

Coalgebraic Constructions of Canonical Nondeterministic Automata

Robert S. R. Myers^a, Jiří Adámek^a, Stefan Milius^{b,1}, Henning Urbat^{a,*}

^a*Institut für Theoretische Informatik
Technische Universität Braunschweig, Germany*

^b*Lehrstuhl für Theoretische Informatik
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany*

Abstract

For each regular language L we describe a family of canonical nondeterministic acceptors (nfas). Their construction follows a uniform recipe: build the minimal dfa for L in a locally finite variety \mathcal{V} , and apply an equivalence between the category of finite \mathcal{V} -algebras and a suitable category of finite structured sets and relations. By instantiating this to different varieties, we recover three well-studied canonical nfes: $\mathcal{V} =$ boolean algebras yields the átomaton of Brzozowski and Tamm, $\mathcal{V} =$ semilattices yields the jiomaton of Denis, Lemay and Terlutte, and $\mathcal{V} = \mathbb{Z}_2$ -vector spaces yields the minimal xor automaton of Vuillemin and Gama. Moreover, we obtain a new canonical nfa called the distromaton by taking $\mathcal{V} =$ distributive lattices. Each of these nfes is shown to be minimal relative to a suitable measure, and we derive sufficient conditions for their state-minimality. Our approach is coalgebraic, exhibiting additional structure and universal properties of the canonical nfes.

Keywords: non-deterministic automata, join-semilattices, coalgebras, minimization

1. Introduction

One of the core topics in classical automata theory is the construction of state-minimal acceptors for a given regular language. It is well known that the difficulty of this task depends on whether one has deterministic or nondeterministic acceptors in mind. First, every regular language L is accepted by a unique minimal *deterministic* finite automaton (dfa). Following a classical construction due to Brzozowski [12], the state set Q_L of the minimal dfa consists

*Corresponding author

Email addresses: me@robmyers.gmail.com (Robert S. R. Myers),
adamek@iti.cs.tu-bs.de (Jiří Adámek), mail@stefan-milius.eu (Stefan Milius),
urbat@iti.cs.tu-bs.de (Henning Urbat)

¹Stefan Milius acknowledges support by the Deutsche Forschungsgemeinschaft (DFG) under project MI 717/5-1

of all left derivatives of L ; see Example 2.12. For *nondeterministic* finite automata (nfas) the situation is significantly more complex: a regular language may have many non-isomorphic state-minimal nfes, and generally there is no way to identify a “canonical” one among them. However, several authors proposed nondeterministic acceptors that *are* in some sense canonical (though not necessarily state-minimal), e.g. the *átomaton* of Brzozowski and Tamm [11], the *jiromaton*¹ of Denis, Lemay and Terlutte [13], and the *minimal xor automaton* of Vuillemin and Gama [25]. In each case, the respective nfa is formed by closing the set Q_L of left derivatives under certain algebraic operations and taking a minimal set of generators as states. Specifically:

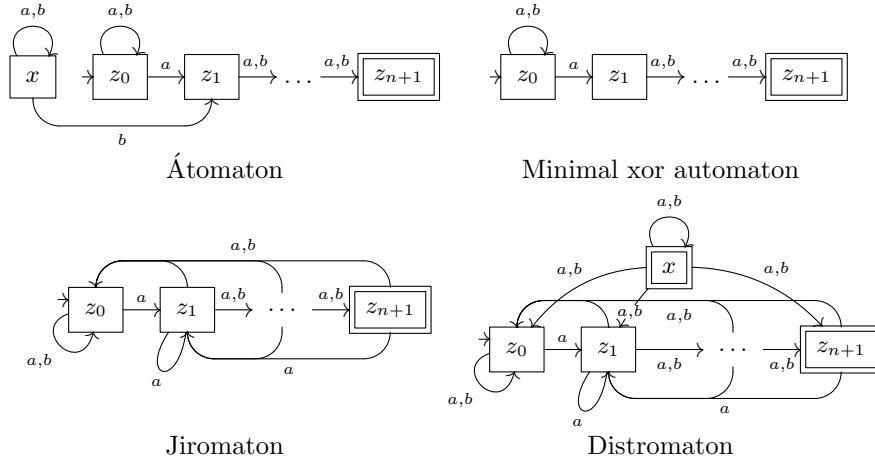
1. The states of the *átomaton* are the atoms of the boolean algebra generated by Q_L , obtained by closing Q_L under finite union, finite intersection and complement.
2. The states of the *jiromaton* are the join-irreducibles of the join-semilattice generated by Q_L , obtained by closing Q_L under finite union.
3. The states of the minimal xor automaton form a basis for the \mathbb{Z}_2 -vector space generated by Q_L . Recall that the \mathbb{Z}_2 -vector space with basis B is the set of all finite subsets of B with \emptyset as the zero vector and addition given by symmetric difference $M \oplus N = (M \setminus N) \cup (N \setminus M)$. Thus the states of the minimal xor automaton are obtained by closing Q_L under symmetric difference and choosing a basis of the resulting \mathbb{Z}_2 -vector space.

Note that the minimal xor automaton differs substantially from the other examples treated in our paper w.r.t. the manner of language acceptance: here we consider acceptance of \mathbb{Z}_2 -weighted languages. That is, a state accepts a word iff the number of accepting paths is odd. In the present paper we demonstrate that all these canonical nfes arise from a coalgebraic construction. For this purpose we first consider *deterministic* automata interpreted in a locally finite variety \mathcal{V} , where *locally finite* means that finitely generated algebras are finite. The three examples above correspond to the variety \mathcal{V} of boolean algebras, join-semilattices and \mathbb{Z}_2 -vector spaces, respectively. A *deterministic \mathcal{V} -automaton* is a coalgebra for the endofunctor $T_\Sigma = \mathbb{2} \times \text{Id}^\Sigma$ on \mathcal{V} , for a fixed two-element algebra $\mathbb{2}$. In Section 2 we describe a Brzozowski-like construction that yields, for every regular language, the minimal deterministic finite \mathcal{V} -automaton accepting it. Next, for certain varieties \mathcal{V} of interest, we derive an equivalence between the full subcategory \mathcal{V}_f of finite algebras and a suitable category $\bar{\mathcal{V}}$ of finite structured sets, whose morphisms are relations preserving the structure. In each case, the objects of $\bar{\mathcal{V}}$ are “small” representations of their counterparts in \mathcal{V}_f , based on specific generators of algebras in \mathcal{V}_f . The equivalence $\mathcal{V}_f \cong \bar{\mathcal{V}}$ then induces an equivalence between deterministic finite \mathcal{V} -automata and coalgebras in $\bar{\mathcal{V}}$ which are *nondeterministic* automata.

¹In [13] the authors called their acceptor “canonical residual finite state automaton”. We propose the shorter “jiromaton” because this is analogous to the *átomaton* terminology.

Hence we have the following two-step procedure for constructing a canonical nfa for a given regular language L : (i) form the minimal deterministic \mathcal{V} -automaton accepting L , and (ii) use the equivalence of \mathcal{V}_f and $\bar{\mathcal{V}}$ to obtain an equivalent nfa. We explain this in Section 3 and show that applying this to different varieties \mathcal{V} yields the three canonical nfes mentioned above. For the átomaton one takes $\mathcal{V} = \mathbf{BA}$ (boolean algebras). Then the minimal deterministic \mathbf{BA} -automaton for L arises from the minimal dfa by closing its states Q_L under boolean operations. The category $\bar{\mathcal{V}} = \bar{\mathbf{BA}}$ is based on Stone duality: $\bar{\mathbf{BA}}$ is the dual of the category of finite sets, so it has as objects all finite sets and as morphisms all converse-functional relations. The equivalence functor $\mathbf{BA}_f \xrightarrow{\cong} \bar{\mathbf{BA}}$ maps each finite boolean algebra to the set of its atoms. This equivalence applied to the minimal deterministic \mathbf{BA} -automaton for L gives precisely the átomaton. Similarly, by taking $\mathcal{V} = \text{join-semilattices}$ and $\mathcal{V} = \text{vector spaces over } \mathbb{Z}_2$ and describing a suitable equivalence $\mathcal{V}_f \cong \bar{\mathcal{V}}$, we recover the jiomaton and the minimal xor automaton, respectively. Finally, for $\mathcal{V} = \text{distributive lattices}$ we get a new canonical nfa called the *distromaton*, which bears a close resemblance to the universal automaton [20].

Example 1.1. Consider the language $L_n = (a + b)^* a (a + b)^n$ where $n \in \omega$. Its minimal dfa has 2^{n+1} states, see Example 2.12, and its state-minimal nfa has $n+2$ states. The átomaton, minimal xor automaton, jiomaton and distromaton of L_n are the nfes with at most $n+3$ states depicted below; see the Examples 3.20-3.23 for detailed explanations.



Generally, the sizes of the four canonical nfes and the minimal dfa are related as follows:

- (a) All the four canonical nfes can have exponentially fewer states than the minimal dfa.
- (b) The minimal xor automaton and jiomaton have no more states than the minimal dfa.

- (c) The átomaton and distromaton have the same number of states, although their structure can be very different. It can happen that the number of states is exponentially larger than that of the minimal dfa.

In Section 4 we characterize the átomaton, jiomaton, minimal xor automaton and distromaton by minimality properties. This provides an explanation of the canonicity of these acceptors that is missing in the original papers. We then use this additional structure to identify conditions on regular languages that guarantee the *state-minimality* of the canonical nfas. That is, there exists a natural class of languages where *canonical* state-minimal nfas exist and can be computed relatively easily.

Related work. Our paper unifies the constructions of canonical nfas given in [11, 13, 25] from a coalgebraic perspective. Previously, several authors have studied coalgebraic methods for constructing minimal and canonical representatives of machines, including Adámek, Bonchi, Hülsbusch, König, Milius and Silva [1], Adámek, Milius, Moss and Sousa [2] and Bezhanishvili, Kupke and Panangaden [7]. Only the first of these three papers, however, treats the case of nondeterministic automata explicitly – in particular, there the átomaton is recovered as an instance of projecting coalgebras in a Kleisli category into a reflective subcategory. This approach is methodologically rather different from the present paper where a categorical equivalence (rather than a reflection) is the basis for the construction of nfas.

In [11] the authors propose a surprisingly simple algorithm for constructing the átomaton of a language L : take the minimal dfa for the reversed language of L , and reverse this dfa. These steps form a fragment of a classical dfa minimization algorithm due to Brzozowski [12]. Recently Bonchi, Bonsangue, Rutten and Silva [9] gave a (co-)algebraic explanation of this procedure, based on the classical duality between observability and reachability of dfas. We provide another explanation in Section 3.3.

A coalgebraic treatment of linear weighted automata (of which the xor automata considered here are a special case) appears in [8]; that paper also provides procedures for computing the minimal linear weighted automaton.

Finally, our work is somewhat related to work on coalgebraic trace semantics [15]. However, while that work considers coalgebras whose carrier is a free algebra of a variety, we consider coalgebras whose carriers are arbitrary algebras from that variety; this means that we move from a Kleisli category to an Eilenberg-Moore category (cf. [10, 16]).

This paper is a reworked full version of the conference paper [23]. Besides including full proofs it contains additional examples, e.g. a coalgebraic construction of the universal automaton (Example 3.27). We also streamlined the presentation and the proofs of the minimality results in Section 4.

Acknowledgement. The authors are grateful to the anonymous referees for numerous suggestions that improved the presentation of our paper.

2. Deterministic Automata

We start with recalling the concept of a finite automaton. Throughout this paper let us fix a finite input alphabet Σ .

Definition 2.1. (a) A *nondeterministic finite automaton (nfa)* is a triple

$$N = (Z, R_a, F)$$

consisting of a finite set Z of states, transition relations $R_a \subseteq Z \times Z$ for each $a \in \Sigma$ and a set of final states $F \subseteq Z$. *Morphisms* of nfes are the usual bisimulations, i.e., relations that preserve and reflect transitions and final states. If N is equipped with a set of initial states $I \subseteq Z$ we write $N = (Z, R_a, F, I)$ and call N a *pointed nfa*. Any pointed nfa N accepts a language $\mathcal{L}_N(I) \subseteq \Sigma^*$ in the usual way.

