powerdomain  $\mathcal{H}(X_a)$ . An adequate domain of limits in the sense of Geeraerts *et al.* [13], i.e., one that additionally satisfies  $(L_2)$  is, up to isomorphism, any subset of  $\mathcal{H}(X_a)$  containing  $\mathcal{S}(X_a)$  plus the special closed set  $X$  itself as top element. We contend that  $\mathcal{S}(X_a)$  is, in general, the sole WADL worth considering.

*Ideal completions.* We have already argued that  $\mathcal{S}(X)$ , for any Noetherian space  $X$ , was in a sense of completion of  $X$ , adding missing limits. Another classical construction to add limits to some poset  $X$  is its *ideal completion*  $Idl(X)$ . The elements of the ideal completion of  $X$  are its *ideals*, i.e., its downward-closed directed families, ordered by inclusion.  $Idl(X)$  can be visualized as a form of Cauchy completion of  $X$ : we add all missing limits of directed families  $(x_i)_{i \in I}$  from  $X$ , by declaring these families to be their limits, equating two families when they have the same downward-closure. In  $Idl(X)$ , the finite elements are the elements of  $X$ ; formally, the map  $\eta_{Idl} : X \rightarrow Idl(X)$  that sends  $x$  to  $\downarrow x$  is an embedding, and the finite elements of  $Idl(X)$  are those of the form  $\eta_{Idl}(x)$ . It turns out that sobrification and ideal completion coincide, in a strong sense:

**Proposition 3.3** ([17]). *For any poset  $X$ ,  $\mathcal{S}(X_a) = Idl(X)$ .*

This is not just an isomorphism: the irreducible closed subsets of  $X_a$  are *exactly* the ideals. Note also that  $Idl(X)$  is always an algebraic dcpo [5, Proposition 2.2.22, Item 4].

When  $X$  is wqo, any downward-closed subset of  $X$  is a *finite* union of ideals. So  $(Idl(X) \setminus X, \subseteq, \text{id})$  is a WADL on  $X$ . Proposition 3.2 and Proposition 3.3 entail this, and a bit more:

**Theorem 3.4.** *For any wqo  $X$ ,  $\mathcal{S}(X_a) = Idl(X)$  is the smallest WADL on  $X$ .*

*Well-based continuous cpos.* There is a natural notion of limit in dcpos: whenever  $(x_i)_{i \in I}$  is a directed family, consider  $\sup_{i \in I} x_i$ . Starting from a wqo  $X$ , it is then natural to look at some dcpo  $Y$  that would contain  $X$  as a basis. In particular,  $Y$  would be continuous. This prompts us to define a *well-based continuous dcpo* as one that has a well-ordered basis—namely the original poset  $X$ .

This has several advantages. First, in general there are several notions of “sets of limits” of a given subset  $A \subseteq Y$ , but we shall see that they all coincide in continuous posets. Such sets of limits are important, because these are what we would like Karp-Miller-like procedures to compute, through acceleration techniques. Here are the possible notions. First, define  $\text{Lub}_Y(A)$  as the set of all least upper bounds in  $Y$  of directed families in  $A$ . Second,  $\text{Ind}_Y(A)$ , the *inductive hull* of  $A$  in  $Y$ , is the smallest sub-dcpo of  $Y$  containing  $A$ . Finally, the (Scott-topological) closure  $\text{cl}(A)$  of  $A$ . It is well-known that  $\text{cl}(A)$  is the smallest *downward closed* sub-dcpo of  $Y$  containing  $A$ .



(Recall that any open is upward closed, so that any closed set must be downward closed.) In any dcpo  $Y$ , one has  $A \subseteq \text{Lub}_Y(A) \subseteq \text{Ind}_Y(A) \subseteq \text{cl}(A)$ , and all inclusions are strict in general. E.g., in  $Y = \mathbb{N}_\omega$ , take  $A$  to be the set of even numbers. Then  $\text{Lub}_Y(A) = \text{Ind}_Y(A) = A \cup \{\omega\}$  while  $\text{cl}(A) = \mathbb{N}_\omega$ . While  $\text{Lub}_Y(A) = \text{Ind}_Y(A)$  in this case, there are cases where  $\text{Lub}_Y(A)$  is itself not closed under least upper bounds of directed families, and one has to iterate the  $\text{Lub}_Y$  operator to compute  $\text{Ind}_Y(A)$ . On continuous posets however, all these notions coincide [10, Appendix A].

**Proposition 3.5.** *Let  $Y$  be a continuous poset. Then, for every downward-closed subset  $A$  of  $Y$ ,  $\text{Ind}_Y(A) = \text{Lub}_Y(A) = \text{cl}(A)$ .*

We shall use this in Section 6. The key point now is that, again, well-based continuous dcpos coincide with completions of the form  $\mathcal{S}(X_a)$  or  $\text{Idl}(X)$ , and are therefore WADLs [10, Appendix B]. This even holds for continuous dcpos having a well-founded (not well-ordered) basis:

**Proposition 3.6.** *Any continuous dcpo  $Y$  with a well-founded basis is order-isomorphic to  $\text{Idl}(X)$  for some well-ordered set  $X$ . One may take the subset of finite elements of  $X$  for  $Y$ . If  $Y$  is well-based, then  $X$  is well-ordered.*

#### 4. Some Concrete WADLs

We now build WADLs for several concrete posets  $X$ . Following Proposition 3.2, it suffices to characterize  $\mathcal{S}(X_a)$ . Although  $\mathcal{S}(X_a) = \text{Idl}(X)$  (Proposition 3.3), the mathematics of  $\mathcal{S}(X_a)$  is easier to deal with than  $\text{Idl}(X)$ .

