Nominal Topology for Data Languages

Fabian Birkmann
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Stefan Milius
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Henning Urbat
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Abstract

We propose a novel topological perspective on data languages recognizable by orbit-finite nominal monoids. For this purpose, we introduce pro-orbit-finite nominal topological spaces. Assuming globally bounded support sizes, they coincide with nominal Stone spaces and are shown to be dually equivalent to a subcategory of nominal boolean algebras. Recognizable data languages are characterized as topologically clopen sets of pro-orbit-finite words. In addition, we explore the expressive power of pro-orbit-finite equations by establishing a nominal version of Reiterman's pseudovariety theorem.

2012 ACM Subject Classification Theory of computation → Formal languages and automata theory

Keywords and phrases Nominal sets, Stone duality, Profinite space, Data languages

Digital Object Identifier 10.4230/LIPIcs.ICALP.2023.114

Category Track B: Automata, Logic, Semantics, and Theory of Programming

Related Version *Full Version*: `https://arxiv.org/abs/2304.13337`

Funding Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 470467389.

1 Introduction

While automata theory is largely concerned with formal languages over finite alphabets, the extension to *infinite alphabets* has been identified as a natural approach to modelling structures involving *data*, such as nonces [26], channel names [23], object identities [22], process identifiers [11], URLs [5], or values in XML documents [31]. For example, if $\mathbb{A}$ is a (countably infinite) set of data values, typical languages to consider might be

$$L_0 = \{\, vaaw \mid a \in \mathbb{A},\ v, w \in \mathbb{A}^* \,\} \qquad \text{("some data value occurs twice in a row"), or}$$
$$L_1 = \{\, avaw \mid a \in \mathbb{A},\ v, w \in \mathbb{A}^* \,\} \qquad \text{("the first data value occurs again").}$$

Automata for data languages enrich finite automata with register mechanisms that allow to store data and test data values for equality (or more complex relations, e.g. order) [24, 31]. In a modern perspective first advocated by Bojańczyk, Klin, and Lasota [9], a convenient abstract framework for studying data languages is provided by the theory of *nominal sets* [36].

Despite extensive research in the past three decades, no universally acknowledged notion of *regular* data language has emerged so far. One reason is that automata models with data notoriously lack robustness, in that any alteration of their modus operandi (e.g. deterministic vs. nondeterministic, one-way vs. two-way) usually affects their expressive power. Moreover, machine-independent descriptions of classes of data languages in terms of algebra or model theory are hard to come by. However, there is one remarkable class of data languages that closely mirrors classical regular languages: data languages recognizable by orbit-finite nominal monoids [7]. Originally introduced from a purely algebraic angle, recognizable data languages have subsequently been characterized in terms of *rigidly guarded* $MSO^{\sim}$, a fragment of monadic second-order logic with equality tests [13], *single-use register automata* [10] (both one-way and two-way), and *orbit-finite regular list functions* [10]. In addition, several landmark results from the algebraic theory of regular languages, namely the McNaughton-Papert-Schützenberger theorem [29, 40], the Krohn-Rhodes theorem [25], and Eilenberg's variety theorem [14] have been extended to recognizable data languages [7, 10, 13, 44].

In the present paper, we investigate recognizable data languages through the lens of *topology*, thereby providing a further bridge to classical regular languages. The topological approach to the latter is closely tied to the algebraic one, which regards regular languages as the languages recognizable by finite monoids. Its starting point is the construction of the topological space $\widehat{\Sigma^*}$ of *profinite words*. Informally, this space casts all information represented by regular languages over Σ and their recognizing monoids into a single mathematical object. Regular languages can then be characterized by purely topological means: they may be interpreted as precisely the clopen subsets of $\widehat{\Sigma^*}$, in such way that algebraic recognition by finite monoids becomes a continuous process. Properties of regular languages are often most conveniently classified in terms of the topological concept of *profinite equations*, that is, equations between profinite words; see [3, 4, 33] for a survey of profinite methods in automata theory. Moreover, since $\widehat{\Sigma^*}$ forms a Stone space, the power of *Stone duality* – the dual equivalence between Stone spaces and boolean algebras – becomes available. This allows for the use of duality-theoretic methods for the study of regular languages and their connection to logic and model theory, which in part even extend to *non-regular* languages [18–21, 35].

On a conceptual level, the topological view of regular languages rests on a single category-theoretic fact: Stone spaces admit a universal property. In fact, they arise from the category of finite sets as the free completion under codirected limits, a.k.a. its *Pro-completion*:

$$\mathbf{Stone} \simeq \mathrm{Pro}(\mathbf{Set}_{\mathrm{f}}). \tag{1.1}$$

In the world of data languages, the role of finite sets is taken over by orbit-finite nominal sets. This strongly suggests to base a topological approach on their free completion $\mathrm{Pro}(\mathbf{Nom}_{\mathrm{of}})$. However, this turns out to be infeasible: the category $\mathrm{Pro}(\mathbf{Nom}_{\mathrm{of}})$ is not concrete over nominal sets (Proposition 3.6), hence it cannot be described via any kind of nominal topological spaces. This is ultimately unsurprising given that the description (1.1) of Stone spaces as a free completion depends on the axiom of choice, which is well-known to fail in the topos of nominal sets. As a remedy, we impose *global bounds* on the support sizes of nominal sets, that is, we consider the categories $\mathbf{Nom}_k$ and $\mathbf{Nom}_{\mathrm{of},k}$ of (orbit-finite) nominal sets where every element has a support of size k, for some fixed natural number k. This restriction is natural from an automata-theoretic perspective, as it corresponds to imposing a bound k on the number of registers of automata, and it fixes exactly the issue making unrestricted nominal sets non-amenable (Lemma 3.9). Let us emphasize, however, that the category $\mathbf{Nom}_k$ is not proposed as a new foundation for names and variable binding; for instance, it generally fails to be a topos.

The first main contribution of our paper is a generalization of (1.1) to k-bounded nominal sets. For this purpose we introduce *nominal Stone spaces*, a suitable nominalization of the classical concept, and prove that k-bounded nominal Stone spaces form the Pro-completion of the category of k-bounded orbit-finite sets. We also derive a nominal version of Stone duality, which relates k-bounded nominal Stone spaces to *locally k-atomic orbit-finitely complete nominal boolean algebras*. Hence we establish the following equivalences of categories:

$$\mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA} \simeq^{\mathrm{op}} \mathbf{nStone}_k \simeq \mathrm{Pro}(\mathbf{Nom}_{\mathrm{of},k}).$$

The above equivalences are somewhat remarkable since even the category of k-bounded nominal sets does not feature choice. They hold because the presence of bounds allows us to reduce topological properties of nominal Stone spaces, most notably compactness, to their classical counterparts.

Building on the above topological foundations, which we regard to be of independent interest, we subsequently develop first steps of a topological theory of data languages. Specifically, we introduce nominal Stone spaces of (bounded) *pro-orbit-finite words* and prove their clopen subsets to correspond to data languages recognizable by bounded equivariant monoid morphisms, generalizing the topological characterization of classical regular languages (Theorem 5.14). Moreover, we investigate the expressivity of *pro-orbit-finite equations* and show that they model precisely classes of orbit-finite monoids closed under finite products, submonoids, and *multiplicatively support-reflecting quotients* (Theorem 6.8). This provides a nominal version of Reiterman's celebrated pseudovariety theorem [37] for finite monoids.

Related work. The perspective taken in our paper draws much of its inspiration from the recent categorical approach to algebraic recognition based on monads [8, 38, 42]. The importance of Pro-completions in algebraic language theory has been isolated in the work of Chen et al. [12] and Urbat et al. [42]. In the latter work the authors introduce *profinite monads* and present a general version of Eilenberg's variety theorem parametric in a given Stone-type duality. The theory developed there applies to algebraic base categories, but not to the category of nominal sets.

Our version of nominal Stone duality builds on the orbit-finite restriction of the duality between nominal sets and complete atomic nominal boolean algebras due to Petrişan [17]. It is fundamentally different from the nominal Stone duality proposed by Gabbay, Litak, and Petrişan [16], which relates *nominal Stone spaces with* $\mathsf{И}$ to *nominal boolean algebras with* $\mathsf{И}$. The latter duality is not amenable for the theory of data languages; see Remark 3.17.

Reiterman's pseudovariety theorem has recently been generalized to the level of finite algebras for a monad [1, 12] and, in a more abstract disguise, finite objects in a category [30]. For nominal sets, varieties of algebras over binding signatures have been studied by Gabbay [16] and by Kurz and Petrişan [27], resulting in nominal Birkhoff-type theorems [6]. Urbat and Milius [44] characterize classes of orbit-finite monoids called *weak* pseudovarieties by sequences of nominal word equations. This gives a nominal generalization of the classical Eilenberg-Schützenberger theorem [15], which in fact is a special case of the general HSP theorem in [30]. Nominal pro-orbit-finite equations as introduced in the present paper are strictly more expressive than sequences of nominal word equations (Example 6.11), hence our nominal Reiterman theorem is not equivalent to the nominal Eilenberg-Schützenberger theorem. Moreover, we note that the nominal Reiterman theorem does not appear to be an instance of any of the abstract categorical frameworks mentioned above.

2 Preliminaries

We assume that readers are familiar with basic notions from category theory, e.g. functors, natural transformations, and (co)limits, and from point-set topology, e.g. metric and topological spaces, continuous maps, and compactness. In the following we recall some facts about Pro-completions, the key categorical concept underlying our topological approach to data languages. Moreover, we give a brief introduction to the theory of nominal sets [36].

Pro-completions. A small category I is *cofiltered* if (i) I is non-empty, (ii) for every pair of objects $i, j \in I$ there exists a span $i \leftarrow k \to j$, and (iii) for every pair of parallel arrows $f, g \colon j \to k$, there exists a morphism $h \colon i \to j$ such that $f \cdot h = g \cdot h$. Cofiltered preorders are called *codirected*; thus a preorder I is codirected if $I \neq \emptyset$ and every pair $i, j \in I$ has a lower bound $k \leq i, j$. For instance, every meet-semilattice with bottom is codirected. A diagram $D\colon I \to \mathbf{C}$ in a category $\mathbf{C}$ is *cofiltered* if its index category I is cofiltered. A *cofiltered limit* is a limit of a cofiltered diagram. *Codirected limits* are defined analogously. The two concepts are closely related: a category has cofiltered limits iff it has codirected limits, and a functor preserves cofiltered limits iff it preserves codirected limits [2, Cor. 1.5]. The dual concept is that of a *filtered colimit* or a *directed colimit*, respectively.

▶ **Example 2.1.**

1. In the category $\mathbf{Set}$ of sets and functions, every filtered diagram $D\colon I \to \mathbf{Set}$ has a colimit cocone $c_i\colon D_i \to \operatorname{colim} D$ $(i \in I)$ given by $\operatorname{colim} D = \big(\coprod_{i\in I} D_i\big)/{\sim}$ and $c_i(x) = [x]_\sim$, where the equivalence relation $\sim$ on the coproduct (i.e. disjoint union) $\coprod_{i\in I} D_i$ relates $x \in D_i$ and $y \in D_j$ iff there exist morphisms $f\colon i \to k$ and $g\colon j \to k$ in I such that $Df(x) = Dg(y)$.
2. Every cofiltered diagram $D\colon I \to \mathbf{Set}$ has a limit whose cone $p_i\colon \lim D \to D_i$ $(i \in I)$ is given by the *compatible families* of D and projection maps:

$$\lim D = \{(x_i)_{i\in I} \mid x_i \in D_i \text{ and } Df(x_i) = x_j \text{ for all } f\colon i \to j \text{ in } I\} \quad \text{and} \quad p_j((x_i)_{i\in I}) = x_j.$$

3. In the category $\mathbf{Top}$ of topological spaces and continuous maps, the limit cone of a cofiltered diagram $D\colon I \to \mathbf{Top}$ is formed by taking the limit in $\mathbf{Set}$ and equipping $\lim D$ with the *initial topology*, viz. the topology generated by the basic open sets $p_i^{-1}[U_i]$ for $i \in I$ and $U_i \subseteq D_i$ open.

