

# Rational Spaces and Set Constraints<sup>\*</sup>

Dexter Kozen

*Computer Science Department, Cornell University, Ithaca, NY 14853-7501, USA*

Set constraints are inclusions between expressions denoting sets of ground terms. They have been used extensively in program analysis and type inference. In this paper we investigate the topological structure of the spaces of solutions to systems of set constraints. We identify a family of topological spaces called *rational spaces*, which formalize the notion of a topological space with a regular or self-similar structure, such as the Cantor discontinuum or the space of runs of a finite automaton. We develop the basic theory of rational spaces and derive generalizations and proofs from topological principles of some results in the literature on set constraints.

## 1 Introduction

Set constraints are inclusions between expressions denoting sets of ground terms. They have been used extensively in program analysis and type inference for many years [3,4,13,15,16,25–27,29]. Considerable recent effort has focussed on the complexity of the satisfiability problem [1,2,5–8,11,12,14,28]. Set constraints have also recently been used to define a constraint logic programming language over sets of ground terms that generalizes ordinary logic programming over an Herbrand domain [21].

Set constraints exhibit a rich mathematical structure. There are strong connections to automata theory [11,12], type theory [23,24], first-order monadic logic [6,7], Boolean algebras with operators [17,18], and modal logic [19]. There are algebraic and topological formulations, corresponding roughly to “soft” and “hard” typing respectively, which are related by Stone duality [19].

---

<sup>\*</sup> In P. D. Mosses, M. Nielsen, and M. I. Schwartzbach, editors, *Proc. Sixth Int. Joint Conf. Theory and Practice of Software Develop. (TAPSOFT'95)*, volume 915 of *Lect. Notes in Comput. Sci.*, pages 42–61. Springer, May 1995.

Many results in the literature on set constraints are topological in flavor. For example, Aiken *et al.* [2] prove that mixed positive and negative set constraints are strictly more expressive than positive constraints alone. The proof of this result is based on a compactness lemma which states that a system of positive constraints is satisfiable if and only if all finite subsets are satisfiable. It is well known that the compactness of classical propositional and predicate logic is essentially a topological result: logical compactness is equivalent to the topological compactness of a dual space. This is also the case here.

In [11], Gilleron *et al.* introduce *tree set automata*. Among other results, they establish various closure properties of these automata and show that every satisfiable finite system of set constraints has a regular solution. The space of runs of a tree set automaton can be viewed as a topological space, and analogs of these results hold in a more general context.

One might wish to incorporate these observations into a theory from which such results can be derived from general topological principles. This quest leads us to the definition of *rational spaces*. This definition is meant to capture the idea of a topological space with a regular or self-similar structure. The Cantor discontinuum is a simple example of such a space. Another example is provided by the space of runs of a Büchi automaton or tree set automaton. Sets of solutions to (finite) systems of set constraints can also be represented as (finitary) rational spaces.

Once this definition has been made and the basic theory developed, it is possible to rederive some of the results mentioned above from general principles, and in some cases give generalizations. For example, the result of [11] that every satisfiable finite system of set constraints has a regular solution is tantamount to the fact that every nonempty finitary rational space contains a rational point. (A *rational point* is a finitary singleton rational subspace.) In fact, every finitary rational space is a complete metric space, and is the completion of its rational points. The significance of this statement in terms of set constraints is that every finite system of set constraints is determined by its regular solutions.

This paper is organized as follows. In §2, we review the basic definitions of set constraints, termset algebras, regular sets, hypergraphs, and tree set automata. In §3 we introduce rational spaces, give several examples, and develop their basic theory, including the notions of rational maps, rational subspaces, rational products, and rational equivalence. In §4 we prove our main theorem, which characterizes the spaces of solutions of systems of set constraints in terms of rational spaces. In §5, we give several applications. Finally, in §6 we draw conclusions and discuss future work.

## 2 Preliminary Definitions

### 2.1 Set Expressions and Set Constraints

Let  $\Sigma$  be a finite ranked alphabet consisting of symbols  $f$ , each with an associated finite arity. Symbols in  $\Sigma$  of arity 0, 1, 2, and  $n$  are called *nullary*, *unary*, *binary*, and *n-ary*, respectively. Nullary elements are called *constants*. To avoid trivial special cases, we assume throughout that  $\Sigma$  contains at least one constant and at least one nonconstant. The use of the expression  $f(x_1, \dots, x_n)$  carries the implicit assumption that  $f$  is of arity  $n$ .

The set of ground terms over  $\Sigma$  is denoted  $T_\Sigma$ . If  $X = \{x, y, \dots\}$  is a set of variables, then  $T_\Sigma(X)$  denotes the set of terms over  $\Sigma$  and  $X$ , considering the elements of  $X$  as symbols of arity 0.

Let  $\mathbf{B} = (\cup, \cap, \sim, 0, 1)$  be the usual signature of Boolean algebra. Other Boolean operators such as  $-$  (set difference) and  $\oplus$  (symmetric difference) are defined as usual. Let  $\Sigma + \mathbf{B}$  denote the signature consisting of the disjoint union of  $\Sigma$  and  $\mathbf{B}$ . A *set expression* over variables  $X$  is any element of  $T_{\Sigma+\mathbf{B}}(X)$ .

A *positive set constraint* is a formal inclusion  $\varphi \subseteq \psi$ , where  $\varphi$  and  $\psi$  are set expressions. We also allow equational constraints  $\varphi = \psi$ , although inclusions and equations are interdefinable. A *negative set constraint* is the negation of a positive set constraint:  $\psi \not\subseteq \varphi$  or  $\varphi \neq \psi$ .

We interpret set expressions over the powerset  $2^{T_\Sigma}$  of  $T_\Sigma$ . This forms an algebra of signature  $\Sigma + \mathbf{B}$ , where the Boolean operators have their usual set-theoretic interpretations and elements  $f \in \Sigma$  are interpreted as functions

$$\begin{aligned} f &: (2^{T_\Sigma})^n \rightarrow 2^{T_\Sigma} \\ f(A_1, \dots, A_n) &= \{f(t_1, \dots, t_n) \mid t_i \in A_i, 1 \leq i \leq n\}. \end{aligned} \tag{1}$$

A *set valuation* is a map  $\sigma : X \rightarrow 2^{T_\Sigma}$  assigning a subset of  $T_\Sigma$  to each variable in  $X$ . Any set valuation  $\sigma$  extends uniquely to a  $(\Sigma + \mathbf{B})$ -homomorphism

$$\sigma : T_{\Sigma+\mathbf{B}}(X) \rightarrow 2^{T_\Sigma}$$

by induction on the structure of set expressions. We say that the set valuation  $\sigma$  *satisfies* the positive constraint  $\varphi \subseteq \psi$  if  $\sigma(\varphi) \subseteq \sigma(\psi)$ , and satisfies the negative constraint  $\varphi \not\subseteq \psi$  if  $\sigma(\varphi) \not\subseteq \sigma(\psi)$ . We write  $\sigma \models \alpha$  if the set valuation  $\sigma$  satisfies the constraint  $\alpha$ . A system  $\mathcal{S}$  of set constraints is *satisfiable* if there is a set valuation  $\sigma$  that satisfies all the constraints in  $\mathcal{S}$ ; in this case we write

$\sigma \models \mathcal{S}$  and say  $\sigma$  is a *solution* of  $\mathcal{S}$ . We write  $\mathcal{S} \models \alpha$  if  $\sigma \models \alpha$  whenever  $\sigma \models \mathcal{S}$ , and say  $\alpha$  is a *logical consequence* of  $\mathcal{S}$ . We write  $\mathcal{S} \models \mathcal{S}'$  if  $\mathcal{S} \models \alpha$  for all  $\alpha \in \mathcal{S}'$ . We say  $\mathcal{S}$  and  $\mathcal{S}'$  are *logically equivalent* if  $\mathcal{S} \models \mathcal{S}'$  and  $\mathcal{S}' \models \mathcal{S}$ .

## 2.2 Termset Algebras

*Termset algebras* were introduced in [19]. These are structures of signature  $\Sigma + \mathbf{B}$  satisfying axioms (2)–(6):

$$\text{axioms of Boolean algebra} \tag{2}$$

$$f(\dots, x \cup y, \dots) = f(\dots, x, \dots) \cup f(\dots, y, \dots) \tag{3}$$

$$f(\dots, x - y, \dots) = f(\dots, x, \dots) - f(\dots, y, \dots) \tag{4}$$

$$\bigcup_{f \in \Sigma} f(1, \dots, 1) = 1 \tag{5}$$

$$f(1, \dots, 1) \cap g(1, \dots, 1) = 0, \quad f \neq g. \tag{6}$$

A termset algebra is called *entire*<sup>1</sup> if it satisfies (7).

$$f(x_1, \dots, x_n) = 0 \implies \bigvee_{i=1}^n (x_i = 0). \tag{7}$$

The ellipses in (3) and (4) indicate that the explicitly given arguments occur in corresponding places, and that implicit arguments in corresponding places agree.

The algebra  $2^{T_\Sigma}$  discussed in §2.1 forms a model of these axioms. Another model is given by the subalgebra of regular subsets of  $T_\Sigma$  described in §2.3 below.