$\mathbb{N}^k$ . We start with  $X = \mathbb{N}^k$ , with the pointwise ordering. We have already recalled from [15] that  $\mathcal{S}(\mathbb{N}_a^k)$  was, up to isomorphism,  $(\mathbb{N}_\omega)^k$ , ordered with the pointwise ordering, where  $\omega$  is a new element above any natural number. This is the structure used in the standard Karp-Miller construction for Petri nets [16].

$\Sigma^*$ . Let  $\Sigma$  be a finite alphabet. The *divisibility ordering*  $|$  on  $\Sigma^*$ , a.k.a. the subsequence (non-continuous subword) ordering, is defined by  $a_1 a_2 \dots a_n | w_0 a_1 w_1 a_2 \dots a_n w_n$ , for any letters  $a_1, a_2, \dots, a_n \in \Sigma$  and words  $w_0, w_1, \dots, w_n \in \Sigma^*$ . There is a more general definition, where letters themselves are quasi-well-ordered. Our definition is the special case where the wqo on letters is  $=$ , and is the one required in verifying lossy channel systems [4]. Higman's Lemma states that  $|$  is wqo on  $\Sigma^*$ .

Any upward closed subset  $U$  of  $\Sigma^*$  is then of the form  $\uparrow E$ , with  $E$  finite. For any element  $w = a_1 a_2 \dots a_n$  of  $E$ ,  $\uparrow w$  is the regular language  $\Sigma^* a_1 \Sigma^* a_2 \Sigma^* \dots \Sigma^* a_n \Sigma^*$ . Forward analysis of lossy channel systems is instead based on simple regular expressions (SREs). Recall from [1] that an *atomic expression* is any regular expression of the form  $a^?$ , with  $a \in \Sigma$ , or  $A^*$ , where  $A$  is a non-empty subset of  $\Sigma$ . When  $A = \{a_1, \dots, a_m\}$ , we take  $A^*$  to denote  $(a_1 + \dots + a_m)^*$ ;  $a^?$  denotes  $\{a, \epsilon\}$ . A *product* is any regular expression of the form  $e_1 e_2 \dots e_n$  ( $n \in \mathbb{N}$ ), where each  $e_i$  is an atomic expression. A *simple regular expression*, or *SRE*, is a sum, either  $\emptyset$  or  $P_1 + \dots + P_k$ , where  $P_1, \dots, P_k$  are products. Sum is interpreted as union. That SREs and products are relevant here is no accident, as the following proposition shows.

**Proposition 4.1.** *The elements of  $\mathcal{S}(\Sigma_a^*)$  are exactly the denotations of products. The downward closed subsets of  $\Sigma^*$  are exactly the denotations of SREs.*

*Proof.* The second part is well-known. If  $F = P_1 + \dots + P_k$  is irreducible closed, then by irreducibility  $k$  must equal 1, hence  $F$  is denoted by a product. Conversely, it is easy to show that any product denotes an ideal, hence an element of  $\text{Idl}(X) = \mathcal{S}(X_a)$  (Proposition 3.3). ■

Inclusion between products can then be checked in quadratic time [1]. Inclusion between SREs can be checked in polynomial time, too, because of the remarkable property that  $P_1 + \dots + P_m \subseteq P'_1 + \dots + P'_n$  if and only if, for every  $i$  ( $1 \leq i \leq m$ ), there is a  $j$  ( $1 \leq j \leq n$ ) with  $P_i \subseteq P'_j$  [1, Lemma 1]. Similar lemmas are given by Abdulla *et al.* [3, Lemma 3, Lemma 4] for more general notions of SREs on words on infinite alphabets, and for a similar notion for finite multisets of elements from a finite set (both will be special cases of our constructions of Section 5). This is again no accident, and is a general fact about Noetherian spaces:

**Proposition 4.2.** *Let  $X$  be a Noetherian space, e.g., a wqo with its Alexandroff topology. Every closed subset  $F$  of  $X$  is a finite union of irreducible closed subsets  $C_1, \dots, C_m$ . If  $C'_1, \dots, C'_n$  are also irreducible closed, Then  $C_1 \cup \dots \cup C_m \subseteq C'_1 \cup \dots \cup C'_n$  if and only if for every  $i$  ( $1 \leq i \leq m$ ), there is a  $j$  ( $1 \leq j \leq n$ ) with  $C_i \subseteq C'_j$ .*

*Proof.* For the first part, by the results of [15],  $\mathcal{S}(X)$  is Noetherian and sober, which entails that  $F$  can be written  $\downarrow \{x_1, \dots, x_m\}$ ; now take  $C_i = \eta_{\mathcal{S}}^{-1}(\downarrow x_i)$ ,  $1 \leq i \leq m$  (see [10, Appendix C] for details). The second part is an easy consequence of irreducibility. ■

Proposition 4.2 suggests to represent closed subsets of  $X$  as finite subsets  $A$  of  $\mathcal{S}(X)$ , interpreted as the closed set  $\bigcup_{C \in A} C$ . When  $X = \Sigma_a^*$ ,  $A$  is a finite set of products, i.e., an SRE. When  $X = \mathbb{N}_a^k$ ,  $A$  is a finite subset of  $\mathbb{N}_\omega^k$ , interpreted as  $\downarrow A \cap \mathbb{N}^k$ .