An object C of a category $\mathbf{C}$ is *finitely copresentable* if the contravariant hom-functor $\mathbf{C}(-, C)\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$ preserves directed colimits. In more elementary terms, this means that for every codirected diagram $D\colon I \to \mathbf{C}$ with limit cone $p_i\colon L \to D_i$ $(i \in I)$,

1. every morphism $f\colon L \to C$ factorizes as $f = g \circ p_i$ for some $i \in I$ and $g\colon D_i \to C$, and
2. the factorization is essentially unique: given another factorization $f = h \cdot p_i$, there exists $j \leq i$ such that $g \cdot D_{j,i} = h \cdot D_{j,i}$.

A *Pro-completion* of a small category $\mathbf{C}$ is a free completion under codirected (equivalently cofiltered) limits. It is given by a category $\mathrm{Pro}(\mathbf{C})$ with codirected limits together with a full embedding $E\colon \mathbf{C} \hookrightarrow \mathrm{Pro}(\mathbf{C})$ satisfying the following universal property:

1. every functor $F\colon \mathbf{C} \to \mathbf{D}$, where the category $\mathbf{D}$ has codirected limits, extends to a functor $\bar{F}\colon \mathrm{Pro}(\mathbf{C}) \to \mathbf{D}$ that preserves codirected limits and satisfies $F = \bar{F} \circ E$;
2. $\bar{F}$ is essentially unique: For every functor G that preserves codirected limits and satisfies $F = G \circ E$, there exists a natural isomorphism $\alpha\colon \bar{F} \cong G$ such that $\alpha E = \mathrm{id}_F$.

$$\begin{tikzcd} \mathbf{C} \arrow[r, hook, "E"] \arrow[dr, "F"'] & \mathrm{Pro}(\mathbf{C}) \arrow[d, dashed, "\bar{F}"] \\ & \mathbf{D} \end{tikzcd} \qquad \begin{tikzcd} \mathbf{C} \arrow[r, hook, "E"] \arrow[dr, "F"'] & \mathrm{Pro}(\mathbf{C}) \arrow[d, dashed, "G"] \\ & \mathbf{D} \end{tikzcd}$$

The universal property determines $\mathrm{Pro}(\mathbf{C})$ uniquely up to equivalence of categories. We note that every object EC $(C \in \mathbf{C})$ is finitely copresentable in $\mathrm{Pro}(\mathbf{C})$, see e.g. [1, Thm A.4]. The dual of Pro-completions are *Ind-completions*: free completions under *directed colimits*.

▶ **Example 2.2.** The Pro-completion $\mathrm{Pro}(\mathbf{Set}_\mathrm{f})$ of the category of finite sets is the full subcategory of $\mathbf{Top}$ given by *profinite spaces* (topological spaces that are codirected limits of finite discrete spaces). Profinite spaces are also known as *Stone spaces* or *boolean spaces* and can be characterized by topological properties: they are precisely compact Hausdorff spaces with a basis of clopen sets. This equivalent characterization depends on the axiom of choice (or rather the ultrafilter theorem, a weak form of choice), as does *Stone duality*, the dual equivalence between the categories of Stone spaces and boolean algebras. The duality maps a Stone space to its boolean algebra of clopen sets, equipped with the set-theoretic boolean operations. Its inverse maps a boolean algebra the set of ultrafilters (equivalently, prime filters) on it, equipped with a suitable profinite topology.

Profinite words. The topological approach to classical regular languages is based on the space $\widehat{\Sigma^*}$ of *profinite words* over the alphabet Σ. This space is constructed as the codirected limit of all finite quotient monoids of Σ^*, the free monoid of finite words generated by Σ. Formally, let $\Sigma^*{\downarrow}\mathbf{Mon}_\mathrm{f}$ be the codirected poset of all surjective monoid morphisms $e\colon \Sigma^* \twoheadrightarrow M$, where M is a finite monoid; the order on $\Sigma^*{\downarrow}\mathbf{Mon}_\mathrm{f}$ is defined by $e \leq e'$ if $e' = e \cdot h$ for some h. Then $\widehat{\Sigma^*}$ is the limit of the diagram $D\colon \Sigma^*{\downarrow}\mathbf{Mon}_\mathrm{f} \to \mathrm{Pro}(\mathbf{Set}_\mathrm{f})$ sending $e\colon \Sigma^* \twoheadrightarrow M$ to the underlying set of M, regarded as a finite discrete topological space. The space $\widehat{\Sigma^*}$ is completely metrizable; in fact, it is the Cauchy completion of the metric space (Σ^*, d) where $d(v, w) = \sup\{\, 2^{-|M|} \mid M \text{ is a finite monoid separating } v, w \,\}$. Here a monoid M *separates* $v, w \in \Sigma^*$ if there exists a morphism $h\colon \Sigma^* \to M$ such that $h(v) \neq h(w)$. Regular languages over Σ correspond to clopen subsets of $\widehat{\Sigma^*}$, or equivalently to continuous maps $L\colon \widehat{\Sigma^*} \to 2$ into the discrete two-element space.

Nominal Sets. Fix a countable set $\mathbb{A}$ of *names*, and denote by $\mathrm{Perm}\,\mathbb{A}$ the group of finite permutations, i.e. bijections $\pi\colon \mathbb{A} \to \mathbb{A}$ fixing all but finitely many names. Given $S \subseteq \mathbb{A}$ write

$$\mathrm{Perm}_S\,\mathbb{A} = \{\pi \in \mathrm{Perm}\,\mathbb{A} \mid \pi(a) = a \text{ for all } a \in S\}$$

for the the subgroup of permutations fixing S. A $\mathrm{Perm}\,\mathbb{A}$*-set* is a set X with a group action, that is, an operation $\cdot\colon \mathrm{Perm}\,\mathbb{A} \times X \to X$ such that $\mathrm{id} \cdot x = x$ and $\pi \cdot (\sigma \cdot x) = (\pi \circ \sigma) \cdot x$ for every $x \in X$ and $\pi, \sigma \in \mathrm{Perm}\,\mathbb{A}$. The *trivial* group action on X is given by $\pi \cdot x = x$ for all $x \in X$ and $\pi \in \mathrm{Perm}\,\mathbb{A}$.

A subset $S \subseteq \mathbb{A}$ is a *support* of $x \in X$ if every permutation $\pi \in \mathrm{Perm}_S\,\mathbb{A}$ acts trivially on x, that is, $\pi \cdot x = x$. The idea is that x is some syntactic object (e.g. a word, a tree, or a λ-term) whose free variables are contained in S. A $\mathrm{Perm}\,\mathbb{A}$-set X is a *nominal set* if every element $x \in X$ has a finite support. This implies that every $x \in X$ has a *least* finite support, denoted by $\mathrm{supp}\, x \subseteq \mathbb{A}$.

For a nominal set X its *nominal powerset* $\mathcal{P}_\mathrm{fs}X \subseteq \mathcal{P}X$ consists of all subsets of $U \subseteq X$ which are finitely supported under the action $\pi \cdot U := \{\pi \cdot x \mid x \in U\}$. For example, for the nominal set $\mathbb{A}$ of names with the action $\pi \cdot a = \pi(a)$, its nominal powerset $\mathcal{P}_\mathrm{fs}\mathbb{A}$ consists of all finite and cofinite subsets of $\mathbb{A}$. A subset $U \subseteq X$ is *equivariant* if it has empty support. If there exists a finite subset $S \subseteq \mathbb{A}$ supporting every $x \in U$ then U is *uniformly finitely supported*, and S also supports U. Given a finite set $S \subseteq \mathbb{A}$ of names and a subset $U \subseteq X$, we define the S*-hull* of U by $\mathrm{hull}_S\, U = \{\pi \cdot x \mid x \in U, \pi \in \mathrm{Perm}_S\,\mathbb{A}\}$. This is the smallest S-supported subset of X containing U.

For finite $S \subseteq \mathbb{A}$ the S*-orbit* of an element $x \in X$ is the set $\mathrm{orb}_S\, x = \{\pi \cdot x \mid \pi \in \mathrm{Perm}_S\,\mathbb{A}\}$. The $\emptyset$-orbit of x is called its *orbit*, denoted $\mathrm{orb}\, x$. We write $\mathrm{orb}_S\, X = \{\mathrm{orb}_S\, x \mid x \in X\}$ for the set of all S-orbits of X, and $\mathrm{orb}\, X$ for the set of all orbits. The S-orbits form a partition of X. A finitely supported subset $Y \subseteq X$ is *orbit-finite* if it intersects only finitely many orbits of X. In particular, the nominal set X is *orbit-finite* if $\mathrm{orb}\, X$ is a finite set. This implies that for every finite subset $S \subseteq \mathbb{A}$ the set $\mathrm{orb}_S\, X$ is finite. Moreover, X contains only finitely many elements with support S.

▶ **Example 2.3.** The set $\mathbb{A}^*$ of finite words over $\mathbb{A}$ forms a nominal set with the group action $\pi \cdot (a_1 \cdots a_n) = \pi(a_1) \cdots \pi(a_n)$. The languages $L_0, L_1 \subseteq \mathbb{A}^*$ from the Introduction are equivariant subsets. Given a fixed name $a \in \mathbb{A}$, the subset $L_2 = \{awa \mid w \in \mathbb{A}^*\}$ is finitely supported with $\operatorname{supp} L_2 = \{a\}$. All the above sets have an infinite number of orbits. An example of an orbit-finite set is given by $\mathbb{A}^2 = \mathbb{A} \times \mathbb{A} \subseteq \mathbb{A}^*$; its two orbits are $\{aa \mid a \in \mathbb{A}\}$ and $\{ab \mid a \neq b \in \mathbb{A}\}$.

A map $f \colon X \to Y$ between nominal sets is *finitely supported* if there exists a finite set $S \subseteq \mathbb{A}$ such that $f(\pi \cdot x) = \pi \cdot f(x)$ for all $x \in X$ and $\pi \in \operatorname{Perm}_S \mathbb{A}$, and *equivariant* if it is supported by $S = \emptyset$. Equivariant maps satisfy $\operatorname{supp} f(x) \subseteq \operatorname{supp} x$ for all $x \in X$. Nominal sets and equivariant maps form a category $\mathbf{Nom}$, with the full subcategory $\mathbf{Nom}_{\mathrm{of}}$ of orbit-finite nominal sets. The category $\mathbf{Nom}$ is complete and cocomplete. Colimits and finite limits are formed like in $\mathbf{Set}$; general limits are formed by taking the limit in $\mathbf{Set}$ and restricting to finitely supported elements. The category $\mathbf{Nom}_{\mathrm{of}}$ is closed under finite limits and finite colimits in $\mathbf{Nom}$. *Quotients* and *subobjects* in $\mathbf{Nom}$ are represented by surjective and injective equivariant maps. Every equivariant map f has an image factorization $f = m \cdot e$ with m injective and e surjective; we call e the *coimage* of f.