Some immediate consequences of (2)–(6) are

$$f(\dots, 0, \dots) = 0 \tag{8}$$

$$f(\dots, \sim x, \dots) = f(\dots, 1, \dots) - f(\dots, x, \dots) \tag{9}$$

$$f(\dots, x \oplus y, \dots) = f(\dots, x, \dots) \oplus f(\dots, y, \dots) \tag{10}$$

$$f(\dots, x \cap y, \dots) = f(\dots, x, \dots) \cap f(\dots, y, \dots) \tag{11}$$

$$f(x_1 \cap y_1, \dots, x_n \cap y_n) = f(x_1, \dots, x_n) \cap f(y_1, \dots, y_n) \tag{12}$$

$$x \subseteq y \implies f(\dots, x, \dots) \subseteq f(\dots, y, \dots) \tag{13}$$

$$\sim f(x_1, \dots, x_n) = \bigcup_{g \neq f} g(1, \dots, 1)$$

<sup>1</sup> The term *closed* was used in [19]. This terminology will be unsuitable in the present context because of the potential of confusion with topological closure.

$$\cup \bigcup_{i=1}^n f(\underbrace{1, \dots, 1}_{i-1}, \sim x_i, \underbrace{1, \dots, 1}_{n-i}) . \quad (14)$$

Property (14) is called the *generalized DeMorgan law*.

Let  $X$  be a fixed set of variables. Let  $F$  be a subset of  $T_{\Sigma+\mathbb{B}}(X)$  closed under subexpressions. Let  $F'$  denote the set of conjunctions of elements of  $F$  and their negations. A *literal* is a variable or a negation of a variable.

**Proposition 1** *Any set constraint all of whose subexpressions are in  $F$  is equivalent under the termset algebra axioms (2)–(6) to a finite system of constraints of the form  $\varphi \cap f(\varphi_1, \dots, \varphi_n) = 0$ , where  $f \in \Sigma$ ,  $\varphi, \varphi_1, \dots, \varphi_n \in F'$ , and  $\varphi$  is a conjunction of literals.*

**PROOF.** If the constraint is an equation, write it as two inclusions. Write an inclusion  $\varphi \subseteq \psi$  as  $\varphi \cap \sim \psi = 0$ . Use the DeMorgan laws and the law  $\sim \sim \varphi = \varphi$  of Boolean algebra to push negations occurring outside the scope of any  $f \in \Sigma$  down until any such negation is applied to a variable or an expression  $f(\varphi_1, \dots, \varphi_n)$ . Use the generalized DeMorgan law (14) at the outermost level only to transform the expression  $\sim f(\varphi_1, \dots, \varphi_n)$  to a disjunction of expressions  $g(\psi_1, \dots, \psi_m)$ , where each  $\psi_i$  is either  $\sim \varphi_j$  or 1. Note both of these expressions are in  $F'$  (1 is the null conjunction). The expression is now a monotone Boolean combination of literals and expressions  $f(\varphi_1, \dots, \varphi_n)$ , where  $\varphi_1, \dots, \varphi_n \in F'$ . Use the distributive laws of Boolean algebra outside the scope of any  $f \in \Sigma$  only to transform the expression to disjunctive normal form. Break the resulting disjunction up into several constraints, one for each disjunct, using the rule  $\varphi \cup \psi = 0 \iff \varphi = 0 \wedge \psi = 0$ . Discard any constraint containing  $f(\varphi_1, \dots, \varphi_n)$  and  $g(\psi_1, \dots, \psi_m)$ ,  $f \neq g$ , because the constraint is automatically true by (6). In each of the remaining constraints, combine all non-literals into one expression using (12); i.e., replace  $f(\varphi_1, \dots, \varphi_n) \cap f(\psi_1, \dots, \psi_n)$  with  $f(\varphi_1 \cap \psi_1, \dots, \varphi_n \cap \psi_n)$ . Replace any constraint  $\varphi = 0$ , where  $\varphi$  is a conjunction of literals alone, with the constraints  $\varphi \cap f(1, \dots, 1) = 0$  for all  $f \in \Sigma$ , as justified by (5). The resulting constraints are of the desired form.  $\square$

### 2.3 Regular Sets

A subset of  $T_{\Sigma}$  is *regular* if it is described by a finite bottom-up tree automaton [10]; equivalently, if it is some set  $x_1$  described by a system of simultaneous set equations of the form

$$x_i = \varphi_i(x_1, \dots, x_m) , \quad 1 \leq i \leq n \quad (15)$$

in which each variable  $x_i$  occurs on the left hand side of exactly one equation and each right hand side is a disjunction of set expressions of the form  $f(y_1, \dots, y_n)$ , where  $f \in \Sigma$  and  $y_i \in \{x_1, \dots, x_m\}$ ,  $1 \leq i \leq n$ . It can be proved by induction on the depth of terms that any such system has a unique solution. For example, the system

$$x = a \cup g(y) \quad y = g(x) \quad (16)$$

has the unique regular solution

$$\sigma(x) = \{g^n(a) \mid n \text{ even}\} \quad \sigma(y) = \{g^n(a) \mid n \text{ odd}\} .$$

#### 2.4 Tree Set Automata and Hypergraphs

*Tree set automata* were introduced in [11]. They are strongly related to the *hypergraphs* introduced in [1], the only essential difference being the inclusion of an acceptance condition in the former. The relation of finite tree set automata and finite hypergraphs to set constraints has been well studied [1,2,11,12,21,28].

**Definition 2** Let  $\Sigma$  be a finite ranked alphabet. A  $\Sigma$ -hypergraph is a pair  $(D, E)$ , where  $D$  is a set (not necessarily finite) and  $E$  is an indexed collection of hyperedges

$$E_f : D^n \rightarrow 2^D, \quad n = \text{arity}(f),$$

one for each  $f \in \Sigma$ .

Thus  $E_a$  gives a subset of  $D$  for constants  $a$ ,  $E_g$  is an ordinary binary edge relation for unary  $g$ , etc.

**Definition 3** A hypergraph  $(D, E)$  is said to be *entire*<sup>2</sup> if every  $E_f(d_1, \dots, d_n)$  is nonempty, *deterministic* if every  $E_f(d_1, \dots, d_n)$  is a singleton, and *unrestricted* if every  $E_f(d_1, \dots, d_n) = D$ .

**Definition 4** A run of the hypergraph  $(D, E)$  is a map  $\theta : T_\Sigma \rightarrow D$  such that for all  $f(t_1, \dots, t_n) \in T_\Sigma$ ,

$$\theta(f(t_1, \dots, t_n)) \in E_f(\theta(t_1), \dots, \theta(t_n)) . \quad (17)$$

<sup>2</sup> The term *closed* was used in [1,2,19,21,28]. This also corresponds to the condition  $COND_\omega$  of [11].

There exists a run in  $(D, E)$  if and only if  $(D, E)$  has an entire induced subhypergraph: an entire induced subhypergraph on  $D' \subseteq D$  allows the definition of a run  $\theta : T_\Sigma \rightarrow D'$  by induction; conversely, for any run  $\theta$ , the induced subhypergraph on the image  $\theta(T_\Sigma)$  is entire.

**Definition 5** *Let  $X$  be a set of variables. A  $\Sigma, X$ -tree set automaton  $\mathcal{M}$  is a tuple  $(D, E, A)$ , where  $(D, E)$  is a  $\Sigma$ -hypergraph and  $A$  is an indexed family  $A_x, x \in X$  of subsets of  $D$ , called accept sets. A tuple  $L_x, x \in X$  of subsets of  $T_\Sigma$  is accepted by  $\mathcal{M}$  if there exists a run  $\theta$  such that  $L_x = \{t \mid \theta(t) \in A_x\}, x \in X$ .*

In previous work [1,2,11,12,21,28],  $D$  and  $X$  are assumed to be finite, but we will find it convenient not impose these restrictions.

### 3 Rational Spaces

#### 3.1 Definition of Rational Spaces

Let  $\Sigma$  be a fixed finite ranked alphabet.

**Definition 6** *A topological  $\Sigma$ -hypergraph is a  $\Sigma$ -hypergraph  $(D, E)$ , finite or infinite, endowed with a topology on  $D$  whose hyperedges*

$$\{(d, d_1, \dots, d_n) \mid d \in E_f(d_1, \dots, d_n)\}$$

*are closed in the product topology on  $D^{n+1}$ .*

**Definition 7** *A space of runs over  $\Sigma$  is the space  $\mathcal{R}(\mathcal{D})$  of runs of a topological  $\Sigma$ -hypergraph  $\mathcal{D}$ , where the topology on  $\mathcal{R}(\mathcal{D})$  is inherited from the product topology on  $\mathcal{D}^{T_\Sigma}$ . The space  $\mathcal{R}(\mathcal{D})$  is called finitary if  $\mathcal{D}$  is finite.*

Recall that the product topology on  $\mathcal{D}^{T_\Sigma}$  is the smallest topology such that all projections  $\pi_t : \theta \mapsto \theta(t)$  are continuous. In other words, it is generated by subbasic open sets

$$\{\theta \mid \theta(t) \in x\}, \quad t \in T_\Sigma, \quad x \text{ open in } \mathcal{D}. \quad (18)$$

The space  $\mathcal{R}(\mathcal{D})$  of runs of  $\mathcal{D}$ , being a space of functions  $T_\Sigma \rightarrow \mathcal{D}$ , is a subspace of this space. The topology is thus generated by subbasic open sets (18) restricted to  $\mathcal{R}(\mathcal{D})$ .