*Finite Trees.* All the examples given above are well-known. Here is one that is new, and also more involved than the previous ones. Let  $\mathcal{F}$  be a finite signature of function symbols with their arities. We let  $\mathcal{F}_k$  the set of function symbols of arity  $k$ ;  $\mathcal{F}_0$  is the set of *constants*, and is assumed to be non-empty. The set  $\mathcal{T}(\mathcal{F})$  is the set of ground terms built from  $\mathcal{F}$ . Kruskal's Tree Theorem states that this is well-quasi-ordered by the *homeomorphic embedding* ordering  $\trianglelefteq$ , defined as the smallest relation such that, whenever  $u = f(u_1, \dots, u_m)$  and  $v = g(v_1, \dots, v_n)$ ,  $u \trianglelefteq v$  if and only if  $u \trianglelefteq v_j$  for some  $j$ ,  $1 \leq j \leq n$ , or  $f = g$ ,  $m = n$ , and  $u_1 \trianglelefteq v_1, u_2 \trianglelefteq v_2, \dots, u_m \trianglelefteq v_m$ . (As for  $\Sigma^*$ , we take a special case, where each function has fixed arity.)

The structure of  $\mathcal{S}(\mathcal{T}(\mathcal{F})_a)$  is described using an extension of SREs to the tree case. This uses regular tree expressions as defined in [6, Section 2.2]. Let  $\mathcal{K}$  be a countably infinite set of additional constants, called *holes*  $\square$ . Most tree regular expressions are self-explanatory, except Kleene star  $L^{*,\square}$  and concatenation  $L.\square L'$ . The latter denotes the set of all terms obtained from a term  $t$  in  $L$  by replacing all occurrences of  $\square$  by (possibly different) terms from  $L'$ . The language of a hole  $\square$  is just  $\{\square\}$ .  $L^{*,\square}$  is the infinite union of the languages of  $\square, L, L.\square L, L.\square L.\square L$ , etc.

**Definition 4.3 (STRE).** *Tree products and product iterators are defined inductively by:*

- Every hole  $\square$  is a tree product.
- $f^?(P_1, \dots, P_k)$  is a tree product, for any  $f \in \Sigma_k$  and any tree products  $P_1, \dots, P_k$ . We take  $f^?(P_1, \dots, P_k)$  as an abbreviation for  $f(P_1, \dots, P_k) + P_1 + \dots + P_k$ .
- $(\sum_{i=1}^n C_i)^{*,\square}. \square P$  is a tree product, for any tree product  $P$ , any  $n \geq 1$ , and any product iterators  $C_i$  over  $\square$ ,  $1 \leq i \leq n$ . We write  $\sum_{i=1}^n C_i$  for  $C_1 + C_2 + \dots + C_n$ .
- $f(P_1, \dots, P_k)$  is a product iterator over  $\square$  for any  $f \in \Sigma_k$ , where: 1. each  $P_i$ ,  $1 \leq i \leq k$  is either  $\square$  itself or a tree product such that  $\square$  is not in the language of  $P_i$ ; and 2.  $P_i = \square$  for some  $i$ ,  $1 \leq i \leq k$ .

A *simple tree regular expression* (STRE) is a finite sum of tree products.

A tree regular expression is *closed* iff it has no free hole, where a hole is free in  $f(L_1, \dots, L_k)$ ,  $L_1 + \dots + L_k$ , or in  $f^?(L_1, \dots, L_k)$  iff it is free in some  $L_i$ ,  $1 \leq i \leq k$ ; the only free hole in  $\square$  is

$\square$  itself; the free holes of  $L^{*,\square}$  are those of  $L$ , plus  $\square$ ; the free holes of  $L.\square L'$  are those of  $L'$ , plus those of  $L$  except  $\square$ . E.g.,  $f^?(a^?, b^?)$  and  $(f(\square, g^?(a^?)) + f(g^?(b^?), \square))^*, \square.f^?(a^?, b^?)$  are closed tree products. Then [10, Appendix D]:

**Theorem 4.4.** *The elements of  $\mathcal{S}(\mathcal{T}(\mathcal{F})_a)$  are exactly the denotations of closed tree products. The downward closed subsets of  $\mathcal{T}(\mathcal{F})$  are exactly the denotations of closed STREs. Inclusion is decidable in polynomial time for tree products and for STREs.*

## 5. A Hierarchy of Data Types

The sobrification WADL can be computed in a compositional way, as we now show. Consider the following grammar of data types of interest in verification:

$D ::=$	$\mathbb{N}$	natural numbers
	$A_{\leq}$	finite set $A$ , quasi-ordered by $\leq$
	$D_1 \times \dots \times D_k$	finite product
	$D_1 + \dots + D_k$	finite, disjoint sum
	$D^*$	finite words
	$D^{\oplus}$	finite multisets

By *compositional*, we mean that the sobrification of any data type  $D$  is computed in terms of the sobrifications of its arguments. E.g.,  $\mathcal{S}(D_a^*)$  will be expressed as some extended form of products over  $\mathcal{S}(D_a)$ . The semantics of data types is the intuitive one. Finite products are quasi-ordered by the pointwise quasi-ordering, finite disjoint sums by comparing elements in each summand—elements from different summands are incomparable. For any poset  $X$  (even infinite),  $X^*$  is the set of finite words over  $X$  ordered by the *embedding* quasi-ordering  $\leq^*$ :  $w \leq^* w'$  iff, writing  $w$  as the sequence of  $m$  letters  $a_1 a_2 \dots a_m$ , one can write  $w'$  as  $w_0 a'_1 w_1 a'_2 w_2 \dots w_{m-1} a'_m w'_m$  with  $a_1 \leq a'_1$ ,  $a_2 \leq a'_2$ ,  $\dots$ ,  $a_m \leq a'_m$ .  $X^{\oplus}$  is the set of finite multisets  $\{x_1, \dots, x_n\}$  of elements of  $X$ , and is quasi-ordered by  $\leq^{\oplus}$ , defined as:  $\{x_1, x_2, \dots, x_m\} \leq^{\oplus} \{y_1, y_2, \dots, y_n\}$  iff there is an injective map  $r : \{1, \dots, m\} \rightarrow \{1, \dots, n\}$  such that  $x_i \leq y_{r(i)}$  for all  $i$ ,  $1 \leq i \leq m$ . When  $\leq$  is just equality,  $m \leq^{\oplus} m'$  iff every element of  $m$  occurs at least as many times in  $m'$  as in  $m$ : this is the  $\leq^m$  quasi-ordering considered, on finite sets  $X$ , by Abdulla *et al.* [3, Section 2].