A nominal set is *strong* if for all $x \in X$ and $\pi \in \operatorname{Perm} \mathbb{A}$ one has $\pi \cdot x = x$ iff $\pi \in \operatorname{Perm}_S \mathbb{A}$, where $S = \operatorname{supp} x$. (Note that the "if" direction holds in every nominal set.) For example, the nominal set $\mathbb{A}^{\#n} = \{f \colon n \to \mathbb{A} \mid f \text{ injective}\}$ with pointwise action is strong and has a single orbit. Up to isomorphism, (orbit-finite) strong nominal sets are precisely (finite) coproducts of such sets.

3 Nominal Stone Spaces

In this section, we establish the topological foundations for our pro-orbit-finite approach to data languages. We start by recalling the basic definitions of nominal topology [17, 32].

▶ **Definition 3.1.**

1. *A* nominal topology *on a nominal set X is an equivariant subset $\mathcal{O}_X \subseteq \mathcal{P}_{fs}X$ closed under finitely supported union (that is, if $\mathcal{U} \subseteq \mathcal{O}_X$ is finitely supported then $\bigcup \mathcal{U} \in \mathcal{O}_X$) and finite intersection. Sets $U \in \mathcal{O}_X$ are called* open *and their complements* closed*; sets that are both open and closed are* clopen*. A nominal set X together with a nominal topology $\mathcal{O}_X$ is a* nominal topological space*. An equivariant map $f \colon X \to Y$ between nominal topological spaces is* continuous *if for every open set U of Y its preimage $f^{-1}[U]$ is an open set of X. Nominal topological spaces and continuous maps form the category* $\mathbf{nTop}$.
2. *A* subbasis *of a nominal topological space $(X, \mathcal{O}_X)$ is an equivariant subset $\mathcal{B} \subseteq \mathcal{O}_X$ such that every open set of X is a finitely supported union of finite intersections of sets in $\mathcal{B}$. If additionally every finite intersection of sets in $\mathcal{B}$ is a finitely supported union of sets in $\mathcal{B}$, then $\mathcal{B}$ is called a* basis*. In this case, every open set of X is a finitely supported union of elements of $\mathcal{B}$.*

▶ **Example 3.2.**

1. A topological space may be viewed as a nominal topological space equipped with the trivial group action. Then every (open) subset has empty support and every union is finitely supported, so we recover the axioms of classical topology.
2. Every nominal set X equipped with the *discrete topology*, where all finitely supported subsets are open, is a nominal topological space. It has a basis given by all singleton sets.

3. A *nominal (pseudo-)metric space* is given by a nominal set X with a (pseudo-)metric[1] $d\colon X \times X \to \mathbb{R}$ which is equivariant as a function into the set $\mathbb{R}$, regarded as a nominal set with the trivial group action. As usual, the open ball around $x \in X$ with radius $r > 0$ is given by $B_r x = \{y \in X \mid d(x, y) < r\}$. Since $\pi \cdot B_r(x) = B_r(\pi \cdot x)$ for all $\pi \in \operatorname{Perm} \mathbb{A}$ and $x \in X$, every nominal (pseudo-)metric space carries a nominal topology whose basic opens are the open balls.

▶ Remark 3.3. Every nominal topological space induces two families of ordinary topological spaces, one by taking only opens with a certain support and the other by forming orbits. In more detail, let $S \subseteq \mathbb{A}$ be a finite set of names and let X be a nominal topological space with topology $\mathcal{O}$.

1. The underlying set of the nominal space X carries a classical topology $\mathcal{O}_S$ consisting of all S-supported open sets of $\mathcal{O}$. We denote the resulting topological space by $|X|_S$.
2. The set $\operatorname{orb}_S X$ of S-orbits can be equipped with the quotient topology $\mathcal{O}_{\operatorname{orb}_S}$ induced by the projection $X \twoheadrightarrow \operatorname{orb}_S X$ mapping each $x \in X$ to its S-orbit $\operatorname{orb}_S x$. In this topology, a set $O \subseteq \operatorname{orb}_S X$ of S-orbits is open iff its union $\bigcup O$ is open in X.

These constructions give rise to functors $|{-}|_S, \operatorname{orb}_S\colon \mathbf{nTop} \to \mathbf{Top}$. They allow us to switch between nominal and classical topology.

As noted in Example 2.2, the Pro-completion of the category $\mathbf{Set}_\mathrm{f}$ is the category of profinite spaces. One may expect that the Pro-completion of $\mathbf{Nom}_\mathrm{of}$ analogously consists of all *pro-orbit-finite* spaces, that is, nominal topological spaces that are codirected limits of orbit-finite discrete spaces. However, this fails due to a simple fact: while codirected limits of non-empty finite sets are always non-empty (which is a consequence of Tychonoff's theorem, thus the axiom of choice), codirected limits of non-empty orbit-finite nominal sets may be empty.

▶ Remark 3.4. Similar to $\mathbf{Top}$, codirected limits in $\mathbf{nTop}$ are formed by taking the limit in $\mathbf{Nom}$ equipping it with the initial topology.

▶ **Example 3.5.** Consider the ω^op-chain $1 \leftarrow \mathbb{A} \leftarrow \mathbb{A}^{\#2} \leftarrow \mathbb{A}^{\#3} \leftarrow \cdots$ in $\mathbf{Nom}_\mathrm{of}$ with connecting maps omitting the last component. Its limit in $\mathbf{Set}$ (see Example 2.1) is given by $\mathbb{A}^{\#\omega}$, the set of all injective functions from ω to $\mathbb{A}$. Clearly no such function has finite support, thus the limit in $\mathbf{Nom}$ (and therefore also in $\mathbf{nTop}$) is empty.

This entails that it is in fact impossible to characterize $\operatorname{Pro}(\mathbf{Nom}_\mathrm{of})$ by any sort of spaces. By definition of the free completion $\operatorname{Pro}(\mathbf{Nom}_\mathrm{of})$, the inclusion functor $I\colon \mathbf{Nom}_\mathrm{of} \hookrightarrow \mathbf{Nom}$ extends uniquely to a functor $\bar{I}\colon \operatorname{Pro}(\mathbf{Nom}_\mathrm{of}) \to \mathbf{Nom}$ preserving codirected limits. The analogous functor $\bar{I}\colon \operatorname{Pro}(\mathbf{Set}_\mathrm{f}) \to \mathbf{Set}$ is the forgetful functor of the category of profinite spaces. In contrast, we have

▶ **Proposition 3.6.** *The category* $\operatorname{Pro}(\mathbf{Nom}_\mathrm{of})$ *is* not *concrete: the functor* $\bar{I}$ *is not faithful.*

Proof. Consider the chain $1 \leftarrow \mathbb{A} \leftarrow \mathbb{A}^{\#2} \leftarrow \cdots$ of Example 3.5. Let $D\colon \omega^\mathrm{op} \to \mathbf{Nom}_\mathrm{of}$ denote the corresponding diagram, and let $E\colon \mathbf{Nom}_\mathrm{of} \hookrightarrow \operatorname{Pro}(\mathbf{Nom}_\mathrm{of})$ be the embedding. To prove that $\bar{I}$ is not faithful, let 2 be the two element nominal set. We show that $|\operatorname{Pro}(\mathbf{Nom}_\mathrm{of})(\lim ED, E2)| > |\mathbf{Nom}(\bar{I}(\lim ED), \bar{I}E2)|$. Indeed, we have

$$\begin{aligned}
\operatorname{Pro}(\mathbf{Nom}_\mathrm{of})(\lim_{n<\omega} ED_n, E2) &\cong \operatorname*{colim}_{n<\omega} \operatorname{Pro}(\mathbf{Nom}_\mathrm{of})(ED_n, E2) && E2 \text{ finitely copresentable}\\
&\cong \operatorname*{colim}_{n<\omega} \mathbf{Nom}_\mathrm{of}(D_n, 2) && E \text{ full embedding}\\
&\cong 2
\end{aligned}$$

[1] Recall that a pseudometric differs from a metric by not requiring $d(x, y) \neq 0$ for $x \neq y$.

because $\mathbf{Nom}_{\mathrm{of}}(D_0, 2) \cong 2$ and the two elements are not merged by the colimit injection. However,

$$\begin{aligned}
\mathbf{Nom}(\bar{I}(\lim_{n<\omega} ED_n), \bar{I}E2) &\cong \mathbf{Nom}(\lim_{n<\omega} \bar{I}ED_n, \bar{I}E2) && \bar{I} \text{ preserves codirected limits} \\
&\cong \mathbf{Nom}(\lim_{n<\omega} ID_n, 2) && I = \bar{I}E \\
&\cong \mathbf{Nom}(\emptyset, 2) && \text{Example 3.5} \\
&\cong 1.
\end{aligned}$$

◀

We thus restrict our focus to well-behaved subcategories of $\mathbf{Nom}_{\mathrm{of}}$. We choose these subcategories in such way that situations like in Example 3.5, where unrestricted accumulation of supports results in empty codirected limits, are avoided.

▶ **Definition 3.7.** *A nominal set X is* k-bounded*, for $k \in \mathbb{N}$, if $|\mathrm{supp}\, x| \leq k$ for every $x \in X$.*

For concrete categories $\mathbf{C}$ over $\mathbf{Nom}$ (or $\mathbf{Nom}_{\mathrm{of}}$) we denote by $\mathbf{C}_k$ the full subcategory of $\mathbf{C}$ whose underlying objects are k-bounded. For instance, $\mathbf{Nom}_k$ is the category of k-bounded nominal sets, and $\mathbf{nTop}_k$ is the category of k-bounded nominal topological spaces.

▶ Remark 3.8.

1. The full subcategories $\mathbf{Nom}_k \hookrightarrow \mathbf{Nom}$ and $\mathbf{Nom}_{\mathrm{of},k} \hookrightarrow \mathbf{Nom}_{\mathrm{of}}$ are coreflective [28, Section IV.3]: the coreflector (viz. the right adjoint of the inclusion functor) sends a nominal set X to its subset $X_k = \{x \in X \mid |\mathrm{supp}\, x| \leq k\}$. Hence $\mathbf{Nom}_k$ is complete: limits are formed by taking the limit in $\mathbf{Nom}$ and applying the coreflector. Analogously, $\mathbf{Nom}_{\mathrm{of},k}$ is finitely complete.
2. In contrast to $\mathbf{Nom}$, the category $\mathbf{Nom}_k$ generally fails to be a topos because it is not cartesian closed. For instance, the functor $\mathbb{A}^{\#2} \times (-)$ on $\mathbf{Nom}_2$ does not preserve coequalizers, hence it is not a left adjoint.
3. The category $\mathbf{Nom}$ is known to be equivalent to the category of pullback-preserving presheaves $\mathbb{I} \to \mathbf{Set}$, where $\mathbb{I}$ is the category of finite sets and injective functions [36, Theorem 6.8]. By inspecting the proof it is easy to see that this restricts to an equivalence between $\mathbf{Nom}_k$ and the category of k-*generated* pullback-preserving presheaves $\mathbb{I} \to \mathbf{Set}$. Here a presheaf $F\colon \mathbb{I} \to \mathbf{Set}$ is k-generated if for every finite set S and every $x \in FS$ there exists a set S' of cardinality at most k and an injective map $f\colon S' \to S$ such that $x \in Ff[FS']$.

With regard to codirected limits, the restriction to bounded nominal sets fixes the issue arising in Example 3.5:

▶ **Lemma 3.9.** *Codirected limits in $\mathbf{Nom}_k$ are formed at the level of $\mathbf{Set}$.*

We proceed to give a topological characterization of $\mathrm{Pro}(\mathbf{Nom}_{\mathrm{of},k})$ in terms of nominal Stone spaces, generalizing the corresponding result (1.1) for $\mathrm{Pro}(\mathbf{Set}_{\mathrm{f}})$. To this end, we introduce suitable nominalizations of the three characteristic properties of Stone spaces: compactness, Hausdorffness, and existence of a basis of clopens. The nominal version of compactness comes natural and is compatible with the functors $|-|_S$ and orb_S of Remark 3.3.