**Definition 8** A rational space is a space of runs  $\mathcal{R}(\mathcal{D})$  such that  $\mathcal{D}$  is compact and Hausdorff.

We remark here that rational spaces are more than just Cantor spaces; the representation in terms of  $\mathcal{D}$  is germane. Formally, this will be reflected in the restricted class of morphisms defined below, called *rational maps*. First, however, let us look at some examples of rational spaces.

**Example 9** A simple but revealing example is the Cantor discontinuum. This space can be constructed by starting from the closed real unit interval  $[0, 1]$  and applying the following operation: delete the “middle third”, i.e., the open interval  $(\frac{1}{3}, \frac{2}{3})$ ; then delete the middle thirds of the remaining intervals; and so on. The Cantor discontinuum consists of points that are never deleted.



Equivalently, these are all real numbers in the unit interval whose ternary (base 3) expansion does not contain the digit 1. This is a classical example of an uncountable nowhere dense set of reals.

The topology on the Cantor discontinuum is inherited from the usual topology on the real line. With that topology, it is homeomorphic to the topological product  $2^\omega$ , or the space of infinite paths of the infinite binary tree with the intervals as basic open sets, where in this context an interval is a set consisting of all extensions of some finite prefix.

The Cantor discontinuum has a representation as a finitary rational space over  $\Sigma = \{a, g\}$ , where  $a$  is a constant and  $g$  is unary. For  $\mathcal{D}$  we take the discrete space  $\{0, 1, 2\}$  with  $E_g(d) = E_a(\ ) = \{0, 2\}$ . Corresponding to each point  $x$  we take the run whose value on  $g^n(a)$  is the  $n^{\text{th}}$  digit in the ternary expansion of  $x$ .

**Example 10** Consider an infinite tree that is binary branching at even levels and ternary branching at odd levels. The basic open sets are the intervals. Equivalently, take the subspace of the real unit interval consisting of all numbers whose ternary expansion contains no 2 in an even position. This space has a representation as a finitary rational space with  $\Sigma = \{a, g\}$ ,  $\mathcal{D}$  the discrete space  $\{0, 1, 2\} \times \{0, 1\}$ , and

$$E_a(\ ) = E_g((d, 1)) = \{(0, 0), (1, 0)\}$$

$$E_g((d, 0)) = \{(0, 1), (1, 1), (2, 1)\} .$$



The value on  $g^n(a)$  of the run corresponding to  $x$  is  $(d, n \bmod 2)$ , where  $d$  is the  $n^{\text{th}}$  digit in the ternary expansion of  $x$ .

Examples (9) and (10) exhibit a treelike structure, because there is only one nonconstant symbol and it is unary. Already with two unary symbols, this intuitive picture is no longer accurate.

**Example 11** *The following is an example of a non-finitary rational space. Let  $X$  be a set of variables ranging over subsets of  $T_\Sigma$ , and consider the family of set valuations  $\sigma : X \rightarrow 2^{T_\Sigma}$ . Endow this set with smallest topology whose closed sets include all sets of the form  $\{\sigma \mid \sigma \models \mathcal{S}\}$  for  $\mathcal{S}$  a system of set constraints over  $X$ . One can show that the topology on this space is generated by subbasic clopen (closed and open) sets  $\{\sigma \mid \sigma \models t \subseteq x\}$  and  $\{\sigma \mid \sigma \models t \subseteq \sim x\}$  for  $t \in T_\Sigma$  and  $x \in X$ . It was shown in [2] that this space is compact.*

We will show in §4 below that this space has the following representation as a rational space. Take  $\mathcal{D}$  to be the topological product  $2^X$  with  $E_f(d_1, \dots, d_n) = \mathcal{D}$ . (If  $X$  is countable,  $\mathcal{D}$  is just the Cantor discontinuum.) Associate with every set valuation  $\sigma$  the run

$$\lambda t. \lambda x. \begin{cases} 1, & \text{if } t \in \sigma(x) \\ 0, & \text{otherwise.} \end{cases} \quad (19)$$

This space is not finitary unless  $X$  is finite. However, it is a product of finitary spaces, one for each  $x \in X$ , as can be seen by reversing the binding order of  $t$  and  $x$  in the  $\lambda$ -expression (19).

**Example 12** *Every termset algebra has a set-theoretic representation as a topological term automaton whose topology is Stone [19]. These automata were introduced in [23,24]. They differ from tree set automata in that they are top-down instead of bottom-up, but they have a naturally defined hypergraph structure in which the hyperedges are closed, giving rational spaces.*

### 3.2 Basic Properties

**Proposition 13** *Let  $\mathcal{R}(\mathcal{D})$  be a space of runs.*

- (i) *If  $\mathcal{D}$  is Hausdorff, then so is  $\mathcal{R}(\mathcal{D})$ .*
- (ii) *If  $\mathcal{D}$  is compact, then so is  $\mathcal{R}(\mathcal{D})$ .*
- (iii) *If  $\mathcal{D}$  is finite and discrete (all sets are open), then  $\mathcal{R}(\mathcal{D})$  is a complete metric space (all Cauchy sequences converge) under the metric*

$$d(\theta, \eta) = 2^{-\text{depth}(t)} , \quad (20)$$

where  $t$  is a term of minimal depth on which  $\theta$  and  $\eta$  differ, or 0 if no such term exists.

**PROOF.** (i) Any product of Hausdorff spaces is Hausdorff and any subspace of a Hausdorff space is Hausdorff, and  $\mathcal{R}(\mathcal{D})$  is a subspace of the product  $\mathcal{D}^{T_\Sigma}$ .

(ii) Regarding  $E_f$  as an  $(n + 1)$ -ary relation, we can write

$$\mathcal{R}(\mathcal{D}) = \bigcap_{f(t_1, \dots, t_n) \in T_\Sigma} \{ \theta \in \mathcal{D}^{T_\Sigma} \mid (\theta(f(t_1, \dots, t_n)), \theta(t_1), \dots, \theta(t_n)) \in E_f \} .$$

The set in this conjunction corresponding to  $f(t_1, \dots, t_n)$  is the continuous preimage of the closed set  $E_f$  under the projection  $(\pi_{f(t_1, \dots, t_n)}, \pi_{t_1}, \dots, \pi_{t_n}) : \mathcal{D}^{T_\Sigma} \rightarrow D^{n+1}$ , thus is closed. By Tychonoff's Theorem,  $\mathcal{D}^{T_\Sigma}$  is compact, and any closed subspace of a compact space is compact, therefore  $\mathcal{R}(\mathcal{D})$  is compact.

(iii) It is easily verified that  $d(\theta, \eta) = 0$  iff  $\theta = \eta$ ,  $d(\theta, \eta) = d(\eta, \theta)$ , and the triangle inequality holds. The two topologies coincide: every basic open neighborhood in the metric topology is open in the product topology, since

$$N_\epsilon(\theta) = \{ \eta \mid d(\eta, \theta) < \epsilon \} = \bigcap_{\text{depth}(t) \leq -\log \epsilon} \{ \eta \mid \eta(t) = \theta(t) \} ,$$

and vice-versa, since if  $\theta(t) = u$ , then

$$\{ \eta \mid \eta(t) = u \} \supseteq \{ \eta \mid d(\eta, \theta) < 2^{-\text{depth}(t)} \} = N_{2^{-\text{depth}(t)}}(\theta) .$$

The space  $\mathcal{R}(\mathcal{D})$  is complete because it is compact by (ii), and any compact metric space is complete.  $\square$

**Corollary 14** *Every rational space is compact and Hausdorff. Every finitary rational space is a complete metric space under the metric (20).*

### 3.3 Rational Maps

The spaces of Examples 9 and 10 are homeomorphic, since both spaces are compact Hausdorff spaces with countable clopen bases and no isolated points, and all such spaces are homeomorphic. This is a consequence under Stone duality of the fact that all countable atomless Boolean algebras are isomorphic. Indeed, the space of paths in a tree that is binary branching at prime levels

and ternary branching at nonprime levels is also homeomorphic to the Cantor space, but it is not clear at all how to assign a finitary rational space structure to it.

The homeomorphisms relating these spaces apparently do not preserve the structure inherent in the representation as spaces of runs. Thus the relation of *homeomorphism* is too coarse. Reflecting on this observation, one is led to the realization that rational spaces cannot be defined independent of some representation; i.e., there is no purely topological definition.

These observations motivate the definition of a restricted class of maps called *rational maps*, which are continuous maps preserving the rational structure. We take these maps as our morphisms in the category of rational spaces.