The analogue of products and SREs for  $D^*$  is given by the following definition, which generalizes the  $\Sigma^*$  case of Section 4. Note that  $D$  is in general an *infinite* alphabet, as in [3]. The following definition should be compared with [1]. The only meaningful difference is the replacement of  $(a + \epsilon)$ , where  $a$  is a letter, with  $C^?$ , where  $C \in \mathcal{S}(X_a)$ . It should also be compared with the *word language generators* of [3, Section 6]. Indeed, the latter are exactly our products on  $A^{\oplus}$ , where  $A$  is a finite alphabet (in our notation,  $A_{\leq}$ , with  $\leq$  given as equality).

**Definition 5.1** (Product, SRE). Let  $X$  be a topological space. Let  $X^*$  be the set of finite words on  $X$ . For any  $A, B \subseteq X^*$ , let  $AB$  be  $\{ww' \mid w \in A, w' \in B\}$ ,  $A^*$  be the set of words on  $A$ ,  $A^? = A \cup \{\epsilon\}$ .

*Atomic expressions* are either of the form  $C^?$ , with  $C \in \mathcal{S}(X)$ , or  $A^*$ , with  $A$  a non-empty finite subset of  $\mathcal{S}(X)$ . *Products* are finite sequences  $e_1 e_2 \dots e_k$ ,  $k \in \mathbb{N}$ , and *SREs* are finite sums of products. The denotation of atomic expressions is given by  $\llbracket C^? \rrbracket = C^?$ ,  $\llbracket A^* \rrbracket = (\bigcup_{C \in A} \llbracket C \rrbracket)^*$ ; of products by  $\llbracket e_1 e_2 \dots e_k \rrbracket = \llbracket e_1 \rrbracket \llbracket e_2 \rrbracket \dots \llbracket e_k \rrbracket$ ; of SREs by  $\llbracket P_1 + \dots + P_k \rrbracket = \bigcup_{i=1}^k \llbracket P_i \rrbracket$ .

Atomic expressions are ordered by  $C^? \sqsubseteq C'^?$  iff  $C \subseteq C'$ ;  $C^? \sqsubseteq A'^*$  iff  $C \subseteq C'$  for some  $C' \in A'$ ;  $A^* \not\sqsubseteq C'^?$ ;  $A^* \sqsubseteq A'^*$  iff for every  $C \in A$ , there is a  $C' \in A'$  with  $C \subseteq C'$ . Products are quasi-ordered by  $eP \sqsubseteq e'P'$  iff (1)  $e \not\sqsubseteq e'$  and  $eP \sqsubseteq P'$ , or (2)  $e = C^?$ ,  $e' = C'^?$ ,  $C \subseteq C'$  and  $P \sqsubseteq P'$ , or (3)  $e' = A'^*$ ,  $e \sqsubseteq A'^*$  and  $P \sqsubseteq e'P'$ . We let  $\equiv$  be  $\sqsubseteq \cap \supseteq$ .



**Definition 5.2** ( $\otimes$ -Product,  $\otimes$ -SRE). Let  $X$  be a topological space. For any  $A, B \subseteq X$ , let  $A \odot B = \{m \uplus m' \mid m \in A, m' \in B\}$ ,  $A^*$  be the set of multisets comprised of elements from  $A$ ,  $A^\circ = \{\{x\} \mid x \in A\} \cup \{\emptyset\}$ , where  $\emptyset$  is the empty multiset.

The  $\otimes$ -products  $P$  are the expressions of the form  $A^* \odot C_1^\circ \odot \dots \odot C_n^\circ$ , where  $A$  is a finite subset of  $\mathcal{S}(X)$ ,  $n \in \mathbb{N}$ , and  $C_1, \dots, C_n \in \mathcal{S}(X)$ . Their denotation  $\llbracket P \rrbracket$  is  $(\bigcup_{C \in A} C)^* \odot \llbracket C_1 \rrbracket^\circ \odot \dots \odot \llbracket C_n \rrbracket^\circ$ . They are quasi-ordered by  $P \sqsubseteq P'$ , where  $P = A^* \odot C_1^\circ \odot C_2^\circ \odot \dots \odot C_m^\circ$  and  $P' = A'^* \odot C'_1 \odot C'_2 \odot \dots \odot C'_n$ , iff: (1) for every  $C \in A$ , there is a  $C' \in A'$  with  $C \subseteq C'$ , and (2) letting  $I$  be the subset of those indices  $i$ ,  $1 \leq i \leq m$ , such that  $C_i \subseteq C'$  for no  $C' \in A'$ , there is an injective map  $r : I \rightarrow \{1, \dots, n\}$  such that  $C_i \subseteq C'_{r(i)}$  for all  $i \in I$ . Let  $\equiv$  be  $\sqsubseteq \cap \supseteq$ .

**Theorem 5.3.** For every data type  $D$ ,  $\mathcal{S}(D_a)$  is Noetherian, and is computed by:  $\mathcal{S}(\mathbb{N}_a) = \mathbb{N}_\omega$ ;  $\mathcal{S}(A_{\leq a}) = A_{\leq}$ ;  $\mathcal{S}((D_1 \times \dots \times D_k)_a) = \mathcal{S}(D_{1a}) \times \dots \times \mathcal{S}(D_{ka})$ ;  $\mathcal{S}((D_1 + \dots + D_k)_a) = \mathcal{S}(D_{1a}) + \dots + \mathcal{S}(D_{ka})$ ;  $\mathcal{S}(D^*)$  is the set of products on  $D$  modulo  $\equiv$ , ordered by  $\sqsubseteq$  (Definition 5.1);  $\mathcal{S}(D^\circ)$  is the set of  $\otimes$ -products on  $D$  modulo  $\equiv$ , ordered by  $\sqsubseteq$  (Definition 5.2).