▶ **Definition 3.10.** *An* open cover *of a nominal topological space $(X, \mathcal{O})$ is a finitely supported set $\mathcal{C} \subseteq \mathcal{O}$ that covers X, i.e. $\bigcup \mathcal{C} = X$. A* subcover *of $\mathcal{C}$ is a finitely supported subset of $\mathcal{C}$ that also covers X. A nominal topological space X is* compact *if every open cover $\mathcal{C}$ of X has an orbit-finite subcover: there exist $U_1, \ldots, U_n \in \mathcal{C}$ such that $X = \bigcup_{i=1}^{n} \bigcup \mathrm{orb}\, U_i$.*

▶ **Lemma 3.11.** *For every nominal topological space X the following conditions are equivalent:*
1. *The space X is compact.*
2. *Every uniformly finitely supported open cover of X has a finite subcover.*
3. *For every finite set $S \subseteq \mathbb{A}$ the topological space $|X|_S$ is compact.*
4. *For every finite set $S \subseteq \mathbb{A}$ the topological space $\operatorname{orb}_S X$ is compact.*

The Hausdorff property is more subtle: rather than just separation of points, we require separation of S-orbits ("thick points") by disjoint S-supported open neighbourhoods.

▶ **Definition 3.12.** *A nominal topological space X is* (nominal) Hausdorff *if for every finite set $S \subseteq \mathbb{A}$ and every pair $x_1, x_2 \in X$ of points lying in different S-orbits, there exist disjoint S-supported open sets $U_1, U_2 \subseteq X$ such that $x_i \in U_i$ for $i = 1, 2$.*

Note that the nominal Hausdorff condition is clearly equivalent to being able to separate disjoint *S-orbits*: If $\operatorname{orb}_S x_1 \neq \operatorname{orb}_S x_2$, then any two disjoint open S-supported neighbourhoods U_1, U_2 of x_1, x_2 satisfy $\operatorname{orb}_S x_i \subseteq U_i$ for $i = 1, 2$. Note also that $\operatorname{orb}_S x = \{x\}$ whenever $\operatorname{supp} x \subseteq S$, hence the nominal Hausdorff condition implies the ordinary one. For bounded nominal compact Hausdorff spaces, we have a codirected Tychonoff theorem:

▶ **Proposition 3.13.** *For every codirected diagram of non-empty k-bounded nominal compact Hausdorff spaces, the limit in* **nTop** *is a non-empty k-bounded nominal compact Hausdorff space.*

Finally, having a basis of clopen sets is not sufficient in our setting. To see this, note that in an ordinary topological space X every clopen subset $C \subseteq X$ can be represented as $C = f^{-1}[A]$ for some continuous map $f\colon X \to Y$ into a finite discrete space Y and some subset $A \subseteq Y$. (In fact, one may always take $Y = 2$ and $A = \{1\}$.) This is no longer true in the nominal setting, see Remark 3.15 below. Therefore, in lieu of clopens we work with *representable* subsets:

▶ **Definition 3.14.** *A subset $R \subseteq X$ of a nominal space X is* representable *if there exists a continuous map $f\colon X \to Y$ into an orbit-finite discrete space Y such that $R = f^{-1}[A]$ for some $A \in \mathcal{P}_{fs} Y$.*

▶ Remark 3.15.
1. Every representable set is clopen, but the converse generally fails. To see this, consider the discrete space $X = \coprod_{n<\omega} \mathbb{A}^{\#n}$. We show that for fixed $a \in \mathbb{A}$ the (clopen) subset $R = \{x \mid a \in \operatorname{supp} x\} \subseteq X$ is not representable. Towards a contradiction suppose that R is represented by $f\colon X \to Y$ as $R = f^{-1}[A]$ for some $A \in \mathcal{P}_{\mathrm{fs}} Y$. Since Y is orbit-finite, we can choose m large enough such that there exists some $x \in \mathbb{A}^{\#m} \setminus R \subseteq X$ for which $\operatorname{supp} f(x) \subsetneq \operatorname{supp} x$. Choose a name $b \in \operatorname{supp} x \setminus \operatorname{supp} f(x)$. Then $a, b \notin \operatorname{supp} f(x)$, and so we have
$$f((a\ b) \cdot x) = (a\ b) \cdot f(x) = f(x).$$
Since $(a\ b) \cdot x \in R$, this shows $f(x) \in A$ and thus $x \in R$. This contradicts the above choice of x.
2. If a nominal space X has a basis of representable sets, then we may assume without loss of generality that the basic open sets are of the form $f^{-1}[y]$ for some $f\colon X \to Y$ and $y \in Y$, where Y is orbit-finite and discrete. Indeed, if $R = f^{-1}[A]$ for $A \in \mathcal{P}_{\mathrm{fs}} Y$, then $R = \bigcup_{y \in A} f^{-1}[y]$. Moreover, given representable sets $R_i = f_i^{-1}[y_i]$, $i = 1, 2$, the set $R_1 \cap R_2$ is equal to $\langle f_1, f_2 \rangle^{-1}[y_1, y_2]$ and therefore representable as well. Hence, to show that representable subsets form a basis it suffices to check whether every open set is a finitely supported union of subsets of the form $f^{-1}[y]$.

▶ **Definition 3.16.** *A* nominal Stone space *is a nominal compact Hausdorff space with a basis of representables. We let* **nStone** *denote the full subcategory of* **nTop** *given by nominal Stone spaces.*

▶ Remark 3.17. Nominal Stone spaces as per Definition 3.16 are conceptually very different from *nominal Stone spaces with* И, introduced by Gabbay et al. [17] as the dual of *nominal boolean algebras with* И. The latter are equipped with a *restriction operator* **n** tightly related to the freshness quantifier И of nominal sets, which enables a nominal version of the ultrafilter theorem and thus a represention of boolean algebras with И via spaces of ultrafilters. In nominal Stone spaces with И, the Hausdorff property is implicit (but would be analogous to that in standard topology), the basis is given by clopen rather than representable sets, and the notion of compactness (called **n**-*compactness*) considers open covers closed under the operator **n**, which are required to have a *finite* subcover. By this definition, the orbit-finite discrete space $\mathbb{A}$ fails to be compact (the **n**-cover $\{\{a\} \mid a \in \mathbb{A}\} \cup \{\emptyset\}$ has no finite subcover). Hence, given that algebraic recognition is based on orbit-finite sets, nominal Stone spaces with И are not suitable for a topological interpretion of data languages.

▶ **Example 3.18.** Every orbit-finite nominal set can be viewed as a nominal Stone space equipped with the discrete topology. We thus regard $\mathbf{Nom}_{\mathrm{of}}$ as a full subcategory of **nStone**. Nontrivial examples of nominal Stone spaces are given by the spaces of pro-orbit-finite words introduced later.

Within the class of nominal Stone spaces, representable and clopen subsets coincide:

▶ **Lemma 3.19.** *If X is a nominal Stone space, then every clopen set $C \subseteq X$ is representable.*

The following theorem is the key result leading to our topological approach to data languages.

▶ **Theorem 3.20.** *For each $k \in \mathbb{N}$, the category of k-bounded nominal Stone spaces is the Pro-completion of the category of k-bounded orbit-finite nominal sets:*

$$\mathrm{Pro}(\mathbf{Nom}_{\mathrm{of},k}) = \mathbf{nStone}_k.$$

Moreover, k-bounded nominal Stone spaces are precisely the nominal topological spaces arising as codirected limits of k-bounded orbit-finite discrete spaces.

For $k = 0$, we recover the corresponding characterization of classical Stone spaces.

4 Nominal Stone Duality

Next, we give a dual characterization of (bounded) nominal Stone spaces. It builds on the known duality between nominal sets and complete atomic nominal boolean algebras due to Petrişan [32].

▶ **Definition 4.1.** *A* nominal boolean algebra *is a nominal set equipped with the structure of a boolean algebra such that all operations are equivariant. It is* (orbit-finitely) complete *if every (orbit-finite) finitely supported subset has a supremum. A* subalgebra *of an (orbit-finitely) complete nominal boolean algebra is an equivariant subset closed under boolean operations and the respective suprema. Let* $\mathbf{nC_{of}BA}$ *and* **nCBA** *denote the categories of (orbit-finitely) complete nominal boolean algebras; their morphisms are equivariant homomorphisms preserving (orbit-finite) suprema.*

▶ **Definition 4.2.** *An element $x \in B$ of a nominal boolean algebra is an* atom *if $x \neq \bot$ and $y < x$ implies $y = \bot$. The (equivariant) set of atoms of B is denoted $\operatorname{At} B$. The algebra B is* atomic *if every element is the supremum of all atoms below it; if additionally $\operatorname{At} B \in \mathbf{Nom}_{\mathrm{of},k}$ we call it k*-atomic*. If $A \subseteq B$ is a k-atomic subalgebra we write $A \leq_{\mathrm{of},k} B$. An algebra $B \in \mathbf{nCBA}$ is called* locally k-atomic *if every element of B is contained in some $A \leq_{\mathrm{of},k} B$. We denote by $\mathbf{nCA}_k\mathbf{BA} \subseteq \mathbf{nCBA}$ the full subcategory of all k-atomic complete nominal boolean algebras, and $\mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA} \subseteq \mathbf{nC}_{\mathrm{of}}\mathbf{BA}$ denotes the full subcategory of all locally k-atomic orbit-finitely complete nominal boolean algebras.*

▶ Remark 4.3.

1. Orbit-finite completeness is equivalent to the weaker condition that suprema of S-orbits exist for all finite subsets $S \subseteq \mathbb{A}$. In fact, every S-supported orbit-finite subset $X \subseteq B$ is a finite union $X = \bigcup_{i=1}^n \operatorname{orb}_S x_i$ of S-orbits, whence $\bigvee X = \bigvee_{i=1}^n \bigvee \operatorname{orb}_S x_i$.
2. Every k-atomic orbit-finitely complete nominal boolean algebra is complete: For every finitely supported subset $X \subseteq B$ we have $\bigvee X = \bigvee \{b \in \operatorname{At}(B) \mid \exists (x \in X).\ b \leq x\}$, which is a supremum of an orbit-finite subset.

▶ **Theorem 4.4.** *For each $k \in \mathbb{N}$, the category of locally k-atomic orbit-finitely complete nominal boolean algebras is the Ind-completion of the category of k-atomic complete nominal boolean algebras:*

$$\mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA} \simeq \operatorname{Ind}(\mathbf{nCA}_k\mathbf{BA}).$$

▶ **Theorem 4.5** (Nominal Stone Duality). *For each $k \in \mathbb{N}$, the category of locally k-atomic orbit-finitely complete nominal boolean algebras is dual to the category of k-bounded nominal Stone spaces:*

$$\mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA} \simeq^{\mathrm{op}} \mathbf{nStone}_k.$$

Proof. The category $\mathbf{Nom}$ of nominal sets is dually equivalent to the category $\mathbf{nCABA}$ of complete atomic nominal boolean algebras [32]. The duality sends a nominal set X to the boolean algebra $\mathcal{P}_{\mathrm{fs}}X$, equippped with the set-theoretic boolean structure. Conversely, a complete atomic nominal boolean algebra B is mapped to the nominal set $\operatorname{At}(B)$ of its atoms, and an $\mathbf{nCABA}$-morphism $h\colon C \to B$ to the equivariant map $\operatorname{At}(B) \to \operatorname{At}(C)$ sending $b \in \operatorname{At}(B)$ to the unique $c \in \operatorname{At}(C)$ such that $c \leq h(b)$. For every $k \in \mathbb{N}$ the duality clearly restricts to one between k-bounded orbit-finite nominal sets and k-atomic complete nominal boolean algebras. Thus Theorem 4.4 and Theorem 3.20 yield

$$\mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA} \simeq \operatorname{Ind}(\mathbf{nCA}_k\mathbf{BA}) \simeq^{\mathrm{op}} \operatorname{Pro}(\mathbf{nCA}_k\mathbf{BA}^{\mathrm{op}}) \simeq \operatorname{Pro}(\mathbf{Nom}_{\mathrm{of},k}) \simeq \mathbf{nStone}_k. \quad ◀$$

▶ Remark 4.6. We give an explicit description of the dual equivalence of Theorem 4.5.