**Definition 15** *Let  $\mathcal{R}(\mathcal{D})$  and  $\mathcal{R}(\mathcal{E})$  be rational spaces over  $\Sigma$ . A rational map  $\mathcal{R}(\mathcal{D}) \rightarrow \mathcal{R}(\mathcal{E})$  is a function  $\hat{h} : \theta \mapsto h \circ \theta$  induced by a continuous map  $h : \mathcal{D} \rightarrow \mathcal{E}$  such that*

$$h(E_f^{\mathcal{D}}(d_1, \dots, d_n)) \subseteq E_f^{\mathcal{E}}(h(d_1), \dots, h(d_n)) . \quad (21)$$

*A rational map  $\mathcal{R}(\mathcal{D}) \rightarrow \mathcal{R}(\mathcal{E})$  is called a rational embedding if it is injective, and a refinement if it is bijective.*

Note that the terms “injective” and “bijective” in this definition refer to  $\hat{h}$ , not to  $h$ . There exist rational embeddings and refinements in which  $h$  is not bijective.

Any rational embedding or refinement is a homeomorphism between its domain and image, since any continuous bijection from a compact space to a Hausdorff space is a homeomorphism.

We mention in particular two special kinds of rational embeddings:

**Narrowing** If  $\mathcal{D} = (D, E)$  and  $\mathcal{D}' = (D', E')$  are two hypergraphs on the same set of vertices  $D$ , and if all  $E_f(d_1, \dots, d_n) \subseteq E'_f(d_1, \dots, d_n)$ , then the identity map  $D \rightarrow D$  induces an embedding  $\mathcal{R}(\mathcal{D}) \rightarrow \mathcal{R}(\mathcal{D}')$ . Such an  $\mathcal{R}(\mathcal{D})$  is called a *narrowing* of  $\mathcal{R}(\mathcal{D}')$ .

**Induced Subspaces** If  $\mathcal{D} = (D, E)$  is the induced subhypergraph of  $\mathcal{D}' = (D', E')$  on some subset  $D \subseteq D'$ , i.e. if  $E_f(d_1, \dots, d_n) = E'_f(d_1, \dots, d_n) \cap D$  for all  $d_1, \dots, d_n \in D$ , then the inclusion map  $D \rightarrow D'$  induces an embedding  $\mathcal{R}(\mathcal{D}) \rightarrow \mathcal{R}(\mathcal{D}')$ . Such an  $\mathcal{R}(\mathcal{D})$  is called an *induced subspace* of  $\mathcal{R}(\mathcal{D}')$ .

**Example 16** The map  $(d, 0) \mapsto 0$ ,  $(d, 1) \mapsto 2$  induces a rational map from the space of Example 10 into the space of Example 9. The image consists of the single point  $.020202\dots$ . This is not a rational embedding, since it is not injective.

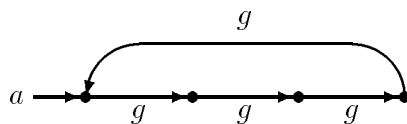
**Example 17** The hypergraph with vertices  $\{0, 1, 2\} \times \{0, 1\}$  and hyperedges

$$\begin{aligned} E_a(\ ) &= E_g((d, 1)) = \{(0, 0)\} \\ E_g((d, 0)) &= \{(0, 1), (2, 1)\} \end{aligned}$$

constitutes a narrowing of the space of Example 10. This is the subspace consisting of all numbers whose ternary expansion contains no 1 or 2 in an even position or 1 in an odd position.

**Example 18** The induced subhypergraph of  $\{0, 1, 2\} \times \{0, 1\}$  on  $\{0, 1\} \times \{0, 1\}$  generates an induced subspace of the space of Example 10. This is the subspace of all numbers whose ternary expansion contains no 2.

**Example 19** Here is an example of a rational embedding which is neither a narrowing nor an induced subspace. Let  $\Sigma = \{a, g\}$ , where  $a$  is a constant and  $g$  is unary. Consider the following hypergraph:



Map the first three vertices to 0 and the last vertex to 2. This gives a rational embedding of a singleton finitary rational space into the Cantor discontinuum. The image is the rational number  $.000200020002\dots$ .

**Example 20** The induced subspace on the subhypergraph  $\{0, 2\}$  is a refinement of the Cantor discontinuum, as is the subspace of Example 18 under the map  $(d, c) \mapsto 2d$ .

### 3.4 Rational Products

**Definition 21** The rational product of any indexed family  $\mathcal{R}(\mathcal{D}_x)$  of rational spaces is the space  $\mathcal{R}(\prod_x \mathcal{D}_x)$ , where

$$E_f^{\prod_x \mathcal{D}_x}(d_1, \dots, d_n) = \prod_x E_f^{\mathcal{D}_x}(\pi_x(d_1), \dots, \pi_x(d_n)) .$$

Intuitively, the runs in each factor space can develop independently. The topological product  $\prod_x \mathcal{D}_x$  is Hausdorff and compact by Tychonoff's Theorem, and

$E_f^{\Pi_x \mathcal{D}_x}$  is closed in the product topology, thus  $\mathcal{R}(\Pi_x \mathcal{D}_x)$  is a rational space.

Strictly speaking, the rational product  $\mathcal{R}(\Pi_x \mathcal{D}_x)$  is not the same as the set-theoretic (Cartesian) product  $\Pi_x \mathcal{R}(\mathcal{D}_x)$ , although the two are in one-to-one correspondence, as can be seen by exchanging the order of  $\lambda$ -bindings as in Example 11.

The projections  $\pi_y : \Pi_x \mathcal{D}_x \rightarrow \mathcal{D}_y$  induce projections  $\hat{\pi}_y : \mathcal{R}(\Pi_x \mathcal{D}_x) \rightarrow \mathcal{R}(\mathcal{D}_y)$ , which are rational maps.

A finite rational product of finitary spaces is finitary.

### 3.5 Rational Subspaces and Rational Points

**Definition 22** A rational subspace of a rational space is any embedded image of another rational space. In other words, a subspace  $\mathcal{R}$  of a rational space  $\mathcal{R}(\mathcal{E})$  is a rational subspace if there exists a rational space  $\mathcal{R}(\mathcal{D})$  and a rational embedding  $\hat{h} : \mathcal{R}(\mathcal{D}) \rightarrow \mathcal{R}(\mathcal{E})$  such that  $\mathcal{R} = \hat{h}(\mathcal{R}(\mathcal{D}))$ .

A rational subspace is entire if it is the embedded image of a rational space defined on an entire hypergraph.

**Definition 23** A rational point of a rational space is a finitary singleton rational subspace.

**Example 24** The rational points of the Cantor discontinuum are the rational numbers, i.e., real numbers whose ternary expansion is ultimately periodic.

**Proposition 25** Let  $\mathcal{R}(\mathcal{D})$  be any rational space. For any closed set  $x \subseteq \mathcal{D}$  and  $t \in T_\Sigma$ , the closed set  $\pi_t^{-1}(x) = \{\theta \mid \theta(t) \in x\}$  is a rational subspace of  $\mathcal{R}(\mathcal{D})$ . Moreover, if  $\mathcal{R}(\mathcal{D})$  is finitary, then so is  $\pi_t^{-1}(x)$ .

**PROOF.** Let  $F$  be the set of subterms of  $t$ . Let  $*$  be a new element not in  $F$ , and let  $D_* = F \cup \{*\}$ . Let  $\mathcal{D}_*$  be the discrete hypergraph on this set with hyperedges

$$E_f(t_1, \dots, t_n) = \begin{cases} \{f(t_1, \dots, t_n)\}, & \text{if } f(t_1, \dots, t_n) \text{ is a subterm of } t \\ , & \text{otherwise.} \end{cases}$$

Now take the rational product  $\mathcal{R}(\mathcal{D}_*) \times \mathcal{R}(\mathcal{D})$ , then take the induced subspace obtained by discarding the open set consisting all vertices of the form  $(t, d)$  for  $d \notin x$ .  $\square$

**Proposition 26** *The intersection of two (finitary) rational subspaces is again a (finitary) rational subspace.*

**PROOF.** Let  $h_1 : \mathcal{D}_1 \rightarrow \mathcal{E}$ ,  $h_2 : \mathcal{D}_2 \rightarrow \mathcal{E}$  be maps inducing rational maps  $\hat{h}_1 : \mathcal{R}(\mathcal{D}_1) \rightarrow \mathcal{R}(\mathcal{E})$ ,  $\hat{h}_2 : \mathcal{R}(\mathcal{D}_2) \rightarrow \mathcal{R}(\mathcal{E})$ . Construct the rational product  $\mathcal{R}(\mathcal{D}_1) \times \mathcal{R}(\mathcal{D}_2)$  with projections  $\pi_i : \mathcal{D}_1 \times \mathcal{D}_2 \rightarrow \mathcal{D}_i$ ,  $i \in \{1, 2\}$ , and restrict to the diagonal set  $\mathcal{C} = \{(d_1, d_2) \mid h_1(d_1) = h_2(d_2)\}$ . Let  $\mathcal{R}(\mathcal{C})$  be the induced rational subspace. The map  $h((d_1, d_2)) = h_1(d_1) = h_2(d_2)$  gives the desired intersection.  $\square$

**Definition 27** *The essential subspace of a rational space  $\mathcal{R}(\mathcal{D})$  is the induced subspace on the set  $\bigcup_{\theta \in \mathcal{R}(\mathcal{D})} \theta(T_\Sigma)$ .*

In other words, we discard vertices not appearing in any run. The essential subspace of  $\mathcal{R}(\mathcal{D})$  is a refinement of  $\mathcal{R}(\mathcal{D})$ .