(b) A *deterministic finite automaton (dfa)* is an nfa whose transition relations are functions. It is called *pointed* if it is equipped with a single initial state.

Although the goal of our paper is constructing canonical nondeterministic automata, we first consider deterministic ones from a coalgebraic perspective. Given an endofunctor $T : \mathcal{V} \rightarrow \mathcal{V}$ of a category \mathcal{V} , a *T-coalgebra* (Q, γ) consists of an object Q of \mathcal{V} and a morphism $\gamma : Q \rightarrow TQ$. A *coalgebra homomorphism* into another coalgebra (Q', γ') is a morphism $h : Q \rightarrow Q'$ of \mathcal{V} such that $\gamma' \circ h = Th \circ \gamma$. This defines a category $\mathbf{Coalg}(T)$. If it exists, the final object νT of $\mathbf{Coalg}(T)$ is called the *final T-coalgebra*.

A coalgebra (Q, γ) is a *subcoalgebra* of (Q', γ') if there exists a coalgebra homomorphism $m : (Q, \gamma) \rightarrow (Q', \gamma')$ with m a monomorphism in \mathcal{V} , and (Q, γ) is a *quotient coalgebra* of (Q', γ') if there exists a coalgebra homomorphism $e : (Q', \gamma') \rightarrow (Q, \gamma)$ with e a strong epimorphism in \mathcal{V} .

Assumption 2.2. From now on \mathcal{V} is a locally finite variety with a specified two-element algebra $\mathbb{2}$ carried by the set $\{0, 1\}$. That is, \mathcal{V} is the category of algebras for some finitary signature and equations, its morphisms being the usual algebra homomorphisms. That \mathcal{V} is *locally finite* means that its finitely generated algebras are finite, or equivalently that its finitely generated *free* algebras are finite.

Note that monomorphisms and strong epimorphisms in \mathcal{V} are precisely the injective and surjective morphisms, respectively. Hence subcoalgebras are represented by injective coalgebra homomorphisms, and quotient coalgebras are represented by surjective coalgebra homomorphisms.

Example 2.3. (a) The category \mathbf{Set}_* of pointed sets is a locally finite variety, given by the signature with one constant and no equations. We denote the chosen point of a pointed set Q by $q_!$. Let $Q = \mathbb{2}$ be pointed by $q_! = 0$.

(b) The category \mathbf{BA} of boolean algebras is a locally finite variety: a boolean algebra on n generators has at most 2^{2^n} elements. $\mathbb{2}$ is the 2-chain $0 < 1$.

- (c) The category $\mathbf{Vect}(\mathbb{Z}_2)$ of vector spaces over the binary field \mathbb{Z}_2 is a locally finite variety. Here $\mathfrak{2} = \mathbb{Z}_2$, the one-dimensional vector space.
- (d) The category \mathbf{JSL} of (join-)semilattices with a least element 0 is locally finite: the finite powerset $\mathcal{P}_f X$ is the free semilattice on X , so a semilattice on n generators has at most 2^n elements. $\mathfrak{2}$ is the 2-chain $0 < 1$.
- (e) The category \mathbf{DL} of distributive lattices with a least and largest element 0 and 1 is locally finite. Again, $\mathfrak{2}$ is the 2-chain $0 < 1$.

Definition 2.4. A (*deterministic*) \mathcal{V} -automaton is a coalgebra for the functor

$$T_\Sigma : \mathcal{V} \rightarrow \mathcal{V}, \quad T_\Sigma = \mathfrak{2} \times \mathbf{ld}^\Sigma = \mathfrak{2} \times \underbrace{\mathbf{ld} \times \cdots \times \mathbf{ld}}_{|\Sigma|\text{-times}}.$$

Remark 2.5. Hence, by the universal property of the product, a deterministic \mathcal{V} -automaton $Q \xrightarrow{\gamma} \mathfrak{2} \times Q^\Sigma$ is given by an algebra Q of states, a homomorphism $\gamma_\varepsilon : Q \rightarrow \mathfrak{2}$ defining final states via $\gamma_\varepsilon^{-1}(\{1\})$ and, for each $a \in \Sigma$, a homomorphism $\gamma_a : Q \rightarrow Q$ representing the a -transitions. We sometimes write $(Q, \gamma_a, \gamma_\varepsilon)$ instead of (Q, γ) . In particular, deterministic \mathbf{Set} -automata are precisely the classical (possibly infinite) deterministic automata without initial states, shortly *da*'s.

Example 2.6. (a) A deterministic \mathbf{Set}_* -automaton is a da whose carrier is a pointed set and whose chosen state q_\dagger is non-final and a fixpoint of all transition functions γ_a (that is, q_\dagger is a *sink state*). These are the partial automata of [24].

- (b) A deterministic \mathbf{BA} -automaton is a da with a boolean algebra structure on the states Q such that (i) the final states form an ultrafilter, (ii) $q \xrightarrow{a} q'$ and $r \xrightarrow{a} r'$ implies $q \vee r \xrightarrow{a} q' \vee r'$ and $\neg q \xrightarrow{a} \neg q'$, and (iii) 0 is a non-final sink state. Recall that a filter $F \subseteq Q$ is an upper set closed under binary meets. An ultrafilter (also called prime filter) is a filter with $0 \notin F$ and $q \vee q' \in F$ iff $q \in F$ or $q' \in F$. The above conditions (i)-(iii) imply that 1 is a final sink state.
- (c) A deterministic $\mathbf{Vect}(\mathbb{Z}_2)$ -automaton is a da with a \mathbb{Z}_2 -vector space structure on the states Q such that (i) the final states $F \subseteq Q$ satisfy $q+r \in F$ iff either $q \in F$ or $r \in F$ but not both, (ii) $q \xrightarrow{a} q'$ and $r \xrightarrow{a} r'$ implies $q+r \xrightarrow{a} q'+r'$, and (iii) 0 is a non-final sink state. Note that these automata are the usual weighted automata with weights in the field \mathbb{Z}_2 , see e.g. [8].
- (d) A deterministic \mathbf{JSL} -automaton is a da with a join-semilattice structure on the states Q such that (i) the final states $F \subseteq Q$ satisfy $q \vee q' \in F$ iff $q \in F$ or $q' \in F$, (ii) $q \xrightarrow{a} q'$ and $r \xrightarrow{a} r'$ implies $q \vee r \xrightarrow{a} q' \vee r'$, and (iii) 0 is a non-final sink state.

- (e) A deterministic DL-automaton is a da with a distributive lattice structure on the states Q such that (i) the final states form a prime filter, (ii) $q \xrightarrow{a} q'$ and $r \xrightarrow{a} r'$ implies $q \vee r \xrightarrow{a} q' \vee r'$ and $q \wedge r \xrightarrow{a} q' \wedge r'$, and (iii) 0 is a non-final sink state and 1 is a final sink state.

Remark 2.7. 1. The final T_Σ -coalgebra in **Set** is $\nu T_\Sigma = \mathcal{P}\Sigma^*$, the automaton of all languages over Σ . Its transitions are $L \xrightarrow{a} a^{-1}L$ for $a \in \Sigma$, where $a^{-1}L$ is the left derivative

$$a^{-1}L = \{v \in \Sigma^* : av \in L\},$$

and the final states are precisely those languages containing ε . Importantly, as shown by Barr [6], νT_Σ arises as the limit of the terminal sequence of T_Σ

$$1 \leftarrow T_\Sigma 1 \leftarrow T_\Sigma T_\Sigma 1 \leftarrow \dots$$

Since for any variety \mathcal{V} the forgetful functor from \mathcal{V} to **Set** creates limits and since T_Σ on \mathcal{V} lifts T_Σ on **Set**, the final T_Σ -coalgebra νT_Σ in \mathcal{V} exists and is a lifting of the one in **Set**. Hence νT_Σ has the underlying set $\mathcal{P}\Sigma^*$ equipped with a canonical \mathcal{V} -operations and the transitions and final states are as above.

2. For finitary endofunctors T , Milius [22] introduced the concept of a locally finitely presentable coalgebra: it is a filtered colimit of coalgebras carried by finitely presentable objects. In the present context the finitely presentable objects are precisely the finite algebras in \mathcal{V} , so we speak about *locally finite coalgebras*. A T_Σ -coalgebra is locally finite iff from each state only finitely many states are reachable by transitions. Indeed, for $\mathcal{V} = \mathbf{Set}$ this was shown in [22], which implies the claim for general locally finite varieties \mathcal{V} since filtered colimits of T_Σ -coalgebras are constructed in their underlying category \mathcal{V} , and hence in **Set**.
3. For every locally finitely presentable category \mathcal{V} and every finitary endofunctor $T : \mathcal{V} \rightarrow \mathcal{V}$ one can construct the *rational fixpoint* ϱT , i.e., the filtered colimit of all T -coalgebras carried by a finitely presentable object, see [5]. At this level of generality it is known that the rational fixpoint is the final object in the category of locally finitely presentable coalgebras.
4. For the endofunctor $T_\Sigma = \{0, 1\} \times \text{Id}^\Sigma$ on **Set** the rational fixpoint ϱT_Σ is the automaton of all regular languages over Σ , with transitions and final states as in νT_Σ . More generally, if \mathcal{V} is a locally finite variety, ϱT_Σ is a lifting of this coalgebra:

Theorem 2.8 (see [3], Theorem 2.10). *For every locally finite variety \mathcal{V} the rational fixpoint ϱT_Σ is carried by the set of regular languages over Σ equipped with canonical \mathcal{V} -operations. The transitions are defined by left derivatives, and the final states are those languages containing the empty word.*

Example 2.9. (a) In \mathbf{Set}_* the carrier of the rational fixpoint ϱT_Σ has the chosen point \emptyset . All transition maps $L \mapsto a^{-1}L$ preserve \emptyset .

- (b) In BA, ϱT_Σ has the usual set-theoretic boolean algebra structure. The principal filter $\uparrow\{\varepsilon\}$ is an ultrafilter and the transition maps $L \mapsto a^{-1}L$ are boolean homomorphisms.
- (c) In $\mathbf{Vect}(\mathbb{Z}_2)$ the vector space structure on ϱT_Σ is given by symmetric difference and \emptyset is the zero vector. The transition maps $L \mapsto a^{-1}L$ are linear.
- (d) In JSL the semilattice structure on ϱT_Σ is union and \emptyset . The final states form the principal upper set $\uparrow\{\varepsilon\}$ and the transitions maps are semilattice homomorphisms.
- (e) In DL we have the usual set-theoretic lattice structure on ϱT_Σ . The final states form a prime filter and the transition maps are lattice homomorphisms.

Notation 2.10. Let (Q, γ) be a locally finite T_Σ -coalgebra. The unique coalgebra homomorphism into ϱT_Σ , see Remark 2.7, is denoted by

$$\mathcal{L}_\gamma : Q \rightarrow \varrho T_\Sigma.$$

It sends $q \in Q$ to the regular language $\mathcal{L}_\gamma(q) \subseteq \Sigma^*$ accepted by the state q .

Definition 2.11. Let $V \in \mathcal{V}$ denote the free algebra of \mathcal{V} on one generator g . Then a *pointed \mathcal{V} -automaton* (Q, γ, q_0) is a \mathcal{V} -automaton (Q, γ) with a morphism $q_0 : V \rightarrow Q$. The latter may be viewed as the initial state $q_0(g) \in Q$. The *language accepted by* (Q, γ, q_0) is $\mathcal{L}_\gamma(q_0)$. We say that a pointed \mathcal{V} -automaton (Q, γ, q_0) is

1. *reachable* if it is generated by q_0 , i.e., no proper subcoalgebra contains q_0 ;
2. *simple* if it has no proper quotients, i.e., every quotient coalgebra $e : (Q, \gamma) \rightarrow (Q', \gamma')$ is an isomorphism;
3. *minimal* if it is reachable and simple.