1. In the direction $\mathbf{nStone}_k \to \mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA}$ it maps a k-bounded nominal Stone space X to the nominal boolean algebra $\operatorname{Clo}(X)$ of clopens (or representables, see Lemma 3.19). A continuous map $f\colon X \to Y$ is mapped to the homomorphism $f^{-1}\colon \operatorname{Clo}(Y) \to \operatorname{Clo}(X)$ taking preimages.
2. The direction $\mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA} \to \mathbf{nStone}_k$ requires some terminology. A finitely supported subset $F \subseteq B$ of an algebra $B \in \mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA}$ is a *nominal orbit-finitely complete prime filter* if (i) $F \neq \emptyset$, (ii) F is upwards closed ($x \in F \wedge x \leq y \Rightarrow y \in F$), (iii) F is downwards directed ($x, y \in F \Rightarrow x \wedge y \in F$), and (iv) for every finitely supported k-bounded orbit-finite subset $X \subseteq B$ such that $\bigvee X \in F$, one has $X \cap F \neq \emptyset$. The equivalence now maps $B \in \mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA}$ to the space $\mathcal{F}_{\mathrm{np}}(B)$ of nominal orbit-finitely complete prime

filters of B, whose topology is generated by the basic open sets $\{F \in \mathcal{F}_{\mathrm{np}}(B) \mid b \in F\}$ for $b \in B$. A morphism $h\colon B \to C$ of $\mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA}$ is mapped to the continuous map $h^{-1}\colon \mathcal{F}_{\mathrm{np}}(C) \to \mathcal{F}_{\mathrm{np}}(B)$ taking preimages.

In Theorem 4.5 we made the support bound k explicit, but we can also leave it implicit. A nominal Stone space is *bounded* if it lies in $\mathbf{nStone}_k$ for some natural number k; similarly, a *locally bounded atomic orbit-finitely complete nominal boolean algebras* is an element of $\mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA}$ for some k.

▶ **Corollary 4.7.** *The category of locally bounded atomic orbit-finitely complete nominal boolean algebras is dual to the category of bounded nominal Stone spaces.*

▶ Remark 4.8. For $k = 0$ we recover the classical Stone duality between boolean algebras and Stone spaces. Indeed, 0-bounded nominal Stone spaces are precisely Stone spaces, and locally 0-atomic orbit-finitely complete nominal boolean algebras are precisely boolean algebras.

5 Pro-Orbit-Finite Words

In this section, we generalize the topological characterization of regular languages to data languages recognizable by orbit-finite nominal monoids [7, 10, 13].

▶ **Definition 5.1.** *A* nominal monoid *M is a monoid object in $\mathbf{Nom}$, that is, it is given by nominal set M equipped with an equivariant associative multiplication $M \times M \to M$ and an equivariant unit $1 \in M$. Nominal monoids and equivariant monoid homomorphisms form a category* $\mathbf{nMon}$.

As for ordinary monoids, the *free monoid* generated by $\Sigma \in \mathbf{Nom}$ is the nominal set Σ^* of finite words (with pointwise group action); its multipliation is concatenation and its unit the empty word.

▶ Remark 5.2. We emphasize the difference between k-bounded nominal monoids – nominal monoids whose carrier is k-bounded – and monoid objects in $\mathbf{Nom}_k$, which are partial nominal monoids where the product $x \cdot y$ is defined iff $|\operatorname{supp} x \cup \operatorname{supp} y| \leq k$.

▶ **Definition 5.3.** *A* data language *over $\Sigma \in \mathbf{Nom}_{\mathrm{of}}$ is a finitely supported subset $L \subseteq \Sigma^*$. It is* recognizable *if there exists an equivariant monoid morphism $h\colon \Sigma^* \to M$ with M orbit-finite and a finitely supported subset $P \subseteq M$ such that $L = h^{-1}[P]$. In this case, we say that the morphism h* recognizes *L.*

For example, the equivariant language L_0 from the Introduction is recognizable, while the language L_1 is not recognizable.

▶ Remark 5.4.
1. The morphism h can be taken to be surjective; otherwise, take its coimage.
2. Via characteristic functions, data languages correspond precisely to finitely supported maps $L\colon \Sigma^* \to 2$, where 2 is the two-element nominal set. Recognizablity then states that L factorizes through some equivariant monoid morphism with orbit-finite codomain.

Recall from Section 2 that the Stone space $\widehat{\Sigma^*}$ of profinite words over a finite alphabet Σ is constructed as the limit in $\mathbf{Stone} \simeq \mathrm{Pro}(\mathbf{Set}_{\mathrm{f}})$ of all finite quotient monoids of Σ^*. The obvious generalization to a *nominal* alphabet $\Sigma \in \mathbf{Nom}_{\mathrm{of}}$, which constructs the limit of all orbit-finite quotient monoids in $\mathrm{Pro}(\mathbf{Nom}_{\mathrm{of}})$, is unlikely to yield a useful object since this category is not concrete (Proposition 3.6); in fact, it is futile from a language-theoretic perspective, cf. Remark 5.15. Instead, our results of Section 3 suggest to restrict the diagram scheme to $\Sigma^*{\downarrow}\mathbf{nMon}_{\mathrm{of},k}$, the poset of k-bounded orbit-finite quotient monoids (where $e \leq e'$ iff e' factorizes through e), and take the limit in $\mathrm{Pro}(\mathbf{Nom}_{\mathrm{of},k}) = \mathbf{nStone}_k$. However, this diagram is not codirected, so its limit may not be a nominal Stone space. We again focus on well-behaved (i.e., codirected), subcategories by introducing *support bounds*.

▶ **Definition 5.5.** *A* support bound *is a map $s\colon \Sigma^* \to \mathcal{P}\mathbb{A}$ such that $s[\Sigma^*] \subseteq \mathcal{P}_k\mathbb{A}$ for some $k \in \mathbb{N}$, where $\mathcal{P}_k\mathbb{A} = \{S \subseteq \mathbb{A} \mid |S| \leq k\}$. We usually identify s with its codomain restrictions to $\mathcal{P}_k\mathbb{A}$ for sufficiently large k. A morphism $h\colon \Sigma^* \to M$ of nominal monoids is* s-bounded *if* $\operatorname{supp} h(w) \subseteq s(w)$ *for all $w \in \Sigma^*$; we write $h\colon \Sigma^* \to_s M$. We denote by $\Sigma^*{\downarrow}_s\mathbf{nMon}_{\mathrm{of},k}$ the subposet of $\Sigma^*{\downarrow}\mathbf{nMon}_{\mathrm{of},k}$ given by s-bounded quotient monoids.*

▶ **Lemma 5.6.** *For every support bound s, the poset $\Sigma^*{\downarrow}_s\mathbf{nMon}_{\mathrm{of},k}$ is codirected.*

Proof. Let $h\colon \Sigma^* \to_s M_h$ and $h'\colon \Sigma^* \to_s M_{h'}$ be two s-bounded quotients in $\Sigma^*{\downarrow}_s\mathbf{nMon}_{\mathrm{of},k}$. Form the coimage $k\colon \Sigma^* \twoheadrightarrow M$ of their pairing $\langle h, h'\rangle\colon \Sigma^* \to M_h \times M_{h'}$. Then for all $w \in \Sigma^*$

$$\operatorname{supp} k(w) = \operatorname{supp}(h(w), h'(w)) = \operatorname{supp} h(w) \cup \operatorname{supp} h'(w) \subseteq s(w).$$

Hence, k is a lower bound for h, h' in the poset $\Sigma^*{\downarrow}_s\mathbf{nMon}_{\mathrm{of},k}$. ◀

▶ **Definition 5.7.** *For an orbit-finite nominal set Σ and a support bound $s\colon \Sigma^* \to \mathcal{P}_k\mathbb{A}$ we define the nominal Stone space $\widehat{\Sigma^*_s}$ to be the limit of the codirected diagram*

$$D\colon \Sigma^*{\downarrow}_s\mathbf{nMon}_{\mathrm{of},k} \to \mathbf{nStone}_k, \qquad (e\colon \Sigma^* \twoheadrightarrow_s M) \;\mapsto\; |M|,$$

where $|M|$ is the nominal set underlying M, regarded as a discrete nominal topological space. The elements of $\widehat{\Sigma^*_s}$ are called the (s-bounded) pro-orbit-finite words *over Σ. We denote by $\hat{e}\colon \widehat{\Sigma^*_s} \to M$ the limit projection associated to $e\colon \Sigma^* \twoheadrightarrow_s M$ in $\Sigma^*{\downarrow}_s\mathbf{nMon}_{\mathrm{of},k}$.*

▶ Remark 5.8.

1. One may equivalently define $\widehat{\Sigma^*_s}$ as the limit of the larger cofiltered diagram D' given by

$$D'\colon \Sigma^*{\downarrow_s}\mathbf{nMon}_{\mathrm{of},k} \to \mathbf{nStone}_k, \qquad (e\colon \Sigma^* \to_s M) \;\mapsto\; |M|,$$

where $\Sigma^*{\downarrow_s}\mathbf{nMon}_{\mathrm{of},k}$ is the category of all equivariant s-bounded monoid morphisms $h\colon \Sigma^* \to_s M$ with k-bounded orbit-finite codomain; a morphism from h to $h'\colon \Sigma^* \to_s M'$ is an equivariant monoid morphism $k\colon M \to M'$ such that $h' = k \cdot h$. In fact, the inclusion $\Sigma^*{\downarrow}_s\mathbf{nMon}_{\mathrm{of},k} \hookrightarrow \Sigma^*{\downarrow_s}\mathbf{nMon}_{\mathrm{of},k}$ is an initial functor, hence the limits of D and D' coincide. Since the limit of D' is formed as in **Set** (Lemma 3.9), the space $\widehat{\Sigma^*_s}$ is carried by the nominal set of compatible families $(x_h)_h$ of D', and the limit projection $\hat{h}$ associated to $h\colon \Sigma^* \to_s M$ is given by $(x_h)_h \mapsto x_h$.
2. The forgetful functor $V\colon \mathbf{nStone}_k \to \mathbf{Nom}_k$ and the inclusion $I\colon \mathbf{Nom}_k \to \mathbf{Nom}$ both preserve codirected limits. The morphisms $\Sigma^*{\downarrow_s}\mathbf{nMon}_{\mathrm{of},k}$ viewed as equivariant functions form a cone for the diagram IVD', so there exists a unique equivariant map $\eta\colon \Sigma^* \to IV\widehat{\Sigma^*_s}$ such that

$$h = \big(\Sigma^* \overset{\eta}{\dashrightarrow} IV\widehat{\Sigma^*_s} \xrightarrow{IV\hat{h}} IVM\big) \qquad \text{for all } h \in \Sigma^*{\downarrow_s}\mathbf{nMon}_{\mathrm{of},k}.$$

In more explicit terms, the map η is given by $\eta(w) = (h(w))_h$ for $w \in \Sigma^*$. For simplicity we omit I and V and write $\eta\colon \Sigma^* \to \widehat{\Sigma^*_s}$. The image of η forms a dense subset of $\widehat{\Sigma^*_s}$. We note that η is generally not injective since we restrict to a subdiagram $\Sigma^*{\downarrow_s}\mathbf{nMon}_{\mathrm{of},k}$ of the diagram $\Sigma^*{\downarrow}\mathbf{nMon}_{\mathrm{of}}$,

3. The space $\widehat{\Sigma_s^*}$ is a nominal monoid with product $\hat{h}(x \cdot y) = \hat{h}(x) \cdot \hat{h}(y)$ and unit $\eta(\varepsilon)$, with ε the empty word. Since the multiplication is readily seen to be continuous, $\widehat{\Sigma_s^*}$ can be regarded as an object of $\mathbf{Mon}(\mathbf{nStone})$, the category of nominal Stone spaces equipped with a continuous monoid structure and continuous equivariant monoid morphisms.