### 3.6 Rational Equivalence

**Definition 28** *Two spaces are rationally equivalent if they have a common refinement.*

**Proposition 29** *Rational equivalence is transitive.*

**PROOF.** It suffices to show the following *Diamond Lemma*: if  $\mathcal{R}(\mathcal{D}_1)$  and  $\mathcal{R}(\mathcal{D}_2)$  are both refinements of  $\mathcal{R}(\mathcal{E})$ , then  $\mathcal{R}(\mathcal{D}_1)$  and  $\mathcal{R}(\mathcal{D}_2)$  have a common refinement. Let  $\mathcal{R}(\mathcal{C})$  be the space constructed in the proof of Proposition 26. The rational maps  $\mathcal{R}(\mathcal{C}) \rightarrow \mathcal{R}(\mathcal{D}_1)$  and  $\mathcal{R}(\mathcal{C}) \rightarrow \mathcal{R}(\mathcal{D}_2)$  induced by the restriction of  $\pi_1$  and  $\pi_2$  to domain  $\mathcal{C}$  are bijective, thus  $\mathcal{R}(\mathcal{C})$  is the desired common refinement.  $\square$

## 4 A Characterization of Set Constraints

In this section we give a complete characterization of the sets of solutions to systems of set constraints in terms of rational spaces. Our main result is a one-to-one correspondence, up to logical equivalence on one side and rational equivalence preserving  $X$  on the other, between (finite) systems of set constraints over variables  $X$  and certain (finitary) subspaces of a certain rational space (Theorem 31).

Let  $\Sigma$  be a fixed finite ranked alphabet and let  $X$  be a fixed set of variables (finite or infinite). Let  $\mathcal{S}$  be a finite system of set constraints over  $\Sigma$  and  $X$ . In [1] it was shown how to construct a finite hypergraph  $(D, E)$  whose runs are in one-to-one correspondence with the solutions  $\sigma : X \rightarrow 2^{T_\Sigma}$  of  $\mathcal{S}$ , where  $X$  is the set of variables occurring in  $\mathcal{S}$ . This result is also implicit in [11]. One can construct  $(D, E)$  by various alternative methods [2,19,21]. Here is yet another method.

Let  $F$  be any subset of  $T_{\Sigma+\mathbf{B}}(X)$  (finite or infinite) containing  $X$  and closed under subexpressions. Let  $2^F$  be the topological product of  $F$  copies of the discrete space  $2 = \{0, 1\}$ .

The space of all unrestricted runs  $\theta : T_\Sigma \rightarrow 2^F$  is homeomorphic to the space of all functions  $\sigma : F \rightarrow 2^{T_\Sigma}$ , taking the product topology on each. This can be seen by writing

$$\theta : T_\Sigma \rightarrow F \rightarrow 2 \quad \sigma : F \rightarrow T_\Sigma \rightarrow 2 \quad (22)$$

and exchanging the order of  $\lambda$ -binding. Their respective topologies are generated by subbasic clopen sets

$$\{\theta \mid \theta(t)(\varphi) = b\} \quad \{\sigma \mid \sigma(\varphi)(t) = b\}$$

for  $b \in \{0, 1\}$ ,  $t \in T_\Sigma$  and  $\varphi \in F$ .

Note that the set of all functions  $\sigma : F \rightarrow 2^{T_\Sigma}$  includes some that are not partial set valuations. In order to be a partial set valuation,  $\sigma$  must be a  $(\Sigma + \mathbf{B})$ -homomorphism on its domain, i.e., must satisfy the following requirements for expressions in  $F$ :

$$\sigma(\varphi \cup \psi) = \sigma(\varphi) \cup \sigma(\psi) \quad (23)$$

$$\sigma(\sim \varphi) = \sim \sigma(\varphi) \quad (24)$$

$$\sigma(f(\varphi_1, \dots, \varphi_n)) = \{f(t_1, \dots, t_n) \mid t_i \in \sigma(\varphi_n), 1 \leq i \leq n\}. \quad (25)$$

Likewise, the set of functions  $\theta : T_\Sigma \rightarrow 2^F$  is the unrestricted rational space  $\mathcal{R}(2^F)$ . Let us narrow this space as follows.

Let  $\mathcal{S}$  be any system of set constraints all of whose subexpressions are in  $F$ . Let  $F'$  denote the set of conjunctions of elements of  $F$  and their negations. For each  $d \in 2^F$ , let  $d' \subseteq F'$  be the smallest set of expressions including  $d$  and  $\{\sim \varphi \mid \varphi \in F - d\}$  and closed under conjunction. Call a set expression  $\varphi$   *$\mathcal{S}$ -consistent* if  $\varphi = 0$  is not a logical consequence of  $\mathcal{S}$  and the equational axioms (2)–(6) of termset algebra. Define the hyperedges

$$E_f^{\mathcal{S}}(d_1, \dots, d_n) = \{d \mid \text{for every } \varphi \in d' \text{ and } \varphi_i \in d'_i, 1 \leq i \leq n, \\ \varphi \cap f(\varphi_1, \dots, \varphi_n) \text{ is } \mathcal{S}\text{-consistent}\} .$$

The  $(n + 1)$ -ary relation  $E_f^{\mathcal{S}}$  is closed, as is any set

$$\{u \mid \text{for all basic open neighborhoods } x \text{ of } u, P(x)\}$$

in any topological space for any property  $P$  of open sets whatsoever. Denote the resulting rational space by  $\mathcal{R}(2^F, \mathcal{S})$ .

**Theorem 30** *Let the function  $\sigma : F \rightarrow 2^{T_{\Sigma}}$  correspond to the unrestricted run  $\theta : T_{\Sigma} \rightarrow 2^F$  under the correspondence (22). Let  $\mathcal{S}$  be any system of set constraints all of whose subexpressions are in  $F$ . Then  $\sigma$  is a partial set valuation satisfying  $\mathcal{S}$  if and only if  $\theta$  is a run of  $\mathcal{R}(2^F, \mathcal{S})$ .*

**PROOF.** If  $\sigma$  is a partial set valuation on  $F$ , then it extends uniquely by induction to a total set valuation  $\sigma' : T_{\Sigma+B}(X) \rightarrow 2^{T_{\Sigma}}$ . Such a map is a  $(\Sigma + \mathbf{B})$ -homomorphism and its image is a subalgebra  $\mathcal{A}$  of the set-theoretic termset algebra  $2^{T_{\Sigma}}$ . Moreover, since  $\sigma \models \mathcal{S}$ ,  $\mathcal{A}$  satisfies the equations and inequalities in  $\mathcal{S}$ .

Let  $f(t_1, \dots, t_n)$  be any ground term. For any  $\varphi \in \theta(f(t_1, \dots, t_n))'$  and  $\varphi_i \in \theta(t_i)'$ ,  $1 \leq i \leq n$ , by (22) we have  $f(t_1, \dots, t_n) \in \sigma'(\varphi)$  and  $t_i \in \sigma'(\varphi_i)$ ,  $1 \leq i \leq n$ , thus  $f(t_1, \dots, t_n) \in \sigma'(\varphi \cap f(\varphi_1, \dots, \varphi_n))$ . Then  $\varphi \cap f(\varphi_1, \dots, \varphi_n)$  is  $\mathcal{S}$ -consistent, since it is nonzero in the termset algebra  $\mathcal{A}$ . Since the  $\varphi, \varphi_1, \dots, \varphi_n$  were arbitrary,  $\theta(f(t_1, \dots, t_n)) \in E_f^{\mathcal{S}}(\theta(t_1), \dots, \theta(t_n))$ , and since  $f(t_1, \dots, t_n)$  was arbitrary,  $\theta$  is a run of  $\mathcal{R}(2^F, \mathcal{S})$ .

Conversely, suppose  $\theta$  is a run of  $\mathcal{R}(2^F, \mathcal{S})$ . Under the correspondence (22), the properties (23)–(25) become

$$\varphi \cup \psi \in \theta(t) \iff \varphi \in \theta(t) \text{ or } \psi \in \theta(t) \quad (26)$$

$$\sim \varphi \in \theta(t) \iff \varphi \notin \theta(t) \quad (27)$$

$$f(\varphi_1, \dots, \varphi_n) \in \theta(t) \iff \exists t_1, \dots, t_n \ t = f(t_1, \dots, t_n) \\ \text{and } \varphi_i \in \theta(t_i), \quad 1 \leq i \leq n \quad (28)$$

for expressions  $\varphi \cup \psi$ ,  $\sim \varphi$ , and  $f(\varphi_1, \dots, \varphi_n) \in F$ . The first two of these follow immediately from fact that  $\theta(t)'$  is consistent with the laws of Boolean algebra (2). For (28), suppose first that  $f(\varphi_1, \dots, \varphi_n) \in \theta(t)$ . For any  $g \in \Sigma$ ,

$$t = g(t_1, \dots, t_m) \implies \theta(t) \in E_g^{\mathcal{S}}(\theta(t_1), \dots, \theta(t_m)) \\ \implies f(\varphi_1, \dots, \varphi_n) \cap g(1, \dots, 1) \text{ is } \mathcal{S}\text{-consistent} \\ \implies m = n \text{ and } f = g \quad \text{by (6).}$$



Thus  $t = f(t_1, \dots, t_n)$  and  $\theta(t) \in E_f^S(\theta(t_1), \dots, \theta(t_n))$ . Now each  $\theta(t_i)'$  contains some  $\psi_i \in \{\varphi_i, \sim \varphi_i\}$ , and  $f(\varphi_1, \dots, \varphi_n) \cap f(\psi_1, \dots, \psi_n)$  must be  $\mathcal{S}$ -consistent. By properties (8) and (12), the only  $\mathcal{S}$ -consistent possibility is  $\varphi_i = \psi_i$ ,  $1 \leq i \leq n$ , thus  $\varphi_i \in \theta(t_i)$ ,  $1 \leq i \leq n$ .