Example 2.12. The minimal pointed dfa in \mathbf{Set} for a regular language L has been described by Brzozowski [12]. Its states Q_L are the left derivatives of L , i.e.,

$$Q_L = \{w^{-1}L : w \in \Sigma^*\} \quad \text{where } w^{-1}L = \{v \in \Sigma^* : wv \in L\},$$

the transitions are given by $K \xrightarrow{a} a^{-1}K$ for $K \in Q_L$ and $a \in \Sigma$, the initial state is L , and a state is final iff it contains the empty word.

The minimal pointed dfa for the language $L_n = (a+b)^*a(a+b)^n$, see Example 1.1, has 2^{n+1} states since L_n has precisely 2^{n+1} left derivatives. Indeed, for all words $w = w_0 \cdots w_n$ of length $n+1$ the left derivatives are pairwise distinct: $w^{-1}L_n$ consists of the words of L_n and all words of length $i = 0, \dots, n$ such that $w_{n-i} = a$:

$$w^{-1}L_n = L_n + \sum_{w_{n-i}=a} (a+b)^i.$$

And there are no other left derivatives, since for longer words vw (where w has length $n + 1$) we have $(vw)^{-1}L_n = w^{-1}L_n$, and for shorter words u choose j such that $w = b^j u$ has length $n + 1$, and get $u^{-1}L_n = w^{-1}L_n$.

Remark 2.13. The variety \mathcal{V} has the factorization system (strong epi, mono) = (surjective, injective). Since T_Σ preserves monomorphisms, this factorization system lifts to $\text{Coalg}(T_\Sigma)$, that is, every coalgebra homomorphism can be decomposed into a surjective homomorphism followed by an injective one.

Lemma 2.14. *A finite pointed \mathcal{V} -automaton (Q, γ, q_0) is reachable iff the algebra Q is generated by the set of all states reachable from q_0 by transitions. It is simple iff \mathcal{L}_γ is injective, that is, distinct states accept distinct languages.*

Proof. Let $M \subseteq Q$ be the set of all states reachable from q_0 by transitions. Then all transition maps γ_a preserve M , hence, being homomorphisms of \mathcal{V} , they preserve the subalgebra $Q' \subseteq Q$ generated by M . Consequently Q' is a subcoalgebra of the T_Σ -coalgebra Q . If Q is reachable, we conclude $Q' = Q$. Conversely, if $Q' = Q$, then Q is reachable: every subcoalgebra containing q_0 also contains M , hence it contains Q' .

Now suppose (Q, γ) is simple. Factorize the unique coalgebra homomorphism $\mathcal{L}_\gamma : (Q, \gamma) \rightarrow \varrho T_\Sigma$ into a surjective homomorphism f followed by an injective one, see Remark 2.13. Then f is bijective by simplicity, so \mathcal{L}_γ is injective. Conversely, suppose \mathcal{L}_γ is injective and $f : (Q, \gamma) \twoheadrightarrow (Q', \gamma')$ is surjective. Then $\mathcal{L}_\gamma = \mathcal{L}_{\gamma'} \circ f$ by finality of ϱT_Σ , so f is injective and hence an isomorphism. \square

Brzozowski's construction of the minimal pointed dfa for a regular language (see Example 2.12) generalizes to deterministic \mathcal{V} -automata as follows. Recall that the rational fixpoint ϱT_Σ is carried by the set of regular languages over Σ .

Construction 2.15. For every regular language $L \subseteq \Sigma^*$ let A_L be the pointed \mathcal{V} -automaton (Q_L, γ, L) where:

1. Q_L is the subalgebra of ϱT_Σ generated by all left derivatives $w^{-1}L$ ($w \in \Sigma^*$).
2. The transitions are $K \xrightarrow{a} a^{-1}K$ for $a \in \Sigma$ and $K \in Q_L$.
3. $K \in Q_L$ is final iff $\varepsilon \in K$.

Lemma 2.16. *For every regular language $L \subseteq \Sigma^*$, A_L is a well-defined finite pointed \mathcal{V} -automaton accepting L .*

Proof. L is regular, so it has only finitely many distinct left derivatives $w^{-1}L$. Hence Q_L is a finite algebra because \mathcal{V} is a locally finite variety. Next we show that $\gamma_a : Q_L \rightarrow Q_L$ and $\gamma_\varepsilon : Q_L \rightarrow \mathbb{2}$ as specified in points 2. and 3. are well-defined \mathcal{V} -morphisms. Recall the final locally finite T_Σ -coalgebra $(\varrho T_\Sigma, \gamma_\varrho)$. Then

$$\gamma_\varepsilon = Q_L \hookrightarrow \varrho T_\Sigma \xrightarrow{(\gamma_\varrho)_\varepsilon} \mathbb{2}$$

is a \mathcal{V} -morphism since ϱT_Σ is a lifting of the automaton of all regular languages, see Lemma 2.8. Furthermore $(\gamma_\varrho)_a : \varrho T_\Sigma \rightarrow \varrho T_\Sigma$ is defined $(\gamma_\varrho)_a(K) = a^{-1}K$

i.e. the left derivative $a^{-1}(-)$ preserves the algebraic operations. Thus Q_L is closed under left derivatives, so γ_a is a well-defined algebra morphism.

To see that A_L accepts L , use the fact that the state L of ϱT_Σ accepts the language L , so the same holds for L in A_L . \square

Example 2.17. (a) In Set_* , we have $Q_L = \{\emptyset\} \cup \{w^{-1}L : w \in \Sigma^*\}$ with the chosen state $q_l = \emptyset$.

(b) In BA, Q_L is the closure of $\{w^{-1}L : w \in \Sigma^*\}$ under finite union and complement.

(c) In $\text{Vect}(\mathbb{Z}_2)$, Q_L is the closure of $\{w^{-1}L : w \in \Sigma^*\}$ under symmetric difference.

(d) In JSL, Q_L is the closure of $\{w^{-1}L : w \in \Sigma^*\}$ under finite union.

(e) In DL, Q_L is the closure of $\{w^{-1}L : w \in \Sigma^*\}$ under finite union and finite intersection.

In (b)–(e) the constants \emptyset and Σ^* appear as empty union and intersection, respectively, and closure under intersection in (b) follows from closure under union and complement.

Construction 2.18 (see [2]). Here we give a two-step minimization of any finite pointed \mathcal{V} -automaton (Q, γ, q_0) :

1. Construct the reachable subcoalgebra $(R, \delta, q_0) \hookrightarrow (Q, \gamma, q_0)$ generated by q_0 . That is, R is the \mathcal{V} -subalgebra generated by the set of all states reachable from q_0 .
2. Factorize the unique T_Σ -coalgebra homomorphism $\mathcal{L}_\delta : (R, \delta) \rightarrow (\varrho T_\Sigma, \gamma_\varrho)$ as in Remark 2.13:

$$(R, \delta) \xrightarrow{s} (R', \delta') \xrightarrow{m} (\varrho T_\Sigma, \gamma_\varrho).$$

Then $(R', \delta', s(q_0))$ is minimal.

Theorem 2.19. *Let $L \subseteq \Sigma^*$ be a regular language. Then A_L is (up to isomorphism) the unique minimal pointed \mathcal{V} -automaton accepting L . It arises from any pointed finite \mathcal{V} -automaton (Q, γ, q_0) accepting L by Construction 2.18.*

Proof. A_L is reachable because every state is a \mathcal{V} -algebraic combination of those states $w^{-1}L$ reachable from L by transitions. It is simple because every state K of A_L accepts the language K , hence different states accept different languages. It follows that A_L is minimal.

Now let (Q, γ, q_0) be any pointed \mathcal{V} -automaton accepting L and (R, δ, q_0) its reachable subautomaton, so every $q' \in R$ arises as a \mathcal{V} -algebraic combination of states reachable from q_0 by transitions. Since the map $\mathcal{L}_\delta : R \rightarrow \varrho T_\Sigma$ is a coalgebra morphism and since $a^{-1}(-)$ taking left-derivatives preserves \mathcal{V} -operations, the languages of states reachable from q_0 are precisely the left derivatives of L . Consequently \mathcal{L}_δ has a codomain restriction $s : (R, \delta) \rightarrow A_L$ which is a

coalgebra morphism since \mathcal{L}_δ is one, and A_L is a subcoalgebra of ϱT_Σ . Moreover, s is surjective since every left derivative $w^{-1}L$ is accepted by a state of R (namely the state reached from q_0 by w). Thus $(R, \delta) \xrightarrow{s} A_L \xrightarrow{m} (\varrho T_\Sigma, \gamma_\varrho)$ is up to isomorphism the factorization of step 2 in Construction 2.18. \square

3. From Deterministic to Nondeterministic Automata

We now know that every regular language L has *many* canonical deterministic acceptors: one for each locally finite variety \mathcal{V} containing a two-element algebra $\mathbf{2}$. However this canonical acceptor A_L is generally larger than the minimal dfa in \mathbf{Set} because one has to close the set of left derivatives under the \mathcal{V} -algebraic operations. In this section we will show how these larger *deterministic* machines induce smaller *nondeterministic* ones. Let us outline our approach:

1. For each of our varieties \mathcal{V} of interest, we describe an equivalence G between the category \mathcal{V}_f of all finite algebras in \mathcal{V} and another category $\overline{\mathcal{V}}$ whose objects are “small” representations of their counterparts in \mathcal{V}_f , and whose morphisms are relations, not functions (see Lemmas 3.4, 3.8 and 3.10).
2. From G we derive an equivalence between pointed deterministic finite \mathcal{V} -automata and pointed coalgebras in $\overline{\mathcal{V}}$ which are *nondeterministic* finite automata (see Lemma 3.17).
3. Applying this equivalence to the minimal deterministic \mathcal{V} -automaton A_L gives a canonical nondeterministic acceptor for L . This is illustrated in Section 3.3.

3.1. The Equivalence between \mathcal{V}_f and $\overline{\mathcal{V}}$

For each of our varieties \mathcal{V} of interest there is a well-known description of the dual category of \mathcal{V}_f : we have Stone duality ($\mathbf{BA}_f \cong \mathbf{Set}_f^{op}$), Priestley duality ($\mathbf{DL}_f \cong \mathbf{Poset}_f^{op}$), where \mathbf{Poset}_f is the category of finite posets and monotone functions, and the self-dualities $\mathbf{JSL}_f \cong \mathbf{JSL}_f^{op}$ and $\mathbf{Vect}_f(\mathbb{Z}_2) \cong \mathbf{Vect}_f(\mathbb{Z}_2)^{op}$. (The self-duality of \mathbf{JSL}_f assigns to every finite semilattice $(Q, \vee, 0)$ its opposite semilattice $(Q, \wedge, 1)$, see [18, 6.3.6]). We now describe each of these dually equivalent categories as a category $\overline{\mathcal{V}}$ of finite structured sets and relations. The idea is to represent the finite algebras in \mathcal{V} in terms of a minimal set of generators.

- Example 3.1.** (a) For any $Q \in \mathbf{Set}_*$ the subset $Q \setminus \{q_t\}$ generates Q ; that means that we can always drop one element.
- (b) Any finite boolean algebra $Q \in \mathbf{BA}_f$ is generated by the set $\mathbf{At}(Q) \subseteq Q$ of its atoms (i.e., minimal elements). Indeed, every element of Q is the join of all atoms below it.

- (c) A finite dimensional vector space $Q \in \mathbf{Vect}_f(\mathbb{Z}_2)$ is generated by any basis $B \subseteq Q$: every element of Q is the sum of a subset of B .
- (d) Any finite join-semilattice Q is generated by the set $J(Q) \subseteq Q$ of its join-irreducible elements. Recall that an element $q \in Q$ is *join-irreducible* if (i) $q \neq 0$ and (ii) $q = r \vee r'$ implies $q = r$ or $q = r'$. Every element of Q is the join of all join-irreducibles below it.
- (e) Analogously, any finite distributive lattice $Q \in \mathbf{DL}_f$ is generated by its join-irreducibles $J(Q)$.

Notation 3.2. (a) Let $\overline{\mathbf{Set}_*}$ be the category \mathbf{Par}_f of finite sets and partial functions.

(b) $\overline{\mathbf{BA}}$ is the category of finite sets and all relations whose converse is a function. Composition is the usual composition of relations.