For any data type  $D$ , equality and ordering (inclusion) in  $\mathcal{S}(D_a)$  is decidable in the polynomial hierarchy.

*Proof.* We show that  $\mathcal{S}(D_a)$  is Noetherian and is computed as given above, by induction on the construction of  $D$ . We in fact prove the following two facts separately: (1)  $\mathcal{S}(D)$  is Noetherian ( $D$ , not  $D_a$ ), where  $D$  is topologized in a suitable way, and (2)  $D = D_a$ .

To show (1), we topologize  $\mathbb{N}$  and  $A_{\leq}$  with their Alexandroff topologies, sums and products with the sum and product topologies respectively;  $X^*$  with the *subword topology*, viz. the smallest containing the open subsets  $X^*U_1X^*U_2X^* \dots X^*U_nX^*$ ,  $n \in \mathbb{N}$ ,  $U_1, U_2, \dots, U_n$  open in  $X$ ; and  $X^\circ$  with the *sub-multiset topology*, namely the smallest containing the subsets  $X^\circ \odot U_1 \odot U_2 \odot \dots \odot U_n$ ,  $n \in \mathbb{N}$ , where  $U_1, U_2, \dots, U_n$  are open subsets of  $X$ . The case of  $\mathbb{N}$  has already been discussed above. When  $A_{\leq}$  is finite, it is both Noetherian and sober. The case of finite products is by [15, Section 6], that of finite sums by [15, Section 4]. The cases of  $X^*$ , resp.  $X^\circ$ , are dealt with in [10, Appendices E, F].

To show (2), we appeal to a series of coincidence lemmas, showing that  $(X^*)_a = X_a^*$  and that  $(X^\circ)_a = X_a^\circ$  notably. The other cases are obvious.

Finally, we show that inclusion and equality are decidable in the polynomial hierarchy. For this, we show in the appendices that inclusion on  $\mathcal{S}(D^*)$  is  $\sqsubseteq$  on products, and is decidable by a polynomial time algorithm modulo calls to an oracle deciding inclusion in  $\mathcal{S}(D)$ . This is by dynamic programming. Inclusion in  $\mathcal{S}(D^\circ)$  is  $\sqsubseteq$  on  $\otimes$ -products, and is decidable by a non-deterministic polynomial time algorithm modulo a similar oracle. We conclude since the orderings on  $\mathbb{N}_\omega$  and on  $A_{\leq}$  are polynomial-time decidable, while inclusion in  $\mathcal{S}(D_1 \times \dots \times D_k) \cong \mathcal{S}(D_1) \times \dots \times \mathcal{S}(D_k)$  and in  $\mathcal{S}(D_1 + \dots + D_k) \cong \mathcal{S}(D_1) + \dots + \mathcal{S}(D_k)$  are polynomial time modulo oracles deciding inclusion in  $\mathcal{S}(D_i)$ ,  $1 \leq i \leq k$ .  $\blacksquare$

Look at some special cases of this construction. First,  $\mathbb{N}^k$  is the data type  $\mathbb{N} \times \dots \times \mathbb{N}$ , and we retrieve that  $\mathcal{S}(\mathbb{N}^k) = \mathbb{N}_\omega^k$ . Second, when  $A$  is a finite alphabet,  $A^*$  is given by products, as given in the  $\Sigma^*$  paragraph of Section 4; i.e., we retrieve the products (and SREs) of Abdulla *et al.* [1]. The more complicated case  $(A^\circ)^*$  was dealt with by Abdulla *et al.* [3]. We note that the elements of  $\mathcal{S}((A^\circ)_a^*)$  are exactly their *word language generators*, which we retrieve here in a principled way. Additionally, we can deal with more complex data structures such as, e.g.,  $((\mathbb{N} \times A_{\leq})^* \times \mathbb{N}^\circ)^*$ .

Finally, note that (1) and (2) are two separate concerns in the proof of Theorem 5.3. If we are ready to relinquish orderings for the more general topological route, as advocated in [15], we could also enrich our grammar of data types with infinite constructions such as  $\mathbb{P}(D)$ , where  $\mathbb{P}(D)$  is interpreted as the powerset of  $D$  with the so-called lower Vietoris topology. In fact,  $\mathcal{S}(\mathbb{P}(X)) \cong$

$\mathcal{H}(X)$  is Noetherian whenever  $X$  is, and its elements can be represented as *finite* subsets  $A$  of  $\mathcal{S}(X)$ , interpreted as  $\bigcup_{C \in A} C$  [10, Appendix G]. In a sense, while  $\mathcal{S}(X_a) = \text{Idl}(X)$  for all ordered spaces  $X$ , the sobrification construction is more robust than the ideal completion.

## 6. Completing WSTS, or: Towards Forward Procedures Computing the Cover

We show how one may use our completions on wqos to deal with forward analysis of well-structured systems. We shall describe this in more detail in another paper. First note that any data type  $D$  of Section 5 is suited to applying the expand, enlarge and check algorithm [13] out of the box to this end, since then  $\mathcal{S}(D_a)$  is (the least) WADL for  $D$ . We instead explore extensions of the Karp-Miller procedure [16], in the spirit of Finkel [9] or Emerson and Namjoshi [7]. While the latter assumes an already built completion, we construct it. Also, we make explicit how this kind of acceleration-based procedure really computes the cover, i.e.,  $\downarrow \text{Post}^*(\downarrow x)$ , in Proposition 6.1.