For the other direction of (28), suppose  $\varphi_i \in \theta(t_i)$ ,  $1 \leq i \leq n$ . Since  $\theta$  is a run, for any  $\psi \in \theta(f(t_1, \dots, t_n))'$ ,  $\psi \cap f(\varphi_1, \dots, \varphi_n)$  is  $\mathcal{S}$ -consistent. But  $\theta(f(t_1, \dots, t_n))'$  contains at least one of  $f(\varphi_1, \dots, \varphi_n)$ ,  $\sim f(\varphi_1, \dots, \varphi_n)$ , and the former is the only  $\mathcal{S}$ -consistent choice. Therefore  $f(\varphi_1, \dots, \varphi_n) \in \theta(f(t_1, \dots, t_n))$ .

Finally we show that  $\sigma \models \mathcal{S}$ . If not, then by Proposition 1, there exists a logical consequence  $\varphi \cap f(\varphi_1, \dots, \varphi_n) = 0$  of  $\mathcal{S}$ ,  $\varphi, \varphi_1, \dots, \varphi_n \in F'$ , such that  $\sigma'(\varphi \cap f(\varphi_1, \dots, \varphi_n)) \neq 0$ . No term with head symbol  $g \neq f$  can be in  $\sigma'(\varphi \cap f(\varphi_1, \dots, \varphi_n))$ , thus there must be a term  $f(t_1, \dots, t_n) \in \sigma'(\varphi \cap f(\varphi_1, \dots, \varphi_n))$ . Then  $f(t_1, \dots, t_n) \in \sigma'(\varphi)$  and  $t_i \in \sigma'(\varphi_i)$ ,  $1 \leq i \leq n$ , and by the correspondence (22),  $\varphi \in \theta(f(t_1, \dots, t_n))'$  and  $\varphi_i \in \theta(t_i)'$ ,  $1 \leq i \leq n$ . But since  $\theta$  is a run,  $\theta(f(t_1, \dots, t_n)) \in E_f^S(\theta(t_1), \dots, \theta(t_n))$ , therefore  $\varphi \cap f(\varphi_1, \dots, \varphi_n)$  is  $\mathcal{S}$ -consistent. This is a contradiction.  $\square$

Theorem 30 implies that if  $F_1 \subseteq F_2$ , where  $F_1$  and  $F_2$  are subexpression-closed families of set expressions over  $X$ , and if  $\mathcal{S}$  is any system of set constraints over  $F_1$ , then  $\mathcal{R}(2^{F_2}, \mathcal{S})$  is a refinement of  $\mathcal{R}(2^{F_1}, \mathcal{S})$  under the natural projection  $2^{F_2} \rightarrow 2^{F_1}$ . In particular, for every subexpression-closed family  $F$  of set expressions over  $X$ ,  $\mathcal{R}(2^F, \emptyset)$  is a refinement of  $\mathcal{R}(2^X, \emptyset)$ .

Similarly, if  $\mathcal{S}_1$  and  $\mathcal{S}_2$  are systems of set constraints over  $F$  and  $\mathcal{S}_1 \models \mathcal{S}_2$ , then any logical consequence of  $\mathcal{S}_2$  is a logical consequence of  $\mathcal{S}_1$ , therefore  $\mathcal{R}(2^F, \mathcal{S}_1)$  is a narrowing of  $\mathcal{R}(2^F, \mathcal{S}_2)$ .

Combining these observations, we see that every  $\mathcal{R}(2^F, \mathcal{S})$  is a rational subspace of  $\mathcal{R}(2^X, \emptyset)$ , since  $\mathcal{R}(2^F, \mathcal{S})$  is a narrowing of  $\mathcal{R}(2^F, \emptyset)$ , which in turn is a refinement of  $\mathcal{R}(2^X, \emptyset)$ .

For  $x \in X$ , let us denote also by  $x$  the clopen set  $\{d \mid x \in d\}$  of any hypergraph  $2^F$ . Let us say that a rational embedding between spaces  $\mathcal{R}(2^F, \mathcal{S})$  *preserves*  $X$  if the map  $h$  on the underlying hypergraphs satisfies  $x = h^{-1}(x)$  for all  $x \in X$ . Note that both the refinement  $\mathcal{R}(2^{F_2}, \mathcal{S}) \rightarrow \mathcal{R}(2^{F_1}, \mathcal{S})$  and the narrowing  $\mathcal{R}(2^F, \mathcal{S}_1) \rightarrow \mathcal{R}(2^F, \mathcal{S}_2)$  preserve  $X$  in this sense. Let us say that a rational equivalence *preserves*  $X$  if the functions  $h_1 : \mathcal{C} \rightarrow 2^{F_1}$  and  $h_2 : \mathcal{C} \rightarrow 2^{F_2}$  from the hypergraph underlying the common refinement  $\mathcal{R}(\mathcal{C})$  satisfy the property that  $h_1^{-1}(x) = h_2^{-1}(x)$  for all  $x \in X$ .

The following is our main theorem.

**Theorem 31** *Let  $X$  be any set of variables. Up to logical equivalence on one side and rational equivalence preserving  $X$  on the other, the systems  $\mathcal{S}$  of set constraints over  $X$  are in one-to-one correspondence with the rational subspaces  $\mathcal{R}(2^F, \mathcal{S})$  of  $\mathcal{R}(2^X, \emptyset)$ . If  $X$  is finite, then the finite systems of set constraints correspond to the finitary subspaces of  $\mathcal{R}(2^X, \emptyset)$ . The correspondence preserves the partial orders of logical entailment on one side and  $X$ -preserving rational embedding on the other.*

**PROOF.** For any system of set constraints  $\mathcal{S}$ , let  $F_{\mathcal{S}}$  be the smallest subexpression-closed subset of  $T_{\Sigma+B}(X)$  containing  $X$  and all expressions occurring in  $\mathcal{S}$ . Consider the map  $\mu : \mathcal{S} \mapsto \mathcal{R}(2^{F_{\mathcal{S}}}, \mathcal{S})$ .

First we show that up to  $X$ -preserving rational equivalence on the right hand side, the map  $\mu$  is well-defined on logical equivalence classes on the left hand side. Suppose  $\mathcal{S}_1 \models \mathcal{S}_2$  and  $\mathcal{S}_2 \models \mathcal{S}_1$ . Let  $F_1 = F_{\mathcal{S}_1}$ ,  $F_2 = F_{\mathcal{S}_2}$ , and  $F = F_1 \cup F_2$ . Then  $\mathcal{R}(2^F, \mathcal{S}_1)$  is a refinement of  $\mathcal{R}(2^{F_1}, \mathcal{S}_1)$  and  $\mathcal{R}(2^F, \mathcal{S}_2)$  is a refinement of  $\mathcal{R}(2^{F_2}, \mathcal{S}_2)$ . But since  $\mathcal{S}_1$  and  $\mathcal{S}_2$  have the same logical consequences, the two spaces  $\mathcal{R}(2^F, \mathcal{S}_1)$  and  $\mathcal{R}(2^F, \mathcal{S}_2)$  coincide, thus form a common refinement of  $\mathcal{R}(2^{F_1}, \mathcal{S}_1)$  and  $\mathcal{R}(2^{F_2}, \mathcal{S}_2)$ . Moreover, the natural refinements  $\mathcal{R}(2^F, \mathcal{S}_1) \rightarrow \mathcal{R}(2^{F_1}, \mathcal{S}_1)$  and  $\mathcal{R}(2^F, \mathcal{S}_2) \rightarrow \mathcal{R}(2^{F_2}, \mathcal{S}_2)$  preserve  $X$ .

To show that the map  $\mu$  is bijective, suppose  $\mathcal{R}(2^{F_1}, \mathcal{S}_1)$  and  $\mathcal{R}(2^{F_2}, \mathcal{S}_2)$  have a common refinement  $\mathcal{R}(\mathcal{D})$  with underlying maps  $h_1 : \mathcal{D} \rightarrow 2^{F_1}$ ,  $h_2 : \mathcal{D} \rightarrow 2^{F_2}$  such that  $h_1^{-1}(x) = h_2^{-1}(x)$ ,  $x \in X$ . It follows that for any run  $\theta$  of  $\mathcal{R}(\mathcal{D})$ ,  $t \in T_{\Sigma}$ , and  $x \in X$ ,

$$\hat{h}_1(\theta)(t) \in x \iff \hat{h}_2(\theta)(t) \in x .$$

Thus under the correspondence (22), the runs of  $\mathcal{R}(2^{F_1}, \mathcal{S}_1)$  and  $\mathcal{R}(2^{F_2}, \mathcal{S}_2)$  correspond to the same family of set valuations. By Theorem 30,  $\mathcal{S}_1$  and  $\mathcal{S}_2$  have the same set of solutions, thus are logically equivalent.