(c) $\overline{\mathbf{Vect}(\mathbb{Z}_2)}$ is the category of finite sets and *all* relations. However here the composition of two relations $R_1 \subseteq X \times Y$ and $R_2 \subseteq Y \times Z$ is defined by

$$R_2 \bullet R_1 := \{(x, z) : \text{the number of } y \in Y \text{ with } (x, y) \in R_1 \text{ and } (y, z) \in R_2 \text{ is odd}\}.$$

If one identifies a relation between finite sets X and Y with the corresponding binary $X \times Y$ -matrix, the above composition \bullet amounts to matrix multiplication over the field \mathbb{Z}_2 .

Notation 3.3. Given a basis GQ of a \mathbb{Z}_2 -vector space Q and a basis vector $z \in GQ$, we write $\pi_z : Q \rightarrow \{0, 1\}$ for the projection onto the z -coordinate. Hence π_z takes a vector to 1 iff its unique representation as a sum of basis vectors contains z as a summand.

Lemma 3.4. *The following functors G are equivalences of categories:*

- (1) $G : \mathbf{Set}_{*f} \rightarrow \mathbf{Par}_f$ where $G(Q, q_1) = Q \setminus \{q_1\}$ and, for any $f : (Q, q_1) \rightarrow (Q', q'_1)$ in \mathbf{Set}_{*f} ,

$$Gf(z) = \begin{cases} f(z) & \text{if } f(z) \neq q'_1, \\ \text{undefined} & \text{otherwise.} \end{cases}$$

- (2) $G : \mathbf{BA}_f \rightarrow \overline{\mathbf{BA}}$ where $GQ = \mathbf{At}(Q)$ is the set of all atoms of Q and, for any $f : Q \rightarrow Q'$ in \mathbf{BA}_f ,

$$Gf[z] = \{z' \in \mathbf{At}(Q') : z' \leq f(z)\}.$$

- (3) $G : \mathbf{Vect}_f(\mathbb{Z}_2) \rightarrow \overline{\mathbf{Vect}(\mathbb{Z}_2)}$ where GQ is a chosen basis of Q and, for any $f : Q \rightarrow Q'$ in $\mathbf{Vect}_f(\mathbb{Z}_2)$,

$$Gf[z] = \{z' \in GQ' : \pi_{z'} \circ f(z) = 1\}.$$

Proof. The equivalence (1) between pointed sets and partial functions is well-known. For the equivalence (2) observe that (i) \mathbf{BA}_f is dually equivalent to \mathbf{Set}_f , (ii) \mathbf{Set}_f is a non-full subcategory of \mathbf{Rel}_f , the category of finite sets and relations, and (iii) the latter category is self-dual by taking converse relations. Then Gf arises by following these three steps. Finally, (3) follows by the well-known equivalence of the category of finite-dimensional vector spaces, and the category whose objects are the natural numbers and whose morphisms are matrices over \mathbb{Z}_2 . \square

Finite join-semilattices are equivalently represented as closure spaces:

Definition 3.5. A *closure space* is a set X equipped with a *closure operator* that associates to each subset $M \subseteq X$ another subset $\overline{M} \subseteq X$ (the *closure* of M) and is monotone, extensive and idempotent, i.e.,

$$(i) M \subseteq M' \text{ implies } \overline{M} \subseteq \overline{M'}, \quad (ii) M \subseteq \overline{M}, \quad (iii) \overline{\overline{M}} = \overline{M}.$$

A closure space X is *strict* if $\overline{\emptyset} = \emptyset$, and *topological* if moreover $\overline{A \cup B} = \overline{A} \cup \overline{B}$ for all $A, B \subseteq X$. It is a T_0 *space* if $\{x\} \in \overline{\{y\}}$ and $y \in \overline{\{x\}}$ implies $x = y$. A subset $M \subseteq X$ is *closed* if $\overline{M} = M$ and *open* if its complement is closed.

Finite posets are well-known to be equivalent to finite T_0 topological spaces. These amount to finite topological closure spaces satisfying T_0 . For finite join-semilattices we instead use *finite strict closure spaces*, i.e., we do not require T_0 or preservation of unions.

Example 3.6. Each finite join-semilattice Q has an associated finite strict closure space GQ where $GQ = J(Q)$ is the set of join-irreducibles of Q , and the closure of a set $M \subseteq GQ$ is given by all join-irreducibles lying under its join:

$$\overline{M} = \{j \in GQ : j \leq \bigvee M\}.$$

For example, the closure space associated to the free join-semilattice $\mathcal{P}n$ is $n = \{0, \dots, n-1\}$ equipped with the identity closure $\overline{M} = M$. Here we identify the join-irreducibles of $\mathcal{P}n$ with n .

Definition 3.7. The category $\overline{\mathbf{JSL}}$ has as objects all finite strict closure spaces and as morphisms from X to Y all *continuous* relations $R \subseteq X \times Y$, that is,

1. $R[x] \subseteq Y$ is closed for all $x \in X$, and
2. $R[\overline{M}] \subseteq \overline{R[M]}$ for all $M \subseteq X$.

The composition $R_2 \bullet R_1$ of two continuous relations $R_1 \subseteq X \times Y$ and $R_2 \subseteq Y \times Z$ is given by

$$R_2 \bullet R_1[x] := \overline{R_2 \circ R_1[x]},$$

where \circ denotes the usual composition of relations. The identity morphism on X is

$$I_X[x] = \overline{\{x\}}.$$

The following equivalence can be derived from a similar one due to Moshier (see Jipsen [17]). A full proof is given in the Appendix.

Lemma 3.8. *The categories JSL_f and $\overline{\text{JSL}}$ are equivalent. The equivalence functor $G : \text{JSL}_f \rightarrow \overline{\text{JSL}}$ maps any finite semilattice Q to the closure space GQ of Example 3.6, and any homomorphism $f : Q \rightarrow Q'$ to the continuous relation $Gf \subseteq GQ \times GQ'$ defined by*

$$Gf[x] = \{y \in GQ' : y \leq f(x)\}.$$

Notation 3.9. $\overline{\text{DL}}$ is the category whose objects are finite posets and whose morphisms from X to Y are those relations $R \subseteq X \times Y$ such that

1. $R[x]$ is down-closed for every $x \in X$,
2. $x \leq y$ in X implies $R[x] \subseteq R[y]$, and
3. R preserves intersections of down-closed subsets.

The identity morphism on X is the relation $\{(x, y) \in X \times X : y \leq x\}$ and composition is the usual relational composition.

Lemma 3.10. *The categories DL_f and $\overline{\text{DL}}$ are equivalent. The equivalence functor $G : \text{DL}_f \rightarrow \overline{\text{DL}}$ maps any finite distributive lattice Q to the poset $GQ = J(Q)$ (considered as a subposet of Q), and any morphism $f : Q \rightarrow Q'$ to the relation $Gf \subseteq GQ \times GQ'$ defined by*

$$Gf[x] = \{y \in GQ' : y \leq f(x)\}.$$

Proof. G is a restriction of the equivalence $\text{JSL}_f \cong \overline{\text{JSL}}$ described above. The closure spaces associated to distributive lattices are precisely the T_0 topological ones, so we can replace them by finite posets. This gives the first two conditions on morphisms, where closed means downwards closed. However semilattice morphisms between distributive lattices need not preserve meets. This is captured by the third condition. \square

3.2. From Determinism to Nondeterminism

The equivalences from the previous section allow us to view finite deterministic \mathcal{V} -automata as coalgebras in $\overline{\mathcal{V}}$. Let us restrict the endofunctor T_Σ of Definition 2.4 to finite algebras and use the same notation:

$$T_\Sigma = \mathbf{2} \times \text{Id}^\Sigma : \mathcal{V}_f \rightarrow \mathcal{V}_f.$$

Then for each of our five equivalences $G : \mathcal{V}_f \rightarrow \overline{\mathcal{V}}$ described above we have the corresponding functor

$$\overline{T}_\Sigma = \mathbf{1} \times \text{Id}^\Sigma : \overline{\mathcal{V}} \rightarrow \overline{\mathcal{V}},$$

where $\mathbf{1} = G\mathfrak{2} \in \overline{\mathcal{V}}$. In all examples $\mathbf{1}$ has carrier $\{1\}$:

\mathcal{V}	$\mathbf{1}$
Set_*	$\mathfrak{2} \setminus \{0\} = \{1\}$
BA	$\text{At}(\mathfrak{2}) = \{1\}$, the unique atom
$\text{Vect}(\mathbb{Z}_2)$	$\{1\}$, the unique basis of $\mathfrak{2} = \mathbb{Z}_2$
JSL	$J(\mathfrak{2}) = \{1\}$
DL	$J(\mathfrak{2}) = \{1\}$

In analogy to Remark 2.5 we denote \overline{T}_Σ -coalgebras $\delta : Z \rightarrow \overline{T}_\Sigma Z = \mathbf{1} \times Z^\Sigma$ as triples $(Z, \delta_a, \delta_\varepsilon)$ with $\delta_\varepsilon : Z \rightarrow \mathbf{1}$ and $\delta_a : Z \rightarrow Z$ for $a \in \Sigma$. Notice that these are *relations* rather than functions, so \overline{T}_Σ -coalgebras are *nondeterministic automata*. The relation δ_ε defines a set of final states: a state $z \in Z$ is final iff $\delta_\varepsilon[z] \neq \emptyset$.

Lemma 3.11. *The categories $\text{Coalg}(T_\Sigma)$ and $\text{Coalg}(\overline{T}_\Sigma)$ are equivalent. The equivalence functor $\mathbb{G} : \text{Coalg}(T_\Sigma) \rightarrow \text{Coalg}(\overline{T}_\Sigma)$ is defined by*

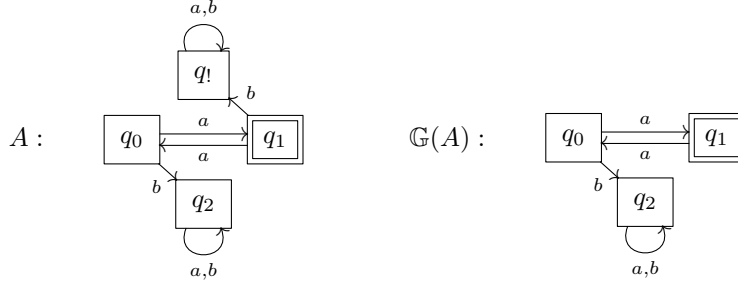
$$\mathbb{G}(Q, \gamma_a, \gamma_\varepsilon) = (GQ, G\gamma_a, G\gamma_\varepsilon) \quad \text{and} \quad \mathbb{G}f = Gf.$$

Proof. A T_Σ -coalgebra homomorphism $f : (Q, \gamma) \rightarrow (Q', \gamma')$ amounts to a \mathcal{V} -morphism $f : Q \rightarrow Q'$ such that $\gamma_\varepsilon = \gamma'_\varepsilon \circ f$ and $f \circ \gamma_a = \gamma'_a \circ f$ for each $a \in \Sigma$. Hence the equivalence $G : \mathcal{V}_f \rightarrow \overline{\mathcal{V}}$ allows us to define an equivalent category by simply applying G to every morphism $\gamma_\varepsilon, \gamma_a$ and f . By the universal property of products in $\overline{\mathcal{V}}$, it follows that $\mathbb{G} : \text{Coalg}(T_\Sigma) \rightarrow \text{Coalg}(\overline{T}_\Sigma)$ defines an equivalence. \square

Example 3.12. (a) If $\mathcal{V} = \text{Set}_*$ then a \overline{T}_Σ -coalgebra $\delta : Z \rightarrow \overline{T}_\Sigma Z$ consists of:

1. A finite set Z of states.
2. A partial function $\delta_\varepsilon : Z \rightarrow \{1\}$ defining a set final states.
3. A partial function $\delta_a : Z \rightarrow Z$ for each $a \in \Sigma$, defining the transitions.