Recall that a *well-structured transition system* (WSTS) is a triple  $S = (X, \leq, (\delta_i)_{i=1}^n)$ , where  $X$  is well-quasi-ordered by  $\leq$ , and each  $\delta_i : X \rightarrow X$  is a partial monotonic transition function. (By “partial monotonic” we mean that the domain of  $\delta_i$  is upward closed, and  $\delta_i$  is monotonic on its domain.) Letting  $\text{Pre}(A) = \bigcup_{i=1}^n \delta_i^{-1}(A)$ ,  $\text{Pre}^0(A) = A$ , and  $\text{Pre}^k(A) = \bigcup_{i \in \mathbb{N}} \text{Pre}^k(A)$ , it is well-known that any upward closed subset of  $X$  is of the form  $\uparrow E$  for some finite  $E \subseteq X$ , and that  $\text{Pre}^*(\uparrow E)$  is an upward-closed subset  $\uparrow E'$ ,  $E'$  finite, that arises as  $\bigcup_{k=0}^m \text{Pre}^k(\uparrow E)$  for some  $m \in \mathbb{N}$ . Hence, provided  $\leq$  is decidable and  $\delta_i^{-1}(\uparrow E)$  is computable for each finite  $E$ , it is decidable whether  $x \in \text{Pre}^*(\uparrow E)$ , i.e., whether one may reach  $\uparrow E$  from  $x$  in finitely many steps. It is equivalent to check whether  $y \in \downarrow \text{Post}^*(\downarrow x)$  for some  $y \in E$ , where  $\text{Post}(A) = \bigcup_{i=1}^n \delta_i(A)$ ,  $\text{Post}^0(A) = A$ , and  $\text{Post}^k(A) = \bigcup_{i \in \mathbb{N}} \text{Post}^k(A)$ .

All the existing symbolic procedures that attempt to compute  $\downarrow \text{Post}^*(\downarrow x)$ , even with a finite number of accelerations (e.g., Fast, Trex, Lash), can only compute subsets of the larger set  $\text{Lub}(\downarrow \text{Post}^*(\downarrow x))$ . In general,  $\text{Lub}(\downarrow \text{Post}^*(\downarrow x))$  does not admit a finite representation. On the other hand, we know that the Scott-closure  $\text{cl}(\text{Post}^*(\downarrow x))$ , as a closed subset of  $\text{Idl}(X)$  (intersected with  $X$  itself), is always finitary. Indeed, it is also a closed subset of  $\mathcal{S}(X_a)$  (Proposition 3.3), which is represented as the downward closure of finitely many elements of  $\mathcal{S}(X_a)$ . Since  $Y = \text{Idl}(X)$  is continuous, Proposition 3.5 allows us to conclude that  $\text{Lub}_Y(\downarrow \text{Post}^*(\downarrow x)) = \text{cl}(\text{Post}^*(\downarrow x))$  is finitary—hence representable provided  $X$  is one of the data types of Section 5.

This leads to the following construction. Any partial monotonic map  $f : X \rightarrow Y$  between quasi-ordered sets lifts to a *continuous* partial map  $\mathcal{S}f : \mathcal{S}(X_a) \rightarrow \mathcal{S}(Y_a)$ : for each irreducible closed subset (a.k.a., ideal)  $C$  of  $\mathcal{S}(X_a)$ , either  $C \cap \text{dom } f \neq \emptyset$  and  $\mathcal{S}f(C) = \downarrow f(C) = \{y \in Y \mid \exists x \in C \cap \text{dom } f \cdot y \leq f(x)\}$ , or  $C \cap \text{dom } f = \emptyset$  and  $\mathcal{S}f(C)$  is undefined. The *completion* of a WSTS  $S = (X, \leq, (\delta_i)_{i=1}^n)$  is then the transition system  $\hat{S} = (\mathcal{S}(X_a), \subseteq, (\mathcal{S}\delta_i)_{i=1}^n)$ .

For example, when  $X = \mathbb{N}^k$ , and  $S$  is a Petri net with transitions  $\delta_i$  defined as  $\delta_i(\vec{x}) = \vec{x} + \vec{d}_i$  (where  $\vec{d}_i \in \mathbb{Z}^k$ ; this is defined whenever  $\vec{x} + \vec{d}_i \in \mathbb{N}^k$ ), then  $\hat{S}$  is the transition system whose set of states is  $\mathcal{S}(X) = \mathbb{N}_\omega^k$ , and whose transition functions are:  $\mathcal{S}\delta_i(\vec{x}) = \vec{x} + \vec{d}_i$ , whenever this has only non-negative coordinates, taking the convention that  $\omega + d = \omega$  for any  $d \in \mathbb{Z}$ .

We may emulate lossy channel systems through the following *functional-lossy* channel systems (FLCS). For simplicity, we assume just one channel and no local state; the general case would only make the presentation more obscure. An FLCS differs from an LCS in that it loses only the least amount of messages needed to enable transitions. Take  $X = \Sigma^*$  for some finite alphabet  $\Sigma$  of messages; the transitions are either of the form  $\delta_i(w) = wa_i$  for some fixed letter  $a_i$  (sending  $a_i$  onto the channel), or of the form  $\delta_i(w) = w_2$  whenever  $w$  is of the form  $w_1a_iw_2$ , with  $w_1$  not containing

$a_i$  (expecting to receive  $a_i$ ). Any LCS is cover-equivalent to the FLCS with the same sends and receives, where two systems are *cover-equivalent* if and only if they have the same sets  $\downarrow \text{Post}^*(F)$  for any downward-closed  $F$ . Equating  $\mathcal{S}(\Sigma_a^*)$  with the set of products, as advocated in Section 4, we find that transition functions of the first kind lift to  $\mathcal{S}\delta_i(P) = Pa_i^?$ , while transition functions of the second kind lift to:  $\mathcal{S}\delta_i(\epsilon)$  is undefined,  $\mathcal{S}\delta_i(a^?P) = \mathcal{S}\delta_i(P)$  if  $a_i \neq a$ ,  $\mathcal{S}\delta_i(a_i^?P) = P$ ,  $\mathcal{S}\delta_i(A^*P) = \mathcal{S}\delta_i(P)$  if  $a_i \notin A$ ,  $\mathcal{S}\delta_i(A^*P) = A^*P$  otherwise. This is exactly how Trex computes successors [1, Lemma 6].