Suppose  $X$  is finite. If  $\mathcal{S}$  is finite, then  $\mathcal{R}(2^{F_{\mathcal{S}}}, \mathcal{S})$  is finitary. Conversely, if  $\mathcal{R}(2^{F_{\mathcal{S}}}, \mathcal{S})$  is finitary, then by Proposition 1,  $\mathcal{S}$  is logically equivalent to a finite system.

Finally, let  $\mathcal{S}_1$  and  $\mathcal{S}_2$  be two systems of set constraints, and let  $F_1 = F_{\mathcal{S}_1}$ ,  $F_2 = F_{\mathcal{S}_2}$ , and  $F = F_1 \cup F_2$ . As argued above, if  $\mathcal{S}_1 \models \mathcal{S}_2$  then  $\mathcal{R}(2^F, \mathcal{S}_1)$  is a narrowing of  $\mathcal{R}(2^F, \mathcal{S}_2)$ , therefore gives a rational embedding preserving  $X$ . Conversely, if  $\hat{h} : \mathcal{R}(2^{F_1}, \mathcal{S}_1) \rightarrow \mathcal{R}(2^{F_2}, \mathcal{S}_2)$  is a rational embedding preserving  $X$  induced by  $h : 2^{F_1} \rightarrow 2^{F_2}$ , then for any run  $\theta$  of  $\mathcal{R}(2^{F_1}, \mathcal{S}_1)$  and  $t \in T_{\Sigma}$ ,

$$\theta(t) \in x \iff h(\theta(t)) \in h(x) \iff \hat{h}(\theta)(t) \in x .$$

Thus under the correspondence (22), the set valuation corresponding to the run  $\theta$  of  $\mathcal{R}(2^{F_1}, \mathcal{S}_1)$  also corresponds to the run  $\hat{h}(\theta)$  of  $\mathcal{R}(2^{F_2}, \mathcal{S}_2)$ . By Theorem 30, every solution of  $\mathcal{S}_1$  is also a solution of  $\mathcal{S}_2$ , thus  $\mathcal{S}_1 \models \mathcal{S}_2$ .  $\square$

We remark that the Stone dual  $\mathbf{St}(F_X/\mathcal{S})$  of the free termset algebra on generators  $X$  modulo  $\mathcal{S}$  [19], embedded in  $2^{T_{\Sigma+B}(X)}$  in the natural way, gives rise to an induced subspace  $\mathcal{R}(\mathbf{St}(F_X/\mathcal{S}))$  of  $\mathcal{R}(2^{T_{\Sigma+B}(X)}, \mathcal{S})$ .

## 5 Applications

In this section we apply the theory of rational spaces to set constraints. We rederive several results in the literature on set constraints and tree set automata in terms of rational spaces. In many cases we are able to generalize the results and give shorter proofs that reveal the topological principles at work.

### 5.1 Positive and Negative Constraints

In [2], Aiken *et al.* prove that mixed positive and negative set constraints are strictly more expressive than positive constraints alone. The proof of this result is based on a compactness lemma which states that a system of positive constraints is satisfiable if and only if all finite subsets are satisfiable. Under the correspondence of Theorem 31, the compactness lemma of [2] is equivalent to the compactness of the space  $\mathcal{R}(2^X, \emptyset)$ .

Now the fact that negative constraints can express something that positive constraints alone cannot is a simple consequence of the observation that all systems of positive set constraints define closed, therefore compact, sets; whereas the set  $A$  of solutions of  $x \neq 0$  is not compact: if  $\sigma_n(x) = \{t \mid \text{depth}(t) > n\}$ , then  $\{\sigma_n \mid n \geq 0\}$  is an infinite subset of  $A$  with no limit point in  $A$ .

### 5.2 Entire Subspaces

In §2.4, we argued that a hypergraph has a run iff it has an entire induced subhypergraph. This essential fact is used in several algorithms for satisfiability of set constraints [1,2,11,12,28]. Similarly, a rational space is nonempty if and only if it contains an entire subspace. This can be generalized as follows.

**Theorem 32** *Every (finitary) rational space is a (finite) union of entire subspaces. Over a language with only constants and unary function symbols, every rational space is essentially entire.*

**PROOF.** The essential subspace of  $\mathcal{R}(\mathcal{D})$  is

$$\mathcal{R}\left(\bigcup_{\theta \in \mathcal{R}(\mathcal{D})} \theta(T_\Sigma)\right) = \bigcup_{\theta \in \mathcal{R}(\mathcal{D})} \mathcal{R}(\theta(T_\Sigma)) , \quad (29)$$

and  $\theta(T_\Sigma)$  is entire. This union is finite if  $\mathcal{D}$  is, since there are only finitely many induced subspaces. Any entire subspace of  $\mathcal{R}(\mathcal{D})$  is a rational subspace of an entire induced subspace of  $\mathcal{R}(\mathcal{D})$ .

Over a language with only constants and unary function symbols, the union  $\bigcup_x \mathcal{D}_x$  of any family  $\mathcal{D}_x$  of entire induced subhypergraphs of  $\mathcal{D}$  is an entire induced subhypergraph of  $\mathcal{D}$ . (Note, however, that in general  $\mathcal{R}(\bigcup_x \mathcal{D}_x) \neq \bigcup_x \mathcal{R}(\mathcal{D}_x)$ ! A counterexample is given below in §5.4.) Applied to (29), this says that the essential subspace of  $\mathcal{R}(\mathcal{D})$  is entire.  $\square$

### 5.3 Density of Rational Points

Gilleron *et al.* [11] prove that every satisfiable finite system of set constraints has a regular solution. Under the correspondence of Theorem 31, this reduces to the following fact about finitary rational spaces:

**Proposition 33** *Every nonempty finitary rational space contains a rational point.*

**PROOF.** If  $\mathcal{R}(\mathcal{D})$  is nonempty, then  $\mathcal{D}$  has an entire induced subspace  $\mathcal{R}(\mathcal{D}')$ . In turn,  $\mathcal{R}(\mathcal{D}')$  has a deterministic narrowing  $\mathcal{R}(\mathcal{D}'')$ , which is a rational point.  $\square$

A rational point corresponds to a deterministic tree set automaton, which has a unique run. Such automata accept regular sets, as can be seen by writing down a system of regular equations of the form (15), one variable for each state.

However, there are more rational points in a finitary rational subspace than just those obtained by narrowings of entire induced subspaces. Indeed, there are only finitely many narrowings of entire induced subspaces, whereas the

number of rational points is infinite in general. For example, the Cantor discontinuum contains countably many rational numbers. An example of a rational point of the Cantor discontinuum that is not a narrowing of an entire induced subspace is given in Example 19. This subspace cannot be represented on a hypergraph of fewer than four vertices.

The following result says that the rational points of any nonempty finitary rational space are dense. In terms of set constraints, the significance of this theorem is that every finite system of positive set constraints is determined by its regular solutions.

**Theorem 34** *Every finitary rational space is the completion of its rational points.*

**PROOF.** Let  $\mathcal{R}(\mathcal{D})$  be a finitary rational space. Recall that  $\mathcal{R}(\mathcal{D})$  is a complete metric space. For any  $x \subseteq \mathcal{D}$  and  $t \in T_\Sigma$ , by Proposition 25 the subbasic open set  $\{\theta \mid \theta(t) \in x\}$  is a finitary rational subspace of  $\mathcal{R}(\mathcal{D})$ . Any basic open set  $A$  of  $\mathcal{R}(\mathcal{D})$  is a finite intersection of subbasic open sets, and  $A$  is a finitary rational subspace of  $\mathcal{R}(\mathcal{D})$  by Proposition 26. By Proposition 33, any nonempty such  $A$  contains a rational point. Thus the set of rational points is dense in  $\mathcal{R}(\mathcal{D})$ , and its closure is  $\mathcal{R}(\mathcal{D})$ .  $\square$

#### 5.4 Closure Properties of Tree Set Automata

In [11], it is claimed without proof that the family of languages accepted by tree set automata is closed under (finite) union, intersection, and cylindrification.

Closure under intersection can be shown as follows. Let  $\mathcal{D}_x$  be an indexed family of tree set automata with tuples  $A_x$  of accept sets. The tuple  $A_x$  is an indexed family  $A_{x_i}$  of subsets of  $\mathcal{D}_x$ . Take the induced subspace of the product  $\prod_x \mathcal{D}_x$  consisting of points  $\bigcap_i (\prod_x A_{x_i} \cup \prod_x \sim A_{x_i})$ . This corresponds to the unification construction of [21,20]. The resulting automaton gives the desired intersection.

Cylindrification is effected as follows. If  $\mathcal{D}$  is a tree set automaton, take the rational product  $\mathcal{D} \times \{0, 1\}$  of  $\mathcal{D}$  with a two-element unrestricted hypergraph. For  $A$  an accept set in  $\mathcal{D}$ , take  $A' = \pi_1^{-1}(A)$ , and add the new accept set  $\pi_2^{-1}(1)$ .