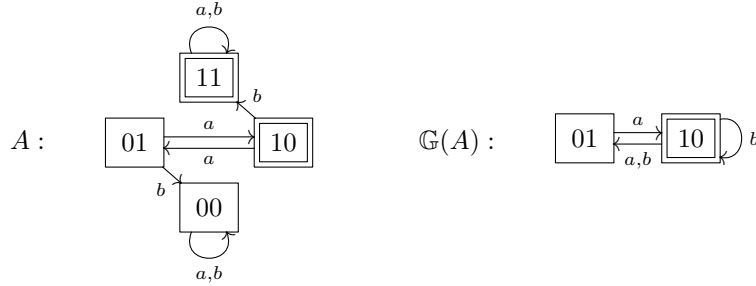
Hence \overline{T}_Σ -coalgebras are *partial dfas*. The equivalence \mathbb{G} assigns to each deterministic Set_* -automaton $A = ((Q, q_i), \gamma)$ the partial dfa $\mathbb{G}(A)$ with states $Q \setminus \{q_i\}$ that arises by removing the sink state q_i . That is, $q \xrightarrow{a} q'$ in $\mathbb{G}(A)$ iff $q' \neq q_i$ and $q \xrightarrow{a} q'$ in A , and the final states of $\mathbb{G}(A)$ are the ones of A . A concrete example:



(b) If $\mathcal{V} = \mathbf{BA}$ then a \overline{T}_Σ -coalgebra $\delta : Z \rightarrow \overline{T}_\Sigma Z$ consists of:

1. A finite set Z of states.
2. A converse-functional relation $\delta_\varepsilon \subseteq Z \times \{1\}$ defining a single final state, viz. the unique state z_0 with $\delta_\varepsilon[z_0] \neq \emptyset$.
3. Converse-functional transition relations $\delta_a \subseteq Z \times Z$ for $a \in \Sigma$.

Hence \overline{T}_Σ -coalgebras are *reverse-deterministic nfas*, i.e., reversing all transitions yields a dfa. The equivalence \mathbb{G} assigns to each deterministic BA-automaton $A = (Q, \gamma)$ an nfa $\mathbb{G}(A)$ whose states are the atoms of Q . Moreover, the final state of $\mathbb{G}(A)$ is the unique atom generating the ultrafilter of all final states of A , and there is a transition $z \xrightarrow{a} z'$ in $\mathbb{G}(A)$ iff $z' \leq \gamma_a(z)$. A concrete example with the boolean algebra $\mathbb{2} \times \mathbb{2}$ as a state space is below:



(c) If $\mathcal{V} = \mathbf{Vect}(\mathbb{Z}_2)$ then a \overline{T}_Σ -coalgebra $\delta : Z \rightarrow \overline{T}_\Sigma Z$ consists of:

1. A finite set Z of states.
2. A relation $\delta_\varepsilon \subseteq Z \times \{1\}$ defining a set of final states.
3. Transition relations $\delta_a \subseteq Z \times Z$ for each $a \in \Sigma$.

Hence \overline{T}_Σ -coalgebras are the classical nfas. The equivalence \mathbb{G} assigns to a deterministic $\mathbf{Vect}(\mathbb{Z}_2)$ -automaton $A = (Q, \gamma)$ an nfa $\mathbb{G}(A)$ whose states $Z \subseteq Q$ form a basis of Q . Furthermore, a state $z \in Z$ is final in $\mathbb{G}(A)$ iff it is final in A , and there is a transition $z \xrightarrow{a} z'$ in $\mathbb{G}(A)$ iff $\pi_{z'} \circ \gamma_a(z) = 1$ (see Notation 3.3), that is, iff one has a transition $z \xrightarrow{a} x$ in A such that z' is a summand of the representation of x as a sum of basis vectors. The concrete example in (b) also works for $\mathbf{Vect}(\mathbb{Z}_2)$: consider A as a $\mathbf{Vect}(\mathbb{Z}_2)$ -automaton with state space $\mathbb{Z}_2 \times \mathbb{Z}_2$ and basis $\{01, 10\}$.

(d) If $\mathcal{V} = \mathbf{JSL}$ then a \overline{T}_Σ -coalgebra $\delta : Z \rightarrow \overline{T}_\Sigma Z$ consists of:

1. A finite strict closure space Z .
2. A continuous relation $\delta_\varepsilon \subseteq Z \times \{1\}$, defining an *open* set of final states.
3. Continuous transition relations $\delta_a \subseteq Z \times Z$.

We call \overline{T}_Σ -coalgebras *nondeterministic closure automata*. The equivalence \mathbb{G} assigns to each deterministic JSL-automaton $A = (Q, \gamma)$ a nondeterministic closure automaton $\mathbb{G}(A)$ whose states are the join-irreducibles $J(Q)$ of Q . A state $z \in J(Q)$ is final in $\mathbb{G}(A)$ iff it is final in A , and $z \xrightarrow{a} z'$ in $\mathbb{G}(A)$ iff $z' \leq \gamma_a(z)$. Note that one can view every nfa as a nondeterministic closure automaton by endowing the set of states with the identity closure. The above example in (b) is also one for JSL.

(e) If $\mathcal{V} = \text{DL}$ then a \overline{T}_Σ -coalgebra $\delta : Z \rightarrow \overline{T}_\Sigma Z$ consists of:

1. A finite poset Z of states.
2. A non-empty relation $\delta_\varepsilon \subseteq Z \times \{1\}$, defining a filter (i.e., a down-directed upper set) of final states.
3. Transition relations $\delta_a \subseteq Z \times Z$ such that:
 - (i) $\delta_a[z]$ is down-closed for each $z \in Z$.
 - (ii) $z \leq z'$ implies $\delta_a[z] \subseteq \delta_a[z']$.
 - (iii) $\delta_a[\bigcap_I Z_i] = \bigcap_I \delta_a[Z_i]$ for all sets I and all down-closed subsets $Z_i \subseteq Z$.

Note that reverse-deterministic nfes are the special case where Z is discrete. An important non-discrete example is the *universal automaton* [20], see Example 3.27.

The equivalence \mathbb{G} assigns to each deterministic DL-automaton $A = (Q, \gamma)$ the \overline{T}_Σ -coalgebra $\mathbb{G}(A)$ whose states are the join-irreducibles $J(Q)$ of Q , ordered as in Q . A state $z \in J(Q)$ is final in $\mathbb{G}(A)$ iff it is final in A , and $z \xrightarrow{a} z'$ in $\mathbb{G}(A)$ iff $z' \leq \gamma_a(z)$. The concrete example in (b) also works for DL.

Remark 3.13. A morphism $f : (Z, \delta) \rightarrow (Z', \delta')$ of \overline{T}_Σ -coalgebras is, by definition, a morphism $f : Z \rightarrow Z'$ of $\overline{\mathcal{V}}$ satisfying $\overline{T}_\Sigma f \circ \delta = \delta' \circ f$, or equivalently,

$$\delta_\varepsilon = \delta'_\varepsilon \circ f \quad \text{and} \quad \delta'_a \circ f = f \circ \delta_a \quad (a \in \Sigma).$$

For $\mathcal{V} = \text{Set}_*$, BA and DL, these morphisms are relations such that (i) transitions are preserved and reflected, and (ii) a state $z \in Z$ is final iff some $z' \in f[z]$ is final. The cases $\mathcal{V} = \text{JSL}$ and $\text{Vect}(\mathbb{Z}_2)$ are different because the composition in $\overline{\mathcal{V}}$ is not the usual composition of relations.

3.3. Canonical Nondeterministic Automata

So far we have seen equivalences between deterministic and nondeterministic automata without initial states. Next, for each of our five running examples $\mathcal{V} = \text{Set}_*$, BA, $\text{Vect}(\mathbb{Z}_2)$, JSL, DL we will extend $\mathbb{G} : \text{Coalg}(T_\Sigma) \rightarrow \text{Coalg}(\overline{T}_\Sigma)$ to an equivalence of pointed coalgebras. Recall from Definition 2.11 that a pointed T_Σ -coalgebra (Q, γ, q_0) comes equipped with a morphism $q_0 : V \rightarrow Q$, where V is the free algebra on one generator in \mathcal{V} .

Notation 3.14. $\text{Coalg}_*(T_\Sigma)$ is the category of pointed T_Σ -coalgebras and point-preserving coalgebra homomorphisms $f : (Q, \gamma, q_0) \rightarrow (Q', \gamma', q'_0)$, i.e., coalgebra homomorphisms $f : (Q, \gamma) \rightarrow (Q', \gamma')$ with $f \circ q_0 = q'_0$.

Using the equivalence $G : \mathcal{V}_f \rightarrow \overline{\mathcal{V}}$, a pointed \overline{T}_Σ -coalgebra is a \overline{T}_Σ -coalgebra (Z, δ) equipped with a morphism $i : GV \rightarrow Z$. And pointed \overline{T}_Σ -coalgebra homomorphisms are those \overline{T}_Σ -coalgebra homomorphisms f from (Z, δ) to (Z', δ') such that $f \circ i = i'$. Just as a morphism $q_0 : V \rightarrow Q$ corresponds to an initial state, it turns out that a morphism $i : GV \rightarrow Z$ corresponds to a *set* of initial states, as one would expect for nfas.

Example 3.15. For each of our running examples \mathcal{V} we describe the possible sets of initial states $I \subseteq Z$ in a \overline{T}_Σ -coalgebra (Z, δ) .

- (a) If $\mathcal{V} = \text{Set}_*$ then $V = \{q!, g\}$ and $GV = \{g\}$. Partial functions $i : \{g\} \rightarrow Z$ are determined by their image $I = i[g]$. Hence I is either empty or any singleton subset.
- (b) If $\mathcal{V} = \text{BA}$ then $V = \{\perp, g, \neg g, \top\}$ and $GV = \{g, \neg g\}$. Given $i \subseteq \{g, \neg g\} \times Z$ that it a converse of a function, the sets $i[g]$ and $i[\neg g]$ form a partition of Z , so i is determined by $I = i[g]$. Hence I is any subset of Z .
- (c) If $\mathcal{V} = \text{Vect}(\mathbb{Z}_2)$ then $V = \{0, g\}$ and $GV = \{g\}$, so the arbitrary relation $i \subseteq \{g\} \times Z$ is determined by its codomain $I = i[g]$. Then I is any subset of Z .
- (d) If $\mathcal{V} = \text{JSL}$ then $V = \{0, g\}$ and $GV = \{g\}$ with closure operator $\text{id}_{\mathcal{P}\{g\}}$. The relation $i \subseteq \{g\} \times Z$ is determined by $I = i[g]$. By continuity $I \subseteq Z$ is a closed subset.
- (e) If $\mathcal{V} = \text{DL}$ then $V = \{0, g, 1\}$ is a 3-chain and $GV = \{g, 1\}$ a 2-chain. Given $i \subseteq \{g, 1\} \times Z$ then $i[g] \subseteq i[1]$ and since i preserves the empty intersection of down-closed subsets, we have $i[1] = Z$. Thus i is determined by $I = i[g]$, and I is a down-closed subset of Z .

By reformulating the condition $f \circ q_0 = q'_0$ of point-preserving coalgebra morphisms in terms of I , we can define the category of pointed \overline{T}_Σ -coalgebras.

Notation 3.16. For each of our five running examples, $\text{Coalg}_*(\overline{T}_\Sigma)$ is the category of all triples (Z, δ, I) where (Z, δ) is a \overline{T}_Σ -coalgebra and $I \subseteq Z$ is restricted as in Example 3.15. The morphisms $f : (Z, \delta, I) \rightarrow (Z', \delta', I')$ are the \overline{T}_Σ -coalgebra homomorphisms $f : (Z, \delta) \rightarrow (Z', \delta')$ such that:

- 1. If $\mathcal{V} = \text{Set}_*$, BA or DL then $I' = f[I]$.
- 2. If $\mathcal{V} = \text{JSL}$ then $I' = \overline{f[I]}$.
- 3. If $\mathcal{V} = \text{Vect}(\mathbb{Z}_2)$ then $I' = \{z' \in Z' : \text{the number of } z \in I \text{ with } (z, z') \in f \text{ is odd}\}$.

Lemma 3.17. *There is an equivalence of pointed coalgebras $\mathbb{G}_* : \text{Coalg}_*(T_\Sigma) \rightarrow \text{Coalg}_*(\overline{T}_\Sigma)$ defined by*

$$\mathbb{G}_*(Q, \gamma, q_0) = (\mathbb{G}(Q, \gamma), I) \quad \text{and} \quad \mathbb{G}_*f = Gf$$

where I corresponds to $Gq_0 : GV \rightarrow GQ = Z$ (cf. Example 3.15).