Now recall from Section 2 that the space $\widehat{\Sigma^*}$ can be constructed as the metric completion of Σ^*, where the metric measures the size of separating monoids. We now investigate to what extent the metric approach applies to the nominal setting, using nominal (pseudo-)metrics; see Example 3.2.

▶ **Definition 5.9.** *Let s be a support bound on Σ^*. We say that a nominal monoid M s-separates $v, w \in \Sigma^*$ if there exists an s-bounded equivariant monoid morphism $h\colon \Sigma^* \to_s M$ such that $h(v) \neq h(w)$. We define a nominal pseudometric d_s on Σ^* by setting*

$$d_s(v, w) = \sup\{\, 2^{-|\operatorname{orb} M|} \mid \text{the orbit-finite nominal monoid } M \text{ } s\text{-separates } v, w \,\}.$$

We let Σ^*/d_s denote the corresponding nominal metric space, obtained as a quotient space of the pseudometric space (Σ^*, d_s) by identifying v, w if $d_s(v, w) = 0$.

▶ Remark 5.10. In contrast to the classical case, d_s is generally not a metric: there may exist words $v \neq w$ which are not s-separated by any orbit-finite nominal monoids. For example, if $\Sigma = \mathbb{A}$ and $s(a_1 \cdots a_n) = a_1$ for $a_1, \ldots, a_n \in \Sigma$, then for every s-bounded h and distinct names $a, b, c \in \mathbb{A}$ we have $h(ab) = h((b\ c) \cdot ac) = (b\ c) \cdot h(ac) = h(ac)$ since $b, c \notin s(ac) \supseteq \operatorname{supp} h(ac)$. Therefore, the additional metrization process is required.

For the next lemma we need some terminology. A nominal metric space is *complete* if every finitely supported Cauchy sequence has a limit. A nominal topological space is *completely metrizable* if its topology is induced by a complete metric. A subset $D \subseteq X$ of a nominal metric space is *(topologically) dense* if every open neighbourhood of a point $x \in X$ contains an element of D.

▶ Remark 5.11. In contrast to classical metric spaces, density is not equivalent to *sequential density* (every point $x \in X$ is a limit of a finitely supported sequence in D). To see this, consider the space $\mathbb{A}^\omega$ of finitely supported infinite words with the prefix metric, that is, $d(v, w) = 2^{-n}$ if n is the length of the longest common prefix of v, w. Let $D \subseteq X$ be the equivariant subset given by

$$D = \{\, x \in \mathbb{A}^\omega \mid |\operatorname{supp} x| \geq 2 \text{ and } |\operatorname{supp} x| \geq |\operatorname{initialblock}(x)| \,\},$$

where $\operatorname{initialblock}(x)$ is the longest prefix of x of the form a^n ($a \in \mathbb{A}$). The set D is dense, but not sequentially dense: $a^\omega \in \mathbb{A}^\omega$ is not the limit of any finitely supported sequence in D.

▶ **Lemma 5.12.**
1. *The space $\widehat{\Sigma_s^*}$ is completely metrizable via the complete nominal metric*

$$\hat{d}_s(x, y) = \sup\{\, 2^{-|\operatorname{orb} M|} \mid \exists (h\colon \Sigma^* \to_s M)\colon \hat{h}(x) \neq \hat{h}(y) \,\}. \tag{5.1}$$

2. *The canonical map η (Remark 5.8) yields a dense isometry $\eta\colon (\Sigma^*, d_s) \to (\widehat{\Sigma_s^*}, \hat{d}_s)$.*

▶ Remark 5.13. In classical topology, it would now be clear that $\widehat{\Sigma_s^*}$ is the metric completion of the metric space Σ^*/d_s, i.e. it satisfies the universal property that every uniformly continuous map from Σ^*/d_s to a complete metric space has a unique uniformly continuous extension to $\widehat{\Sigma_s^*}$. However, this rests on the coincidence of topological and sequential density, which fails over nominal sets as seen in Remark 5.11. We therefore conjecture that $\widehat{\Sigma_s^*}$ is not the nominal metric completion of Σ^*/d_s.

By using support bounds, we obtain a topological perspective on recognizable data languages. Let $\mathrm{Rec}_s\,\Sigma$ denote the set of data languages recognized by s-bounded equivariant monoid morphisms.

▶ **Theorem 5.14.** *For every support bound $s\colon \Sigma^* \to \mathcal{P}_k\mathbb{A}$, the k-bounded nominal Stone space $\widehat{\Sigma^*_s}$ of s-bounded pro-orbit-finite words is dual to the locally k-atomic orbit-finitely complete boolean algebra $\mathrm{Rec}_s(\Sigma^*)$ of s-recognizable languages. In particular, we have the isomorphism*

$$\mathrm{Rec}_s(\Sigma^*) \;\cong\; \mathrm{Clo}(\widehat{\Sigma^*_s}) \qquad \textit{in} \qquad \mathbf{nC}_{\mathrm{of}}\mathbf{A}_{\mathrm{l}k}\mathbf{BA}.$$

Proof (Sketch). The isomorphism is illustrated by the two diagrams below:

$$\begin{tikzcd}[column sep=small] L = h^{-1}[P] \arrow[d, mapsto] & \subseteq & \Sigma^* \arrow[r, "h"] \arrow[d, "\eta"'] & M & \supseteq & P \\ \overline{\eta[L]} = \hat{h}^{-1}[P] & \subseteq & \widehat{\Sigma^*_s} \arrow[ur, dashed, "\hat{h}"'] & & & \end{tikzcd} \qquad \begin{tikzcd}[column sep=small] \eta^{-1}[C] = h^{-1}[P] & \subseteq & \Sigma^* \arrow[r, "h"] \arrow[d, "\eta"'] & M \arrow[d, dashed, "\exists p"] & \supseteq & P = p^{-1}[U] \\ C = f^{-1}[U] \arrow[u, mapsto] & \subseteq & \widehat{\Sigma^*_s} \arrow[r, "f"] \arrow[ur, dashed, "\hat{h}"] & Y & \supseteq & U \end{tikzcd}$$

In more detail, if $L \subseteq \Sigma^*$ is s-recognizable, say $L = h^{-1}[P]$ for an s-bounded morphism h, then its corresponding clopen is the topological closure $\overline{\eta[L]} = \hat{h}^{-1}[P]$ represented by the continuous extension $\hat{h}$. Conversely, every clopen $C \subseteq \widehat{\Sigma^*_s}$ restricts to an s-recognizable language $\eta^{-1}[C] \subseteq \Sigma^*$. We get s-recognizability of $\eta^{-1}[C]$ by factorizing a representation $f\colon \widehat{\Sigma^*_s} \to Y$ of C through a limit projection $\hat{h}$ as $f = p\cdot\hat{h}$, using that Y is finitely copresentable. Thus h recognizes $\eta^{-1}[C]$. ◀

▶ Remark 5.15. In the proof of Theorem 5.14, finite copresentability of orbit-finite sets is crucial to recover recognizable languages from representable subsets, highlighting the importance of working in the Pro-completion $\mathrm{Pro}(\mathbf{Nom}_{\mathrm{of},k}) = \mathbf{nStone}_k$. In a naive approach one might instead want to consider the limit of the diagram $D\colon \Sigma^*{\downarrow}\mathbf{nMon}_{\mathrm{of}} \to \mathbf{nTop}$ of *all* equivariant morphisms from Σ^* to orbit-finite monoids. The resulting space $\widehat{\Sigma^*}$ is still a nominal Hausdorff space with a basis of representables, but it generally fails to be compact, and its representable subsets do not correspond to recognizable data languages. To see this, consider the space $\widehat{\mathbb{A}^*}$ and the orbit-finite nominal monoids $\mathbb{A}^{\leq n}$ (words of length at most n) with multiplication cutting off after n letters. We denote by $h_n\colon \mathbb{A}^* \to \mathbb{A}^{\leq n}$ and $p_{k,n}\colon \mathbb{A}^{\leq k} \twoheadrightarrow \mathbb{A}^{\leq n}$, $n \leq k$, the equivariant monoid morphisms given by projection to the first n letters. For every compatible family $x = (x_h) \in \widehat{\mathbb{A}^*}$ its subfamily $(x_{h_n})_{n\in\mathbb{N}}$ corresponds to a (possibly infinite) word over $\mathbb{A}$ with finite support. Hence there exists a largest natural number $N = N(x)$ such that $|\operatorname{supp} x_{h_N}| = N$. The subsets $C_n = \{x \in \widehat{\mathbb{A}^*} \mid N(x) = n\}$, $n \in \mathbb{N}$, are equivariant clopens since $C_n = \hat{h}_n^{-1}[\mathbb{A}^{\#n}] \cap \hat{h}_{n+1}^{-1}[\mathbb{A}^{\leq n+1} \setminus \mathbb{A}^{\#(n+1)}]$. Thus each C_n is representable (by a continuous map into the two-element discrete space), non-empty (since $\eta(w) = (h(w))_h \in C_n$ for every word $w \in \mathbb{A}^{\#n} \subseteq \mathbb{A}^*$ of pairwise distinct letters), and pairwise disjoint. Hence they form a cover of $\widehat{\mathbb{A}^*}$ that admits no orbit-finite (equivalently, finite) subcover, showing that $\widehat{\mathbb{A}^*}$ is not compact. Moreover, the sets $C_M = \bigcup_{m\in M} C_m$, where $M \subseteq \mathbb{N}$, are equivariant clopens (hence representable) and pairwise distinct. Thus $\widehat{\mathbb{A}^*}$ has uncountably many clopens. On the other hand, there exist only countably many recognizable languages over $\mathbb{A}$ (using that, up to isomorphism, there exist only countably many orbit-finite sets [36, Thm. 5.13] and thus countably many orbit-finite nominal monoids), showing that there is no bijective correspondence between representable sets in $\widehat{\mathbb{A}^*}$ and recognizable data languages over $\mathbb{A}$.