In general, the results of Section 5 allow us to use any domain of datatypes  $D$  for the state space  $X$  of  $S$ . The construction  $\widehat{S}$  then generalizes all previous constructions, which used to be defined specifically for each datatype.

The Karp-Miller algorithm in Petri nets, or the Trex procedure for lossy channel systems, gives information about the cover  $\downarrow \text{Post}^*(\downarrow x)$ . This is true of *any* completion  $\widehat{S}$  as constructed above:

**Proposition 6.1.** *Let  $S$  be a WSTS. Let  $\widehat{\text{Post}}$  be the Post map of the completion  $\widehat{S}$ . For any closed subset  $F$  of  $\mathcal{S}(X_a)$ ,  $\widehat{\text{Post}}(F) = \text{cl}(\text{Post}(F \cap X))$ , and  $\widehat{\text{Post}}^*(F) = \text{cl}(\text{Post}^*(F \cap X))$ . Hence, for any downward closed subset  $F$  of  $X$ ,  $\downarrow \text{Post}(F) = X \cap \widehat{\text{Post}}(F)$ ,  $\downarrow \text{Post}^*(F) = X \cap \widehat{\text{Post}}^*(F)$ .*

*Proof.* Let  $F$  be closed in  $\mathcal{S}(X_a)$ .  $\widehat{\text{Post}}(F) = \bigcup_{i=1}^n \text{cl}(\delta_i(F)) = \text{cl}(\bigcup_{i=1}^n \delta_i(F)) = \text{cl}(\text{Post}(F))$ , since closure commutes with (arbitrary) unions. We then claim that  $\widehat{\text{Post}}^k(F) = \text{cl}(\text{Post}^k(F))$  for each  $k \in \mathbb{N}$ . This is by induction on  $k$ . The cases  $k = 0, 1$  are obvious. When  $k \geq 2$ , we use the fact that, for any continuous partial map  $f$ :  $(*) \text{cl}(f(\text{cl}(A))) = \text{cl}(f(A))$ . Then  $\widehat{\text{Post}}^k(F) = \bigcup_{i=1}^n \text{cl}(\delta_i(\widehat{\text{Post}}^{k-1}(F))) = \bigcup_{i=1}^n \text{cl}(\delta_i(\text{cl}(\text{Post}^{k-1}(F)))) = \bigcup_{i=1}^n \text{cl}(\delta_i(\text{Post}^{k-1}(F)))$  (by  $(*)$ )  $= \text{cl}(\text{Post}^k(F))$ . Finally,  $\widehat{\text{Post}}^*(F) = \bigcup_{k \in \mathbb{N}} \widehat{\text{Post}}^k(F) = \bigcup_{k \in \mathbb{N}} \text{cl}(\text{Post}^k(F)) = \text{cl}(\text{Post}^*(F))$ . We conclude, since for any  $A \subseteq X$ ,  $\downarrow A$  is the closure of  $A$  in  $X_a$ ; the topology of  $X_a$  is the subspace topology of that of  $\mathcal{S}(X_a)$ ; so, writing  $\text{cl}$  for closure in  $\mathcal{S}(X_a)$ ,  $\downarrow A = X \cap \text{cl}(A)$ . ■

Writing  $F$  as the finite union  $C_1 \cup \dots \cup C_k$ , where  $C_1, \dots, C_k \in \mathcal{S}(X_a)$ ,  $\widehat{\text{Post}}(F)$  is computable as  $\bigcup_{1 \leq i_1, \dots, i_n \leq k} \mathcal{S}\delta_1(C_{i_1}) \cup \dots \cup \mathcal{S}\delta_n(C_{i_n})$ , assuming  $\mathcal{S}\delta_i$  computable for each  $i$ . (We take  $\mathcal{S}\delta_j(C_i)$  to mean  $\emptyset$  if undefined, for notational convenience.) Although  $\mathcal{S}\delta_i$  may be uncomputable even when  $\delta_i$  is, it is computable on most WSTS in use. This holds, for example, for Petri nets and lossy channel systems, as exemplified above.