Proof. \mathbb{G} is an equivalence of coalgebras by Lemma 3.11, and \mathbb{G}_* clearly extends this equivalence to pointed coalgebras: the conditions on I from Example 3.15 reflect by definition precisely the corresponding conditions on q_0 . \square

Definition 3.18. A finite (pointed) coalgebra for \overline{T}_Σ in $\overline{\mathcal{V}}$ is called a *finite (pointed) nondeterministic $\overline{\mathcal{V}}$ -automaton*, shortly $\overline{\mathcal{V}}$ -nfa. Given a regular language $L \subseteq \Sigma^*$, the $\overline{\mathcal{V}}$ -nfa $\mathbb{G}_*(A_L)$ (see Construction 2.15) is called the *canonical $\overline{\mathcal{V}}$ -nfa* for L .

We will now spell out the equivalence \mathbb{G}_* for each of our five varieties \mathcal{V} . For the rest of this section fix a finite pointed \mathcal{V} -automaton $A = (Q, \gamma, q_0)$ and a regular language $L \subseteq \Sigma^*$. We give an explicit description of the $\overline{\mathcal{V}}$ -nfa \mathbb{G}_*A and, in particular, of the canonical $\overline{\mathcal{V}}$ -nfa $\mathbb{G}_*(A_L)$.

Example 3.19 (The Minimal Partial Dfa). If $\mathcal{V} = \text{Set}_*$ then \mathbb{G}_*A is the partial dfa $(Q \setminus \{q_\dagger\}, \delta, I)$ that arises from A by deleting the non-final sink state q_\dagger along with all in- and outgoing transitions. Hence the initial states are $I = \{q_0\}$ if $q_0 \neq q_\dagger$ and $I = \emptyset$ if $q_0 = q_\dagger$. Clearly \mathbb{G}_*A (viewed as an nfa) accepts the same language as A .

The canonical $\overline{\text{Set}_*}$ -nfa $\mathbb{G}_*(A_L)$ is the *minimal partial dfa* of L . It has states

$$\mathcal{Q}_L = \{w^{-1}L : w \in \Sigma^*\} \setminus \{\emptyset\},$$

transitions $K \xrightarrow{a} a^{-1}K$ whenever $a^{-1}K \neq \emptyset$, and a state is final iff it contains ε . The initial states are $\{L\}$ if $L \neq \emptyset$ and \emptyset otherwise. Hence the minimal partial dfa for L is obtained from the minimal dfa by deleting its non-final sink state, if it exists.

For example, the minimal partial dfa for the language $L_n = (a+b)^*a(a+b)^n$, see Example 2.12, has $2^{n+1} - 1$ states.

Example 3.20 (The Átomaton). If $\mathcal{V} = \text{BA}$ then \mathbb{G}_*A is the nfa $(\text{At}(Q), \delta, I)$ with initial states $I = \{q \in \text{At}(Q) : q \leq q_0\}$. The canonical $\overline{\text{BA}}$ -nfa $\mathbb{G}_*(A_L)$ is called the *átomaton* of L , see [11]. Its states are

$$\mathcal{Q}_L = \text{At}(Q'_L),$$

the atoms of the finite boolean subalgebra Q'_L of $\mathcal{P}\Sigma^*$ generated by the left derivatives of L . An atom K is an initial state iff $K \subseteq L$, the final states are the atoms containing ε , and a transition $K \xrightarrow{a} K'$ exists iff $K' \subseteq a^{-1}K$. Explicitly constructing \mathcal{Q}_L can be difficult. Fortunately, a simpler method is known [11]:

1. Construct the minimal dfa for the reversed language $L^{rev} = \{w^{rev} : w \in L\}$.

2. Construct its reversed nfa, i.e., flip initial/final states and reverse all transitions.

The átomaton is isomorphic to the resulting nfa as we now explain by means of coalgebras. Let $T'_\Sigma = 2 \times \text{Id}^\Sigma : \text{Set}_f \rightarrow \text{Set}_f$. Then the usual reversal of finite pointed deterministic automata defines a dual equivalence

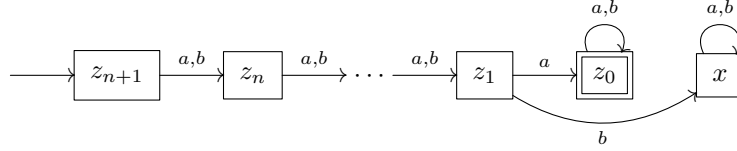
$$H : (\text{Coalg}_*(T'_\Sigma))^{op} \rightarrow \text{Coalg}_*(\overline{T}_\Sigma).$$

To every morphism $f : Z \rightarrow Z'$ it assigns

$$Hf^{op} = \{(z', z) : z \in f^{-1}(\{z'\})\} \subseteq Z' \times Z.$$

Since reachability (no proper subobjects) and simplicity (no proper quotients) are mutually *dual* concepts (see Definition 2.11), a T'_Σ -coalgebra is minimal iff its image under H is minimal, implying the above description.

In our concrete example $L_n = (a+b)^*a(a+b)^n$ it is easy to see that $L_n^{rev} = (a+b)^na(a+b)^*$ has the minimal dfa



Its reversal is the átomaton in Example 1.1.

Example 3.21 (The Minimal Xor Automaton). If $\mathcal{V} = \text{Vect}(\mathbb{Z}_2)$ then \mathbb{G}_*A is the nfa (Z, δ, I) where $Z \subseteq Q$ is a basis and $I = \{z \in Z : \pi_z(q_0) = 1\}$, see Notation 3.3. It accepts the same language as A , however by \mathbb{Z}_2 -weighted non-deterministic acceptance: a word $w \in \Sigma^*$ is accepted iff its number of accepting paths is odd.

The canonical $\overline{\text{Vect}(\mathbb{Z}_2)}$ -nfa $\mathbb{G}_*(A_L)$ is called the *minimal xor automaton* of L , see [25]. Note that its construction depends on the choice of a basis. However, the minimal xor automaton is uniquely determined up to isomorphism in the category of pointed \overline{T}_Σ -coalgebras. We provide a new way to construct it:

1. Construct the átomaton (Z, R_a, F, I) of L and determine the collection $C \subseteq \mathcal{P}Z$ of all subsets of Z that are reachable from I .
2. Form the closure \hat{C} of C under symmetric difference, and find a basis of \hat{C} , i.e., any minimal $\mathcal{Q} \subseteq \mathcal{P}Z$ whose closure under symmetric difference is \hat{C} .
3. Build the nfa $(\mathcal{Q}, R'_a, \mathcal{Q} \cap F, I')$ where $R'_a(y, y')$ iff $\pi_{y'}(R_a[y]) = 1$ and $I' = \{y \in \mathcal{Q} : \pi_y(I) = 1\}$.

The correctness of this construction rests on the observation that closure under boolean operations implies closure under symmetric difference. Hence A_L is a subautomaton of the átomaton of L , see Example 3.20, leading to the

above algorithm. Since the basis \mathcal{Q} has $|\mathcal{Q}| \leq |C| \leq |\{w^{-1}L : w \in \Sigma^*\}|$ it follows that *the minimal xor automaton is never larger than the minimal dfa of L* , see [25].

For our concrete example $L_n = (a + b)^*a(a + b)^n$ take the átomaton of Example 1.1. Its reachable subsets form the set $C = \{S \subseteq Z : x \notin S, z_0 \in S\}$. One can verify that (i) the closure of $\mathcal{Q} = \{\{z_i\} : 0 \leq i \leq n + 1\}$ under symmetric difference is the closure of C and (ii) \mathcal{Q} is minimal. The induced nfa is the minimal xor automaton of Example 1.1.

Example 3.22 (The Jiromaton). If $\mathcal{V} = \text{JSL}$ then \mathbb{G}_*A is the nondeterministic closure automaton $(J(Q), \delta, I)$ where $J(Q)$ is the closure space of Example 3.6 and the initial states are $I = \{z \in J(Q) : z \leq q_0\}$. The canonical $\overline{\text{JSL}}$ -nfa $\mathbb{G}_*(A_L)$ is called the *jiromaton* of L , being based on the join-irreducibles of the corresponding semilattice – in analogy to the átomaton being based on the atoms of the corresponding boolean algebra. Jiromata were first studied by Denis, Lemay and Terlutte [13] who called them *canonical residual finite-state automata*.

The states of the jiromaton are those nonempty left derivatives of L that are not unions of other left derivatives. More precisely, let Q'_L be the subsemilattice of $\mathcal{P}\Sigma^*$ generated by the left derivatives of L . Then the state space of the jiromaton is

$$\mathcal{Q}_L = J(Q'_L).$$

Therefore, the jiromaton has no more states than the minimal dfa. Its structure is analogous to the átomaton: $K \in \mathcal{Q}_L$ is initial iff $K \subseteq L$, final iff $\varepsilon \in K$ and $K \xrightarrow{a} K'$ iff $K' \subseteq a^{-1}K$.

For the language $L_n = (a + b)^*a(a + b)^n$ the join-irreducible left derivatives are

$$z_0 = L_n, \quad \text{and} \quad z_{n+1-i} = L_n + (a + b)^i \quad (i = 0, \dots, n),$$

cf. Example 2.12. This leads to the jiromaton depicted in Example 1.1.

Example 3.23 (The Distromaton). If $\mathcal{V} = \text{DL}$ then $\mathbb{G}_*A = (J(Q), \delta, I)$ with initial states $I = \{z \in J(Q) : z \leq q_0\}$. We call the canonical $\overline{\text{DL}}$ -nfa $\mathbb{G}_*(A_L)$ the *distromaton* of L . Its states

$$\mathcal{Q}_L = J(Q'_L)$$

are the join-irreducibles of the sublattice Q'_L of $\mathcal{P}\Sigma^*$ generated by the left derivatives of L . Hence \mathcal{Q}_L consists of all finite intersections $\bigcap_i w_i^{-1}L$ not arising as finite unions of other such intersections. The structure is again analogous to the átomaton and the jiromaton: $K \in \mathcal{Q}_L$ is initial iff $K \subseteq L$, final iff $\varepsilon \in K$ and $K \xrightarrow{a} K'$ iff $K' \subseteq a^{-1}K$. There is another way to construct the distromaton, analogous to the construction of the átomaton:

1. Take the minimal pointed dfa (Z, δ_a, F, z_0) for the reversed language L^{rev} and order Z by language-inclusion: $z \leq z'$ iff $\mathcal{L}_Z(z) \subseteq \mathcal{L}_Z(z')$.

2. Build the pointed \overline{T}_Σ -coalgebra $(Z^{op}, \delta'_a, \downarrow z_0, F)$ with the poset of states Z^{op} , the final states $\downarrow z_0$ (the down-set of z_0 in Z) and transitions defined by

$$z \xrightarrow{a} z' \quad \text{iff} \quad z \leq \delta_a(z') \text{ in } Z.$$

The initial states F are down-closed in Z^{op} (because $z \in F$ iff $\varepsilon \in \mathcal{L}_Z(z)$) and the final states $\downarrow z_0$ are up-closed in Z^{op} , as required. The proof that this is isomorphic to the distromaton is analogous to our earlier argument regarding the átomaton. Briefly, let $T'_\Sigma = 2 \times \text{ld}^\Sigma : \text{Poset}_f \rightarrow \text{Poset}_f$ where 2 is the two-chain. Then there is a dual equivalence

$$H : (\text{Coalg}_*(T'_\Sigma))^{op} \rightarrow \text{Coalg}_*(\overline{T}_\Sigma),$$

which ‘reverses’ finite pointed deterministic automata equipped with a compatible ordering. The minimal T'_Σ -coalgebra for L is the usual minimal dfa, now equipped with the language-inclusion ordering. Its image under H is again minimal, yielding the above description of the distromaton.