6 A Nominal Reiterman Theorem

As an application of pro-orbit-finite methods, we present a nominal extension of Reiterman's classical pseudovariety theorem [37]. The latter characterizes classes of finite algebras presentable by profinite equations as precisely those closed under finite products, subalgebras, and homomorphic images. This result has been generalized to first-order structures [34] and, recently, to abstract categories [1, 30]. A key insight for the categorical perspective is that equations should be formed over projective objects. (Recall that an object X in a category is projective w.r.t. a class $\mathcal{E}$ of morphisms if for all cospans $X \xrightarrow{f} Y \xleftarrow{e} Z$ with $e \in \mathcal{E}$ there exists a factorization of f through e.) In **Nom**, one takes strong nominal sets, which are projective with respect to support-reflecting quotients (see Definition 6.1.2). For spaces of pro-orbit-finite words we have the support bound as an additional constraint, which makes the situation more complex: In a cospan $\widehat{\Sigma_s^*} \xrightarrow{\hat{h}} N \overset{e}{\twoheadleftarrow} M$ with e support-reflecting, no s-bounded factorization of $\hat{h}$ through e may exist. Surprisingly, there nonetheless exists a suitable type of quotients for nominal monoids, called *MSR quotients*, which is *independent* of the support bound s.

▶ **Definition 6.1.** *A surjective equivariant morphism $e\colon M \twoheadrightarrow N$ of nominal monoids is*

1. support-preserving *if* $\operatorname{supp} e(x) = \operatorname{supp} x$ *for every* $x \in X$*;*
2. support-reflecting *if for every* $y \in Y$ *there exists* $x \in e^{-1}[y]$ *such that* $\operatorname{supp} x = \operatorname{supp} y$*;*
3. multiplicatively support-reflecting *(*MSR *for short) if there exists a nominal submonoid $M' \subseteq M$ such that the domain restriction $e|_{M'}\colon M' \to N$ of e is surjective and support-preserving.*

▶ Remark 6.2. Note that a surjective morphism e is support-reflecting iff it restricts to a support-preserving surjection $e|_{M'}$ for some equivariant subset $M' \subseteq M$. For MSR morphisms one additionally requires that M' may be chosen to form a submonoid. So the implications

$$\text{support-preserving} \implies \text{multiplicatively support-reflecting} \implies \text{support-reflecting}$$

hold, but none of the two converses holds in general; for the first one consider the morphism $\mathbb{A}^* \twoheadrightarrow 1$ into the trivial monoid, and for the second one see Example 6.11.

▶ **Proposition 6.3.** *A surjective equivariant morphism $e\colon M \twoheadrightarrow N$ between orbit-finite nominal monoids is MSR iff all the monoids $\widehat{\Sigma_s^*}$ (where $\Sigma \in \mathbf{Nom}_{\mathrm{of}}$ is strong and $s\colon \Sigma^* \to \mathcal{P}\mathbb{A}$ is a support bound) are projective with respect to e in $\mathbf{Mon}(\mathbf{nStone})$, with M and N regarded as discrete spaces.*

▶ **Definition 6.4.** *An* MSR-pseudovariety of nominal monoids *is a class $\mathcal{V} \subseteq \mathbf{nMon}_{\mathrm{of}}$ of orbit-finite nominal monoids closed under*

1. *finite products: if $M_1, \dots, M_n \in \mathcal{V}$, $n \in \mathbb{N}$, then $M_1 \times \cdots \times M_n \in \mathcal{V}$;*
2. *submonoids: if $M \in \mathcal{V}$ and $N \subseteq M$ is a nominal submonoid, then $N \in \mathcal{V}$:*
3. *MSR quotients: if $M \in \mathcal{V}$ and $e\colon M \twoheadrightarrow N$ is an MSR quotient, then $N \in \mathcal{V}$.*

▶ **Definition 6.5.** *Let $s\colon \Sigma^* \to \mathcal{P}\mathbb{A}$ be a support bound. A* morphic pro-orbit-finite equation*, or* morphic proequation *for short, is a surjective* **nStone***-morphism $\varphi\colon \widehat{\Sigma_s^*} \twoheadrightarrow E$. An orbit-finite monoid M* satisfies *φ if for every s-bounded morphism $h\colon \Sigma^* \to M$, the limit projection $\hat{h}\colon \widehat{\Sigma_s^*} \to M$ factorizes through φ in $\mathbf{nStone}_k$, for some $k \in \mathbb{N}$ such that $M \in \mathbf{Nom}_{\mathrm{of},k}$ and s corestricts to $\mathcal{P}_k\mathbb{A}$:*

$$\hat{h} = \big(\widehat{\Sigma_s^*} \overset{\varphi}{\twoheadrightarrow} E \overset{\exists}{\dashrightarrow} M\big).$$

For a set $\mathcal{T}$ of morphic proequations, taken over possibly different $\widehat{\Sigma_s^*}$, we denote by $\mathcal{V}(\mathcal{T})$ the class of orbit-finite monoids satisfying all proequations in $\mathcal{T}$. A class $\mathcal{V}$ of orbit-finite monoids is presentable by morphic proequations *if $\mathcal{V} = \mathcal{V}(\mathcal{T})$ for some set $\mathcal{T}$ of morphic proequations.*

Note that proequations use support bounds, while the definition of an MSR-pseudovariety does not.

▶ **Theorem 6.6** (Nominal Reiterman)**.** *A class of orbit-finite nominal monoids is an MSR-pseudovariety iff it is presentable by morphic proequations.*

The main technical observations for the proof are that (i) every orbit-finite set is k-bounded for some k, hence finitely copresentable in $\mathbf{nStone}_k$, and (ii) there are "enough" proequations in the sense that every orbit-finite nominal monoid is a quotient of some $\widehat{\Sigma_s^*}$. The quotient is not necessarily MSR, which entails that abstract pseudovariety theorems [1, 30] do not apply to our present setting.

We also give a syntactic version of our nominal Reiterman theorem, which uses explicit proequations in lieu of morphic proequations.

▶ **Definition 6.7.** *An* explicit proequation *is a pair $(x, y) \in \widehat{\Sigma_s^*} \times \widehat{\Sigma_s^*}$ for some strong $\Sigma \in \mathbf{Nom}_{\mathrm{of}}$ and some support bound s, denoted by $x = y$. An orbit-finite monoid M* satisfies *the explicit proequation $x = y$ if*

$$\hat{h}(x) = \hat{h}(y) \qquad \text{for every } s\text{-bounded equivariant monoid morphism } h\colon \Sigma^* \to M.$$

(Here choose a common support size bound k for M and s, so that $\hat{h}$ lies in $\mathbf{nStone}_k$.)

▶ **Theorem 6.8** (Explicit Nominal Reiterman)**.** *A class of orbit-finite nominal monoids is an MSR-pseudovariety iff it is presentable by explicit proequations.*

▶ **Example 6.9.** Recall that in a finite monoid M every element m has a unique idempotent power, denoted by m^ω. This holds analogously for orbit-finite nominal monoids M [7, Theorem 5.1]: one has $m^\omega = m^{(n\cdot k!)!}$ where n is the number of orbits M and k is the maximum support size. (The number $n \cdot k!$ is an upper bound on the number of elements of M with any given finite support [36, Thm. 5.13], hence on the cardinality of the set $\{m^i : i \in \mathbb{N}\}$.) The nominal monoid M is *aperiodic* if $m^\omega \cdot m = m^\omega$ for all $m \in M$. Languages recognizable by aperiodic orbit-finite monoids are captured precisely by first-order logic on data words [7, 13]. One readily verifies that the class of aperiodic orbit-finite monoids forms an MSR-pseudovariety; in fact, it is closed under all quotients. To present it by pro-orbit-finite equations, note that for every $x \in \widehat{\Sigma_s^*}$ the family $x^\omega = (\hat{h}(x)^\omega)_h$ is again compatible, hence $x^\omega \in \widehat{\Sigma_s^*}$. If $s\colon \Sigma^* \to \mathcal{P}_k\mathbb{A}$ and $h\colon \Sigma^* \to_s M$ is an s-bounded equivariant monoid morphism such that M has at most n orbits, then $\hat{h}(x^\omega) = \hat{h}(x)^\omega = \hat{h}(x)^{(n\cdot k!)!} = \hat{h}(x^{(n\cdot k!)!})$, hence $\hat{d}_s(x^\omega, x^{(n\cdot k!)!}) < 2^{-n}$ in the metric (5.1) on $\widehat{\Sigma_s^*}$. This shows that x^ω is the limit of the sequence $(x^{(n\cdot k!)!})_{n\in\mathbb{N}}$ in $\widehat{\Sigma_s^*}$, and moreover that the pseudovariety of aperiodic orbit-finite monoids is presented by the explicit proequations $x^\omega \cdot x = x^\omega$, where $x \in \widehat{\Sigma_s^*}$ and $s\colon \Sigma^* \to \mathcal{P}_k\mathbb{A}$ ranges over all support bounds on strong orbit-finite alphabets. Restricting to $k = 0$, we recover the well-known description of aperiodic finite monoids by the (single) profinite equation $x^\omega \cdot x = x^\omega$.

▶ Remark 6.10. **1.** Pseudovarieties of finite monoids admit an alternative equational characterization based on sequences of word equations rather than profinite equations. A *word equation* is a pair $(v, w) \in \Sigma^* \times \Sigma^*$ of words over some finite alphabet Σ, denoted $v = w$; it is *satisfied* by a monoid M if $h(v) = h(w)$ for every monoid morphism $h\colon \Sigma^* \to M$.

More generally, a sequence $(v_0 = w_0, v_1 = w_1, \ldots)$ of word equations, taken over possibly different finite alphabets, is *eventually satisfied* by M if it satisfies all but finitely many of the equations. As shown by Eilenberg and Schützenberger [15], a class of finite monoids forms a pseudovariety iff it is presentable by a (single) sequence of word equations.

2. Recently, a nominal version of the Eilenberg-Schützenberger theorem by Urbat and Milius [44]. They consider nominal word equations (defined as above, where Σ is now a strong orbit-finite nominal set) and show that sequences of nominal word equations present precisely *weak pseudovarieties*, i.e. classes of orbit-finite nominal monoids closed under finite products, submonoids, and support-reflecting quotients. Clearly every MSR-pseudovariety is weak, but the converse does not hold; hence over nominal sets, sequences of word equations and pro-orbit-finite equations are of different expressivity. The example below illustrates one source of additional expressivity of pro-orbit-finite equations: The support bound s can control how the support changes during multiplication, which is not expressible by sequences of word equations.

▶ **Example 6.11.** An example of an MSR-pseudovariety that is not a weak pseudovariety is given by the class $\mathcal{V}$ of all orbit-finite nominal monoids M such that

$$\forall (m, n \in M)\colon \quad \operatorname{supp}(mn) = \emptyset \quad \iff \quad \operatorname{supp}(m, n) = \emptyset. \tag{6.1}$$

(Note that $\operatorname{supp}(m, n) = \operatorname{supp} m \cup \operatorname{supp} n$ and that "$\Leftarrow$" always holds by equivariance of the monoid multiplication.) It is not difficult to prove that $\mathcal{V}$ is an MSR-pseudovariety. To show that $\mathcal{V}$ is not a weak pseudovariety, we construct a support-reflecting quotient under which $\mathcal{V}$ is not closed. The nominal set $\overline{1} + \overline{\mathbb{A}} = \{\overline{1}\} + \{\overline{a} \mid a \in \mathbb{A}\}$ forms a nominal monoid with multiplication given by projection on the first component and unit $\overline{1}$. We extend the multiplication to the nominal set $M = 1 + \mathbb{A} + \overline{1} + \overline{\mathbb{A}}$ by letting 1 be the unit and setting $x \cdot y = \overline{x} \cdot \overline{y}$ whenever $x, y \neq 1$; here overlining is idempotent ($\overline{\overline{x}} := \overline{x}$). This makes the multiplication associative and equivariant. Thus, M is a nominal monoid. Now let $N = 1 + \mathbb{A} + 0 = \{1\} + \mathbb{A} + \{0\}$ be the nominal monoid with multiplication $x \cdot y = 0$ for $x, y \neq 1$. Thus 0 is an absorbing element. Letting $\mathrm{const}_0\colon \overline{1} + \overline{\mathbb{A}} \to 0$ denote the constant map, we have the equivariant surjective map

$$e = \mathrm{id}_{1+\mathbb{A}} + \mathrm{const}_0\colon M = (1 + \mathbb{A}) + (\overline{1} + \overline{\mathbb{A}}) \twoheadrightarrow (1 + \mathbb{A}) + 0 = N.$$

Note that e is a monoid morphism: it maps 1 to 1 and if $x, y \neq 1$ then $e(x), e(y) \neq 1$ and hence $e(x \cdot y) = e(\overline{x} \cdot \overline{y}) = 0 = e(x) \cdot e(y)$. The quotient e is support-reflecting, but it is not MSR: the subset $1 + \mathbb{A} + \overline{1} \subseteq M$ of support-preserving elements does not form a submonoid of M. Finally, clearly M satisfies (6.1) while N does not.