Contrary to the claim of [11], the family of languages accepted by tree set automata is only closed under union for signatures  $\Sigma$  containing at most one constant or at least one symbol of arity two or greater. In those cases, a marked union  $\bigcup_x \mathcal{D}_x \times \{x\}$  of the hypergraphs with the naturally induced hyperedges

will give the desired union of sets of runs: in the case of one constant symbol  $a$ , if  $\theta(a) \in \mathcal{D}_x$  then the entire image of  $\theta$  must be contained in  $\mathcal{D}_x$ ; and in the case of at least one symbol  $f$  of arity two or greater, there can be no run  $\theta$  taking a value  $d \in \mathcal{D}_x$  and  $e \in \mathcal{D}_y$  for  $x \neq y$ , since  $E_f(d, e) = \emptyset$ , so the run cannot be completed.

For  $\Sigma$  containing at least two constants  $a, b$  and no symbol of arity greater than one, the family of languages accepted by tree set automata is not closed under union. Consider the two systems of set constraints  $\{x = 1\}$  and  $\{x = 0\}$ . Let  $\mathcal{M}$  be any tree set automaton accepting at least the union of the two sets of solutions. Because  $\mathcal{M}$  accepts the solutions of the first system, it must admit a run  $\theta$  with  $\theta(T_\Sigma) \subseteq A_x$ . Because  $\mathcal{M}$  accepts the solutions of the second system, it must admit a run  $\eta$  with  $\eta(T_\Sigma) \subseteq A_{\sim x}$ . Let  $H$  denote any composition of unary functions. Then the function

$$\lambda t. \begin{cases} \theta(t) , & \text{if } t = H(a) \text{ for some } H, \\ \eta(t) , & \text{if } t = H(b) \text{ for some } H \end{cases}$$

is a legal run of  $\mathcal{M}$ , but does not satisfy either of the constraints  $x = 1$  or  $x = 0$ .

## 6 Conclusion

In this paper we have investigated the topological structure of the spaces of solutions of systems of set constraints. We have identified a family of topological spaces called *rational spaces*, which formalize the notion of a topological space with a regular or self-similar structure, such as the Cantor discontinuum. We have developed the basic theory of rational spaces and applied this theory to understand more fully the structure of set constraints.

We have not developed the idea of self-similarity, although it is clear that rational spaces exhibit a high degree of self-similarity in the form of ubiquitous self-embeddings. For example, it should be possible to derive an alternative proof of Theorem 34 by taking fixpoints of contractive self-embeddings. Alfons Geser has also pointed out a strong similarity to fractals, which is another connection that may be worth investigating.

In [9], we apply these results further to provide a Gentzen-style axiomatization involving sequents  $\Phi \vdash \Psi$ , where  $\Phi$  and  $\Psi$  are finite sets of set constraints, and give completeness and incompleteness results for this system.

## Acknowledgement

I am indebted to Ed Brinksma, Allan Cheng, Hartmut Ehrig, Alfons Geser, Anil Nerode, Vaughan Pratt, Vladimiro Sassone, and Richard Shore for valuable comments. The support of the National Science Foundation under grant CCR-9317320 and the U.S. Army Research Office through the ACSyAM branch of the Mathematical Sciences Institute of Cornell University under contract DAAL03-91-C-0027 is gratefully acknowledged.

A previous version of this paper appeared in [22].

## References

- [1] A. Aiken, D. Kozen, M. Vardi, and E. Wimmers. The complexity of set constraints. In E. Börger, Y. Gurevich, and K. Meinke, editors, *Proc. 1993 Conf. Computer Science Logic (CSL'93)*, volume 832 of *Lect. Notes in Comput. Sci.*, pages 1–17. Eur. Assoc. Comput. Sci. Logic, Springer, September 1993.
- [2] A. Aiken, D. Kozen, and E. Wimmers. Decidability of systems of set constraints with negative constraints. *Infor. and Comput.*, 1995. To appear. Also Cornell University Tech. Report 93-1362, June, 1993.
- [3] A. Aiken and B. Murphy. Implementing regular tree expressions. In *Proc. 1991 Conf. Functional Programming Languages and Computer Architecture*, pages 427–447, August 1991.
- [4] A. Aiken and B. Murphy. Static type inference in a dynamically typed language. In *Proc. 18th Symp. Principles of Programming Languages*, pages 279–290. ACM, January 1991.
- [5] A. Aiken and E. Wimmers. Solving systems of set constraints. In *Proc. 7th Symp. Logic in Computer Science*, pages 329–340. IEEE, June 1992.
- [6] L. Bachmair, H. Ganzinger, and U. Waldmann. Set constraints are the monadic class. In *Proc. 8th Symp. Logic in Computer Science*, pages 75–83. IEEE, June 1993.
- [7] W. Charatonik and L. Pacholski. Negative set constraints with equality. In *Proc. 9th Symp. Logic in Computer Science*, pages 128–136. IEEE, July 1994.
- [8] W. Charatonik and L. Pacholski. Set constraints with projections are in *nextime*. In *Proc. 35th Symp. Foundations of Computer Science*, pages 642–653. IEEE, November 1994.
- [9] A. Cheng and D. Kozen. A complete Gentzen-style axiomatization for set constraints. Technical Report TR95-1518, Cornell University, May 1995.

- [10] J. Englefriet. Tree automata and tree grammars. Technical Report DAIMI FN-10, Aarhus University, April 1975.
- [11] R. Gilleron, S. Tison, and M. Tommasi. Solving systems of set constraints using tree automata. In *Proc. Symp. Theor. Aspects of Comput. Sci.*, volume 665, pages 505–514. Springer-Verlag Lect. Notes in Comput. Sci., February 1993.
- [12] R. Gilleron, S. Tison, and M. Tommasi. Solving systems of set constraints with negated subset relationships. In *Proc. 34th Symp. Foundations of Comput. Sci.*, pages 372–380. IEEE, November 1993.
- [13] N. Heintze. *Set Based Program Analysis*. PhD thesis, Carnegie Mellon University, 1993.
- [14] N. Heintze and J. Jaffar. A decision procedure for a class of set constraints. In *Proc. 5th Symp. Logic in Computer Science*, pages 42–51. IEEE, June 1990.
- [15] N. Heintze and J. Jaffar. A finite presentation theorem for approximating logic programs. In *Proc. 17th Symp. Principles of Programming Languages*, pages 197–209. ACM, January 1990.
- [16] N. D. Jones and S. S. Muchnick. Flow analysis and optimization of LISP-like structures. In *Proc. 6th Symp. Principles of Programming Languages*, pages 244–256. ACM, January 1979.
- [17] B. Jónsson and A. Tarski. Boolean algebras with operators. *Amer. J. Math.*, 73:891–939, 1951.
- [18] B. Jónsson and A. Tarski. Boolean algebras with operators. *Amer. J. Math.*, 74:127–162, 1952.
- [19] D. Kozen. Logical aspects of set constraints. In E. Börger, Y. Gurevich, and K. Meinke, editors, *Proc. 1993 Conf. Computer Science Logic (CSL'93)*, volume 832 of *Lect. Notes in Comput. Sci.*, pages 175–188. Eur. Assoc. Comput. Sci. Logic, Springer, September 1993.
- [20] D. Kozen. Set constraints and logic programming. Technical Report 94-1467, Cornell University, November 1994.
- [21] D. Kozen. Set constraints and logic programming (abstract). In J.-P. Jouannaud, editor, *Proc. First Conf. Constraints in Computational Logics (CCL'94)*, volume 845 of *Lect. Notes in Comput. Sci.*, pages 302–303. ESPRIT, Springer, September 1994.
- [22] D. Kozen. Rational spaces and set constraints. In P. D. Mosses, M. Nielsen, and M. I. Schwartzbach, editors, *Proc. Sixth Int. Joint Conf. Theory and Practice of Software Develop. (TAPSOFT'95)*, volume 915 of *Lect. Notes in Comput. Sci.*, pages 42–61. Springer, May 1995.
- [23] D. Kozen, J. Palsberg, and M. I. Schwartzbach. Efficient recursive subtyping. In *Proc. 20th Symp. Princip. Programming Lang.*, pages 419–428. ACM, January 1993.



- [24] D. Kozen, J. Palsberg, and M. I. Schwartzbach. Efficient inference of partial types. *J. Comput. Syst. Sci.*, 49(2):306–324, October 1994.
- [25] P. Mishra. Towards a theory of types in PROLOG. In *Proc. 1st Symp. Logic Programming*, pages 289–298. IEEE, 1984.
- [26] P. Mishra and U. Reddy. Declaration-free type checking. In *Proc. 12th Symp. Principles of Programming Languages*, pages 7–21. ACM, 1985.
- [27] J. C. Reynolds. Automatic computation of data set definitions. In *Information Processing 68*, pages 456–461. North-Holland, 1969.
- [28] K. Stefánsson. Systems of set constraints with negative constraints are *nexptime*-complete. In *Proc. 9th Symp. Logic in Computer Science*, pages 137–141. IEEE, June 1994.
- [29] J. Young and P. O’Keefe. Experience with a type evaluator. In D. Bjørner, A. P. Ershov, and N. D. Jones, editors, *Partial Evaluation and Mixed Computation*, pages 573–581. North-Holland, 1988.