Corollary 3.24. *The átomaton and distromaton of a language L have the same number of states, equal to the number of states of the minimal dfa for the reversed language L^{rev} .*

Example 3.25. (1) In our running example $L_n = (a+b)^*a(a+b)^n$ we start with the minimal dfa for L_n^{rev} (see Example 3.20) and obtain the distromaton in Example 1.1, with the ordering given by

$$x \leq z_i \leq z_0 \quad \text{for all } i = 1, \dots, n,$$

and otherwise incomparable pairs. Note also that this distromaton arises from the jiomaton by adding a final sink state, see also Corollary 4.5.

- (2) Unlike the jiomaton, the átomaton (and the distromaton) can have exponentially more states than the minimal dfa. For example, the minimal dfa for the language L_n^{rev} has $n+2$ states (see Example 3.20) whereas the átomaton has 2^{n+1} states. Indeed, this is the number of states of the minimal dfa for $L_n = (L_n^{rev})^{rev}$.

Remark 3.26. The construction of the jiomaton is based on the category JSL of join-semilattices. Why do we not consider *meet*-semilattices in our paper? The reason is that, with the usual (existential) acceptance mode of nfes, they do not work. And with universal acceptance they can be translated into jiomata.

To see this, let MSL be the category of meet-semilattices with a top element. In analogy to Example ?? the minimal MSL-dfa for a regular language $L \subseteq \Sigma^*$ has states

$$\mathcal{Q}_L = \text{all finite intersections of left derivatives of } L \text{ (including } \Sigma^* \text{)}.$$

The corresponding nfa (called the ‘miromaton’ for the purpose of this remark) has as states the meet-irreducibles of the semilattice \mathcal{Q}_L . Its transitions are

given by $K \xrightarrow{a} K'$ iff $K' \subseteq a^{-1}K$, the initial states are those meet-irreducibles containing L , and the final states are those meet-irreducibles containing ε .

Generally the miromaton does not accept L . For the language L_n of Example 2.12, all states of the miromaton are initial since all derivatives contain L_n . Hence ε is accepted, although ε does not lie in L_n .

However, in the universal acceptance mode, where a word is accepted iff every computation of it terminates in a final state, the miromaton does accept L . The reason is that it is isomorphic to the jiomaton of the complement language $\bar{L} = \Sigma^* \setminus L$. Indeed, the join-irreducible derivatives of \bar{L} are precisely the complements of the states of the miromaton. Consequently, the jiomaton of \bar{L} is isomorphic (via the complement bijection) to the miromaton of L . For more on automata in MSL the reader can consult Klíma and Polák [19].

Example 3.27 (The Universal Automaton). An important construction of a canonical nfa for a regular language L is the *universal automaton* [20]. (The name stems from the fact that all state-minimal nfes accepting L embed into it). It is the nfa whose states \mathcal{Q}_L are again all finite intersections of left derivatives. There is a transition $K \xrightarrow{a} K'$ for $a \in \Sigma$ iff $a^{-1}K \subseteq K'$, the initial states are those states K contained in L , and state K is final iff $\varepsilon \in K$. A slight extension of our approach allows to cover the universal automaton as follows. We precompose the equivalence functor $G : \text{JSL}_f \rightarrow \overline{\text{JSL}}$ of Lemma 3.8 with the free-semilattice functor $F : \text{Poset}_f \rightarrow \text{JSL}_f$ (assigning to a finite poset the join-semilattice of all upwards closed subsets, ordered by inclusion) and the forgetful functor $U : \text{MSL}_f \rightarrow \text{Poset}_f$:

$$\text{MSL}_f \xrightarrow{U} \text{Poset}_f \xrightarrow{F} \text{JSL}_f \xrightarrow{G} \overline{\text{JSL}}.$$

By consecutively applying U , F and G to the minimal MSL-dfa for L , see Remark 3.26, we obtain the following construction of the universal automaton:

- (1) Apply U to obtain the underlying ordered dfa of the minimal MSL-dfa (with the same states \mathcal{Q}_L and the same automata structure).
- (2) Then apply F to construct a JSL-dfa whose states are the upwards closed subsets of \mathcal{Q}_L .
- (3) Since for any finite poset Q we have $J(F(Q)) \cong Q$, applying G yields a nondeterministic closure automaton with states \mathcal{Q}_L . Is easy to see that its transition structure is precisely the structure of the universal automaton.

In our example $L_n = (a + b)^*a(a + b)^n$ the set of states is the same as for the distromaton, hence the universal automaton coincides with the distromaton.

4. State Minimality and Universal Properties

In this section we characterize the canonical nfes by universal properties and discuss their state minimality. We begin with the jiomaton. For every nfa N

with states Q , we denote by

$$\mathcal{L}_N = \{\mathcal{L}_N(I) : I \subseteq Q\}$$

the *language semilattice* of N , consisting of all languages accepted by all choices of initial states. This is a subsemilattice of $\mathcal{P}\Sigma^*$ since $\mathcal{L}_N(-)$ preserves finite unions (including $\mathcal{L}_N(\emptyset) = \emptyset$).

Definition 4.1. A regular language L is *intersection-closed* if every binary intersection of left derivatives of L is a union of left derivatives of L .

Example 4.2. (1) The language $L_n = (a+b)^*a(a+b)^n$ is intersection-closed. To see this, recall the left derivatives of L_n from Example 2.12.

(2) \emptyset , Σ^* and $\{w\}$ for $w \in \Sigma^*$ are intersection-closed.

(3) The language

$$K_n = \{w \in \{0,1\}^n : w \text{ contains an odd number of 1's}\} \quad (n \in \omega)$$

is intersection-closed: its proper left derivatives take the form K_m and $\{0,1\}^m \setminus K_m$ for $0 \leq m < n$, so any two distinct left derivatives have empty intersection.

(4) Fix a natural number n and real numbers k_1, \dots, k_n and t . Then the language

$$L = \{w \in 2^n : \sum_{i=1}^n k_i w_i \geq t\}$$

which models the behaviour of an artificial neuron, see [21], is intersection-closed. Indeed, if $|w| \neq |v|$ or $|w| = |v| > n$ then $w^{-1}L \cap v^{-1}L = \emptyset$. And if $|w| = |v| \leq n$ where w.l.o.g. $\sum_{i \leq |w|} k_i \cdot w_i \leq \sum_{i \leq |v|} k_i \cdot v_i$, then $w^{-1}L \subseteq v^{-1}L$ and therefore $w^{-1}L \cap v^{-1}L = w^{-1}L$.

(5) Every linear subspace $L \subseteq \mathbb{Z}_2^n$ (viewed as a language over the alphabet $\{0,1\}$) is intersection-closed. Indeed, if $|w| \neq |v|$ or if $|w| = |v| > n$ then $w^{-1}L \cap v^{-1}L = \emptyset$. Otherwise if $|w| = |v| \leq n$ and $w^{-1}L \cap v^{-1}L \neq \emptyset$ then $w^{-1}L = v^{-1}L$. To see this, choose $x \in w^{-1}L \cap v^{-1}L$. Then $vx, wx \in L$ and therefore $vx + wx \in L$ (since L is a subspace). It follows that, for every $y \in w^{-1}L$, we have (since $+$ is bit-wise xor)

$$vy = vx + wx + wy \in L,$$

so $y \in v^{-1}L$. Hence $w^{-1}L \subseteq v^{-1}L$, and the other inclusion follows by symmetry.

Theorem 4.3. *The jiraton of any intersection-closed language is a state-minimal nfa.*

Proof. Given an intersection-closed language L , let Q_L be the join-subsemilattice of $\mathcal{P}\Sigma^*$ generated by the left derivatives of L . For every reachable pointed nfa N accepting L we prove that N has at least $|J(Q_L)|$ states, where $J(Q_L)$ is the state set of the jiomaton. This clearly implies that the jiomaton is a state-minimal nfa.

(1) For the set Q of states of N , the function $I \mapsto \mathcal{L}_N(I)$ is an epimorphism

$$\mathcal{L}_N(-) : \mathcal{P}Q \twoheadrightarrow \mathcal{L}_N$$

in JSL . Consequently, \mathcal{L}_N has at most as many join-irreducibles as $\mathcal{P}Q$ (since the join-irreducibles of any finite semilattice form a set of generators of minimal cardinality). It follows that

$$|Q| \geq |J(\mathcal{L}_N)|.$$

To prove our desired inequality $|Q| \geq |J(Q_L)|$, it therefore suffices to show

$$|J(\mathcal{L}_N)| \geq |J(Q_L)|.$$

(2) We verify $Q_L \subseteq \mathcal{L}_N$. Indeed, given any word $w \in \Sigma^*$ let I be the set of states reachable from some initial state on input w . Then

$$\mathcal{L}_N(I) = w^{-1}L,$$

so N accepts all left derivatives of L . Since \mathcal{L}_N is closed under union, the above inclusion follows.

(3) Q_L and \mathcal{L}_N have the same top element, that is,

$$\bigcup Q_L = \mathcal{L}_N(Q).$$

The inclusion \subseteq follows from (2), and the opposite inclusion follows from the observation that every state q of N accepts a subset of a left derivative of L . Indeed, since N is reachable, there exists a word $w \in \Sigma^*$ such that q is reached on input w from some initial state. Then $\mathcal{L}_N(\{q\}) \subseteq w^{-1}L$.

(4) The inclusion

$$Q_L \hookrightarrow \mathcal{L}_N$$

preserves meets. For the empty meet (i.e. the top element) use (3). For non-empty meets one uses the fact that L is intersection-closed and that, in Q_L (as well as \mathcal{L}_N), the meet of two elements is their intersection whenever this intersection lies in Q_L (or \mathcal{L}_N , respectively). Thus for the opposite posets we obtain a monomorphism

$$Q_L^{op} \hookrightarrow \mathcal{L}_N^{op}$$

of JSL_f . The self-duality $X \mapsto X^{op}$ of JSL_f yields an epimorphism

$$\mathcal{L}_N \twoheadrightarrow Q_L$$

of JSL_f . As in (1), this proves the inequality $|J(\mathcal{L}_N)| \geq |J(Q_L)|$. \square

Remark 4.4. The converse of this theorem is generally false: the language $L = \{aa\}$ is not intersection-closed, but its jiomaton is state-minimal, see [4, Example 4.5].

Corollary 4.5. *The átomaton and distromaton of an intersection-closed language have at most one more state than the jiomaton.*

Proof. If L is intersection-closed, then the subsemilattice and the sublattice of $\mathcal{P}\Sigma^*$ generated by the left derivatives of L agree, except that the latter always contains Σ^* . Hence the distromaton has at most one more state than the jiomaton. By Corollary 3.24 the átomaton has the same number of states as the distromaton. \square

By Corollary 3.24 we further deduce:

Corollary 4.6. *Let L be an intersection-closed language, and let n be the number of states of the minimal dfa of L^{rev} . Then a state-minimal nfa for L has (i) n states if every word of Σ^* lies in some left derivative of L , and (ii) $n - 1$ states otherwise.*

Theorem 4.7. *If L is a language whose state-minimal nfes have n states and whose state-minimal dfa has 2^n states, then the jiomaton of L is state-minimal.*

Proof. Let $N = (Q, R_a, F)$ be a state-minimal nfa accepting L via $I \subseteq Q$. Turn it into a pointed T_Σ -coalgebra $A = (\mathcal{P}Q, \gamma, I)$ in JSL via the subset construction. By assumption, since A has 2^n states, this is a state-minimal dfa accepting L ; in particular, it is a reachable pointed T_Σ -coalgebra. Then the surjective morphism $A \rightarrow A_L$ implies that A_L has no more than n join-irreducibles, so the jiomaton is state-minimal. \square

The jiomaton $J(L)$ can be characterized by the following sort of minimality:

Theorem 4.8. (a) *$J(L)$ accepts the least number of languages, i.e., for every nfa N accepting L we have*

$$|\mathcal{L}_{J(L)}| \leq |\mathcal{L}_N|.$$

(b) *Among nfes with the same number of accepted languages, $J(L)$ is state-minimal. That is, for every nfa N accepting L with $|\mathcal{L}_{J(L)}| = |\mathcal{L}_N|$ we have*

$$|J(L)| \leq |N|.$$

Here $|N|$ denotes the number of states of N .