7 Conclusion and Future Work

We have introduced topological methods to the theory of data languages, and also explored some of their subtleties and limitations. Following the spirit of Marshall Stone's slogan *"always topologize"*, the core insight of our paper may be summarized as:

Data languages topologize for bounded supports.

In fact, by restricting to support-bounded orbit-finite nominal sets and analyzing their Pro-completion, we have shown that fundamental results from profinite topology (notably Stone duality and the equivalence between profinite spaces and Stone spaces) generalize to the pro-orbit-finite world. These results are of independent interest; in particular, they are potentially applicable to data languages recognizable by all kinds of orbit-finite structures. For the case of monoids, we derived a topological interpretation of recognizable data languages via clopen sets of pro-orbit-finite words, as well as a nominal version of Reiterman's pseudovariety theorem characterizing the expressive power of pro-orbit-finite equations.

The foundations laid in the present paper open up a number of promising directions for future research. One first goal is to develop a fully fledged duality theory for data languages along the lines of the work of Gehrke et al. [18] on classical regular languages, based on an *extended nominal Stone duality* between pro-orbit-finite monoids and nominal boolean algebras with operators.

Regarding specific applications, we aim to analyze further classes of orbit-finite monoids in terms of pro-orbit-finite equations, following the lines of Example 6.9, in order to classify the corresponding data languages. One natural candidate is the class of $\mathcal{J}$-trivial monoids, with the vision of a nominal version of Simon's theorem [41] relating $\mathcal{J}$-triviality to existential first-order logic on data words.

Finally, we aim to extend our topological theory of recognizable data languages, and the corresponding nominal Reiterman theorem, to algebraic structures beyond orbit-finite monoids. Potential instances include algebras for a signature Σ, which serve as recognizers for data tree languages, infinitary structures such as nominal ω-semigroups [45], modeling languages of infinite data words, and algebraic structures with binders, which we expect to bear interesting connections to data languages with binders and their automata models [39,43].

References

1. Jiří Adámek, Liang-Ting Chen, Stefan Milius, and Henning Urbat. Reiterman's theorem on finite algebras for a monad. *ACM Trans. Comput. Log.*, 22(4):23:1–23:48, 2021.
2. Jiří Adámek and Jiří Rosický. *Locally Presentable and Accessible Categories*. London Mathematical Society Lecture Note Series. Cambridge University Press, 1994.
3. Jorge Almeida. Profinite semigroups and applications. In *Structural Theory of Automata, Semigroups, and Universal Algebra*, pages 1–45. Springer Netherlands, 2005.
4. Jorge Almeida and Alfredo Costa. Profinite topologies. In Jean-Éric Pin, editor, *Handbook of Automata Theory*, pages 615–652. European Mathematical Society Publishing House, Zürich, Switzerland, 2021.
5. Michał Bielecki, Jan Hidders, Jan Paredaens, Jerzy Tyszkiewicz, and Jan Van den Bussche. Navigating with a browser. In *ICALP 2002*, volume 2380 of *LNCS*, pages 764–775. Springer, 2002.
6. Garrett Birkhoff. On the Structure of Abstract Algebras. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(4):433–454, 1935.
7. Mikołaj Bojańczyk. Nominal monoids. *Theory Comput. Syst.*, 53(2):194–222, 2013. `doi:10.1007/s00224-013-9464-1`.
8. Mikołaj Bojańczyk. Recognisable languages over monads. In *DLT 2015*, volume 9168 of *LNCS*, pages 1–13. Springer, 2015.
9. Mikołaj Bojańczyk, Bartek Klin, and Sławomir Lasota. Automata theory in nominal sets. *Log. Methods Comput. Sci.*, 10(3), 2014. `doi:10.2168/LMCS-10(3:4)2014`.
10. Mikołaj Bojańczyk and Rafał Stefański. Single-use automata and transducers for infinite alphabets. In *ICALP 2020*, volume 168 of *LIPIcs*, pages 113:1–113:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. `doi:10.4230/LIPIcs.ICALP.2020.113`.
11. Benedikt Bollig, Peter Habermehl, Martin Leucker, and Benjamin Monmege. A robust class of data languages and an application to learning. *Log. Meth. Comput. Sci.*, 10(4:19):23pp., 2014. `doi:10.2168/LMCS-10(4:19)2014`.

12 Liang-Ting Chen, Jiří Adámek, Stefan Milius, and Henning Urbat. Profinite monads, profinite equations, and Reiterman's theorem. In *FOSSACS 2019*, volume 9634 of *LNCS*, pages 531–547. Springer, 2016. doi:10.1007/978-3-662-49630-5_31.

13 Thomas Colcombet, Clemens Ley, and Gabriele Puppis. Logics with rigidly guarded data tests. *Log. Methods Comput. Sci.*, 11(3), 2015.

14 Samuel Eilenberg. *Automata, Languages, and Machines.* Elsevier Science, 1974.

15 Samuel Eilenberg and Marcel-Paul Schützenberger. On pseudovarieties. *Advances Math.*, 10:413–418, 1976.

16 Murdoch James Gabbay. Nominal algebra and the HSP theorem. *J. Log. Comput.*, 19(2):341–367, 2009. doi:10.1093/logcom/exn055.

17 Murdoch James Gabbay, Tadeusz Litak, and Daniela Petrisan. Stone Duality for Nominal Boolean Algebras with И. In *CALCO 2011*, volume 6859 of *LNCS*, pages 192–207. Springer, 2011. doi:10.1007/978-3-642-22944-2_14.

18 Mai Gehrke, Serge Grigorieff, and Jean-Eric Pin. Duality and equational theory of regular languages. In *ICALP 2008*, volume 5126 of *LNCS*, pages 246–257. Springer, 2008.

19 Mai Gehrke, Serge Grigorieff, and Jean-Eric Pin. A topological approach to recognition. In *ICALP 2010*, volume 6199 of *LNCS*, pages 151–162. Springer, 2010.

20 Mai Gehrke, Daniela Petrişan, and Luca Reggio. The Schützenberger product for syntactic spaces. In *ICALP 2016*, volume 55 of *LIPIcs*, pages 112:1–112:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

21 Mai Gehrke, Daniela Petrişan, and Luca Reggio. Quantifiers on languages and codensity monads. In *LICS 2017*, pages 1–12. IEEE Computer Society, 2017.

22 Radu Grigore, Dino Distefano, Rasmus Petersen, and Nikos Tzevelekos. Runtime verification based on register automata. In *TACAS 2013*, volume 7795 of *LNCS*, pages 260–276. Springer, 2013. doi:10.1007/978-3-642-36742-7_19.

23 Matthew Hennessy. A fully abstract denotational semantics for the pi-calculus. *Theoret. Comput. Sci.*, 278:53–89, 2002. doi:10.1016/S0304-3975(00)00331-5.

24 Michael Kaminski and Nissim Francez. Finite-memory automata. *Theor. Comput. Sci.*, 134(2):329–363, 1994. doi:10.1016/0304-3975(94)90242-9.

25 Kenneth Krohn and John Rhodes. Algebraic theory of machines. I. Prime decomposition theorem for finite semigroups and machines. *Trans. Am. Math. Soc.*, 116:450–464, 1965.

26 Klaas Kürtz, Ralf Küsters, and Thomas Wilke. Selecting theories and nonce generation for recursive protocols. In *FSME 2007*, pages 61–70. ACM, 2007.

27 Alexander Kurz and Daniela Petrisan. On universal algebra over nominal sets. *Math. Struct. Comput. Sci.*, 20(2):285–318, 2010. doi:10.1017/S0960129509990399.

28 Saunders MacLane. *Categories for the Working Mathematician.* Springer-Verlag, 1971.

29 R. McNaughton and S. Papert. *Counter-free Automata.* M.I.T. Press research monographs. M.I.T. Press, 1971.

30 Stefan Milius and Henning Urbat. Equational axiomatization of algebras with structure. In *FOSSACS 2019*, volume 11425 of *LNCS*, pages 400–417. Springer, 2019. doi:10.1007/978-3-030-17127-8_23.

31 Frank Neven, Thomas Schwentick, and Victor Vianu. Finite state machines for strings over infinite alphabets. *ACM Trans. Comput. Log.*, 5(3):403–435, 2004. doi:10.1145/1013560.1013562.

32 Daniela Petrişan. *Investigations into Algebra and Topology over Nominal Sets.* PhD thesis, University of Leicester, 2012.

33 Jean-Eric Pin. Profinite methods in automata theory. In *STACS 2009*, volume 3 of *LIPIcs*, pages 31–50. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany, 2009. doi:10.4230/LIPIcs.STACS.2009.1856.

34 Jean-Eric Pin and Pascal Weil. A Reiterman theorem for pseudovarieties of finite first-order structures. *Algebra Universalis*, 35(4):577–595, 1996.

35 Nicholas Pippenger. Regular languages and Stone duality. *Theory Comput. Syst.*, 30(2):121–134, 1997. doi:10.1007/s002240000045.

36 Andrew M. Pitts. *Nominal Sets: Names and Symmetry in Computer Science*. Cambridge University Press, 2013.

37 Jan Reiterman. The Birkhoff theorem for finite algebras. *Algebra Universalis*, 14(1):1–10, 1982.

38 Julian Salamanca. Unveiling eilenberg-type correspondences: Birkhoff's theorem for (finite) algebras + duality. *CoRR*, 2017. arXiv:1702.02822.

39 Lutz Schröder, Dexter Kozen, Stefan Milius, and Thorsten Wißmann. Nominal automata with name binding. In *FOSSACS 2017*, pages 124–142, 2017.

40 Marcel Paul Schützenberger. On finite monoids having only trivial subgroups. *Information and Control*, 8(2):190–194, 1965.

41 Imre Simon. Piecewise testable events. In H. Brakhage, editor, *Automata Theory and Formal Languages*, pages 214–222. Springer, 1975.

42 Henning Urbat, Jiří Adámek, Liang-Ting Chen, and Stefan Milius. Eilenberg theorems for free. In *MFCS 2017*, volume 83 of *LIPIcs*, pages 43:1–43:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.MFCS.2017.43.

43 Henning Urbat, Daniel Hausmann, Stefan Milius, and Lutz Schröder. Nominal Büchi automata with name allocation. In *CONCUR 2021*, pages 4:1–4:16, 2021. doi:10.4230/LIPIcs.CONCUR.2021.4.

44 Henning Urbat and Stefan Milius. Varieties of data languages. In *ICALP 2019*, volume 132 of *LIPIcs*, pages 130:1–130:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.ICALP.2019.130.

45 Thomas Wilke. An Eilenberg theorem for infinity-languages. In *ICALP 1991*, volume 510 of *LNCS*, pages 588–599. Springer, 1991.