So it is easy to compute  $\downarrow \text{Post}(\downarrow x)$ , as (the intersection of  $X$  with)  $\widehat{\text{Post}}(\downarrow x)$ . Computing  $\downarrow \text{Post}^*(\downarrow x)$  (our goal) is also easily computed as  $\widehat{\text{Post}}^*(\downarrow x)$  (intersected with  $X$  again), using acceleration techniques for loops. This is what the Karp-Miller construction does for Petri nets, what Trex does for lossy channel systems [1]. (We examine termination issues below.) Our framework generalizes all these procedures, using a weak acceleration assumption, whereby we assume that we can compute the least upper bound of the values of loops iterated  $k$  times,  $k \in \mathbb{N}$ . For any continuous partial map  $g : Y \rightarrow Y$  (with open domain) on a dcpo  $Y$ , let the *iteration*  $\overline{g}$  be the map of domain  $\text{dom } g$  such that  $\overline{g}(y)$  is the least upper bound of  $(g^k(y))_{k \in \mathbb{N}}$  if  $y < g(y)$ , and  $g(y)$  otherwise. Let  $\Delta = \{\mathcal{S}\delta_1, \dots, \mathcal{S}\delta_n\}$ ,  $\Delta^*$  be the set of all composites of finitely many maps from  $\Delta$ . Our *acceleration assumption* is that one can compute  $\overline{g}(y)$  for any  $g \in \Delta^*$ ,  $y \in \mathcal{S}(X_a)$ . The following procedure then computes  $\downarrow \text{Post}^*(\downarrow x)$ , as (the intersection of  $X$  with)  $\widehat{\text{Post}}^*(\downarrow x)$ , itself represented as a finite union of elements of  $\mathcal{S}(X_a)$ : initially, let  $A$  be  $\{x\}$ ; then, while  $\widehat{\text{Post}}(A) \not\subseteq \downarrow A$ , choose fairly  $(g, a) \in \Delta^* \times A$  such that  $a \in \text{dom } g$  and add  $\overline{g}(a)$  to  $A$ . If this terminates,  $A$  is a finite set whose downward closure is exactly  $\downarrow \text{Post}^*(\downarrow x)$ . Despite its simplicity, this is the essence of the Karp-Miller procedure, generalized to a large class of spaces  $X$ .

Termination is ensured for flat systems, i.e., systems whose control graph has no nested loop, as one only has to compute the effect of a finite number of loops. In general, the procedure terminates on *cover-flattable* systems, that is systems that are cover-equivalent to some flat system. Petri nets are cover-flattable, while, e.g., not all LCS are: recall that, in an LCS,  $\downarrow \text{Post}^*(\downarrow x)$  is *always* representable as an SRE, however not effectively so.

## 7. Conclusion and Perspectives

We have developed the first comprehensive theory of downward-closed subsets, as required for a general understanding of forward analysis techniques of WSTS. This generalizes previous domain proposals on tuples of natural numbers, on words, on multisets, allowing for nested datatypes, and infinite alphabets. Each of these domains is effective, in the sense that each has finite presentations with a decidable ordering. We have also shown how the notion of sobrification  $\mathcal{S}(X_a)$  was in a sense inevitable (Section 3), and described how this applied to compute downward closures of reachable sets of configurations in WSTS (Section 6). We plan to describe such new forward analysis algorithms, in more detail, in papers to come.

## References

- [1] P. A. Abdulla, A. Bouajjani, and B. Jonsson. On-the-fly analysis of systems with unbounded, lossy fifo channels. In *CAV'98*, Vancouver, Canada, 1998. Springer Verlag LNCS 1427.
- [2] P. A. Abdulla, K. Čerāns, B. Jonsson, and Y.-K. Tsay. Algorithmic analysis of programs with well quasi-ordered domains. *Inf. Comput.*, 160(1-2):109–127, 2000.
- [3] P. A. Abdulla, J. Deneux, P. Mahata, and A. Nylén. Forward reachability analysis of timed Petri nets. In Y. Lakhnech and S. Yovine, editors, *FORMATS/FTRTFT*, pages 343–362. Springer Verlag LNCS 3253, 2004.
- [4] P. A. Abdulla and B. Jonsson. Verifying programs with unreliable channels. In *LICS'93*, pages 160–170, 1993.
- [5] S. Abramsky and A. Jung. Domain theory. In S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Comp. Sci.*, volume 3, pages 1–168. OUP, 1994.
- [6] H. Comon, M. Dauchet, R. Gilleron, F. Jacquemard, D. Lugiez, S. Tison, and M. Tommasi. Tree automata techniques and applications. [www.grappa.univ-lille3.fr/tata](http://www.grappa.univ-lille3.fr/tata), 2004.
- [7] E. A. Emerson and K. S. Namjoshi. On model checking for non-deterministic infinite-state systems. In *LICS'98*, pages 70–80, 1998.
- [8] J. Esparza, A. Finkel, and R. Mayr. On the verification of broadcast protocols. In *LICS'99*, pages 352–359, 1999.
- [9] A. Finkel. Reduction and covering of infinite reachability trees. *Inf. Comput.*, 89(2):144–179, 1990.
- [10] A. Finkel and J. Goubault-Larrecq. Forward analysis for WSTS, part I: Completions. Research report, LSV, ENS Cachan, ENS Cachan, 61 avenue du président Wilson, 94230 Cachan, Dec. 2008. Full version.
- [11] A. Finkel and P. Schnoebelen. Well-structured transition systems everywhere! *Theor. Comp. Sci.*, 256(1–2):63–92, 2001.
- [12] P. Ganty, J.-F. Raskin, and L. van Begin. A complete abstract interpretation framework for coverability properties of WSTS. In *VMCAI'06*, pages 49–64. Springer Verlag LNCS 3855, 2006.
- [13] G. Geeraerts, J.-F. Raskin, and L. van Begin. Expand, enlarge and check: New algorithms for the coverability problem of WSTS. *J. Comp. Sys. Sci.*, 72(1):180–203, 2006.
- [14] G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. Mislove, and D. S. Scott. Continuous lattices and domains. In *Encyc. Math. and its Applications*, volume 93. CUP, 2003.
- [15] J. Goubault-Larrecq. On Noetherian spaces. In *LICS'07*, pages 453–462, 2007.
- [16] R. M. Karp and R. E. Miller. Parallel program schemata. *J. Comp. Sys. Sci.*, 3(2):147–195, 1969.
- [17] M. Mislove. Algebraic posets, algebraic cpo's and models of concurrency. In *Topology and Category Theory in Computer Science*, pages 75–109. Clarendon Press, 1981.