(c) *Among nfes with the same number of accepted languages and states, $J(L)$ has the largest number of transitions. That is, for every nfa N accepting L with $|\mathcal{L}_{J(L)}| = |\mathcal{L}_N|$ and $|J(L)| = |N|$, we have*

$$\mathbf{tr}(J(L)) \geq \mathbf{tr}(N).$$

Here $\mathbf{tr}(N) = \sum_a |R_a|$ is the number of transitions of an nfa $N = (Z, R_a, F)$.

Moreover, these properties determine $J(L)$ uniquely up to isomorphism.

Proof. Each state of $J(L)$ accepts a left derivative of L , and $J(L)$ accepts precisely all unions of left derivatives of L (cf. Example 3.22). Since any nfa N accepting L accepts all unions of left derivatives of L , we conclude $\mathcal{L}_{J(L)} \subseteq \mathcal{L}_N$ and thus $|\mathcal{L}_{J(L)}| \leq |\mathcal{L}_N|$. Suppose now that $|\mathcal{L}_{J(L)}| = |\mathcal{L}_N|$, which means that N accepts precisely the unions of left derivatives of L . Then each left derivative that is a state of $J(L)$ has a distinct state in N accepting it, as it cannot arise as the union of other left derivatives, so $|J(L)| \leq |N|$. Finally, if $|\mathcal{L}_{J(L)}| = |\mathcal{L}_N|$ and $|J(L)| = |N|$, then there is language-preserving bijection between the states of $J(L)$ and N , so we can assume they have the same set of states. Given a transition $K \xrightarrow{a} K'$ in N we must have $K' \subseteq a^{-1}K$, so there is a corresponding transition in J_L . Hence $\text{tr}(N) \leq \text{tr}(J(L))$. Moreover, if $\text{tr}(N) = \text{tr}(J(L))$ then the previous argument shows that N and $J(L)$ are isomorphic. Thus the properties (a), (b) and (c) determine $J(L)$ uniquely up to isomorphism. \square

Although the canonical nfas are generally not state-minimal, they are state-minimal amongst certain subclasses of nfas.

Theorem 4.9. *The átomaton of a regular language L is state-minimal amongst all nfas accepting L whose accepted languages are closed under complement.*

Proof. Let N be an nfa accepting L whose accepted languages are closed under complement (so that \mathcal{L}_N is a boolean subalgebra of $\mathcal{P}\Sigma^*$). It suffices to show the inequalities

$$|Q| \geq |\text{At}(\mathcal{L}_N)| \geq |\text{At}(Q_L)|$$

where Q is the set of states of N and Q_L is the boolean algebra generated by the left derivatives of L . The first inequality follows from item (1) in the proof of Theorem 4.3. For the second one, use that Q_L is a subalgebra of the boolean algebra \mathcal{L}_N . \square

Theorem 4.10 ([25]). *The minimal xor automaton for L is state-minimal amongst nfas accepting L via \mathbb{Z}_2 -weighted acceptance.*

Concerning the jiomaton, we give a mild generalization of a result in [13]. Recall that nfas accepting L also accept all unions of its left derivatives. Then we can conclude from Theorem 4.8:

Corollary 4.11. *The jiomaton of a regular language L is state-minimal amongst nfas accepting precisely the unions of L 's left derivatives.*

Example 4.12. Let N be an nfa accepting L via initial states I . If every singleton set of states is reachable from I then N accepts precisely the unions of left derivatives of L . Thus, N is no smaller than the jiomaton of L .

Theorem 4.13. *The distromaton of a regular language L is state-minimal amongst all nfas accepting L whose accepted languages are closed under intersection.*

Proof. Analogous to the proof of Theorem 4.9. \square

Remark 4.14. How does the size of the jiomaton compare to the átomaton? As demonstrated in Example 3.25, there are cases for which the jiomaton is exponentially smaller than the átomaton. On the other hand, in Example 1.1 the átomaton is strictly smaller than the jiomaton.

5. Conclusions and Future Work

It is well-known that minimal dfa exist and are easy to construct; as a counterpoint, however, state-minimal nfas are not unique and not easily constructed. Instead, in the literature several canonical nfas are studied. We have demonstrated that, from an abstract coalgebraic perspective, such canonical nfas arise in a uniform way from the minimal dfa interpreted in a locally finite variety. In so doing we have unified previous work from three sources [11, 13, 25] and introduced a new canonical nondeterministic acceptor, the distromaton. We also identified a class of languages where canonical *state-minimal* nfas exist. These results depend heavily on a coalgebraic approach to automata theory, providing not only new structural insights and construction methods but also a new perspective on what a state-minimal acceptor actually is.

Our approach in this paper is to start, for a given a regular language L , with the minimal deterministic automaton for L in a locally finite variety \mathcal{V} of algebras. We then use an equivalence between the category of finite algebras of \mathcal{V} and a suitable category $\bar{\mathcal{V}}$ that represents those finite algebras by (often) much smaller structures. This equivalence takes the minimal deterministic automaton in \mathcal{V} to a canonical nondeterministic automaton in the equivalent category $\bar{\mathcal{V}}$. We worked out that equivalence for the varieties \mathcal{V} of boolean algebras, distributive lattices, semilattices, \mathbb{Z}_2 -vector spaces and pointed sets. An interesting open question is whether there is a systematic uniform way to identify, for an arbitrary locally finite variety, the “suitable” equivalent category.

In the particular case of the variety of join-semilattices we introduced nondeterministic closure automata, viz. nfas in the category of closure spaces, mainly as a tool for constructing the jiomaton. However, nondeterministic closure automata bear interesting structural properties themselves which we did not discuss here in depth. In [4] we provide a detailed investigation of these machines, from which we derive additional and more general criteria for the (state-)minimality of nfas.

Another point we aim to investigate in more detail are the algorithmic aspects of the state-minimization of nfas. Although this problem is known to be PSPACE-complete in general, the canonicity of our nfas suggests that – at least for certain natural subclasses of nfas – efficient state-minimization procedures may be in reach. For example, for weighted automata with weights in a field (such as \mathbb{Z}_2) minimization can be computed in polynomial time, see e.g. [14, 8]). We leave the study of further complexity-related issues for future work.

References

- [1] Adámek, J., Bonchi, F., König, B., Hülsbusch, M., Milius, S., Silva, A., 2012. A coalgebraic perspective on minimization and determinization. In: Birkedal, L. (Ed.), Proc. Foundations of Software Science and Computation Structures (FoSSaCS). Vol. 7213 of Lecture Notes Comput. Sci. Springer, pp. 58–73.
- [2] Adámek, J., Milius, S., Moss, L. S., Sousa, L., 2013. Well-pointed coalgebras. *Log. Methods Comput. Sci.* 9 (3:2), 51 pp.
- [3] Adámek, J., Milius, S., Myers, R., Urbat, H., 2014. Generalized Eilenberg Theorem I: Local Varieties of Languages. In: Proc. FOSSACS 2014. Vol. 8412 of Lect. Notes in Comput. Sci. Springer, pp. 366–380.
- [4] Adámek, J., Milius, S., Myers, R. S., Urbat, H., 2014. On continuous nondeterminism and state minimality. In: Proc. Mathematical Foundations of Programming Science (MFPS XXX). Vol. 308 of Electron. Notes Theor. Comput. Sci. pp. 3–23.
- [5] Adámek, J., Milius, S., Velebil, J., 2006. Iterative algebras at work. *Math. Structures Comput. Sci.* 16 (6), 1085–1131.
- [6] Barr, M., 1993. Terminal coalgebras in well-founded set theory. *Theoret. Comput. Sci.* 114 (2), 299–315.
- [7] Bezhanishvili, N., Kupke, C., Panangaden, P., 2012. Minimization via duality. In: Proc. Workshop on Logic, Language, Information and Computation (WoLLIC’12). pp. 191–205.
- [8] Bonchi, F., Bonsangue, M. M., Boreale, M., Rutten, J. J. M. M., Silva, A., 2012. A coalgebraic perspective on linear weighted automata. *Inform. and Comput.* 211, 77–105.
- [9] Bonchi, F., Bonsangue, M. M., Rutten, J. J. M. M., Silva, A., 2012. Brzozowski’s algorithm (co)algebraically. In: Constable, R., Silva, A. (Eds.), *Logic and Program Semantics*. pp. 12–23.
- [10] Bonsangue, M. M., Milius, S., Silva, A., 2013. Sound and complete axiomatizations of coalgebraic language equivalence. *ACM Trans. Comput. Log.* 14 (1:7).
- [11] Brzozowski, J., Tamm, H., 2011. Theory of Automata. In: Mauri, G., Leporati, A. (Eds.), Proc. 15th International Conference on Developments in Language Theory (DLT’11). Vol. 6795 of Lecture Notes Comput. Sci. Springer, pp. 105–116.
- [12] Brzozowski, J. A., 1962. Canonical regular expressions and minimal state graphs for definite events. In: *Mathematical theory of Automata*. Volume 12 of MRI Symposia Series. Polytechnic Press, Polytechnic Institute of Brooklyn, N.Y., pp. 529–561.

- [13] Denis, F., Lemay, A., Terlutte, A., 2002. Residual finite state automata. *Fund. Inform.* XX, 1–30.
- [14] Eisner, J., 2003. Simpler and more general minimisation for weighted finite-state automata. In: *Proc. Joint Meeting of the Human Language Technology Conference and the North American Chapter of the Association for Computational Linguistics (HLT-NAACL)*. pp. 64–71.
- [15] Hasuo, I., Jacobs, B., Sokolova, A., 2007. Generic trace semantics via coinduction. *Log. Methods Comput. Sci.* 3 (4:11).
- [16] Jacobs, B., Silva, A., Sokolova, A., 2012. Trace semantics via determinization. In: Pattinson, D., Schröder, L. (Eds.), *Proc. Coalgebraic Methods in Computer Science (CMCS’12)*. Vol. 7399 of *Lecture Notes Comput. Sci.* Springer, pp. 109–129.
- [17] Jipsen, P., 2012. Categories of algebraic contexts equivalent to idempotent semirings and domain semirings. In: Kahl, W., Griffin, T. G. (Eds.), *Proc. Relational and Algebraic Methods in Computer Science (RAMICS’12)*. Vol. 7560 of *Lecture Notes in Computer Science*. Springer, pp. 195–206.
- [18] Johnstone, P., 1982. *Stone Spaces*. Cambridge University Press.
- [19] Klíma, O., Pol’ak, L., 2008. On varieties of meet automata. *Theoretical Computer Science* 407, 278–289.
- [20] Lombardy, S., Sakarovitch, J., 2008. The universal automaton. In: Flum, J., Grädel, E., Wilke, T. (Eds.), *Logic and Automata*. Vol. 2 of *Texts in Logic and Games*. Amsterdam University Press, pp. 457–504.
- [21] McCulloch, W. S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics* 5 (4), 115–133.
- [22] Milius, S., 2010. A sound and complete calculus for finite stream circuits. In: *Proc. 25th Annual Symposium on Logic in Computer Science (LICS’10)*. IEEE Computer Society, pp. 449–458.
- [23] Myers, R. S. L., Adámek, J., Milius, S., Urbat, H., 2014. Canonical non-deterministic automata. In: Bonsangue, M. M. (Ed.), *Proc. Coalgebraic Methods in Computer Science (CMCS’14)*. Vol. 8446 of *Lecture Notes Comput. Sci.* Springer, pp. 189–210.
- [24] Silva, A., Bonchi, F., Bonsangue, M. M., Rutten, J. J. M. M., 2013. Generalizing determinization from automata to coalgebras. *Log. Methods Comput. Sci.* 9 (1:9).
- [25] Vuillemin, J., Gama, N., May 2010. Efficient Equivalence and Minimization for Non Deterministic Xor Automata. Research report, LIENS.
URL <http://hal.inria.fr/inria-00487031>

Appendix A. Semilattices and Closure Spaces

We prove the equivalence of JSL_f and the category of finite strict closure spaces. Recall from Example 3.6 that for every finite join-semilattice Q we denote by GQ the closure space of all join-irreducible elements of Q with the strict closure operator

$$\overline{M} = \{j \in GS : j \leq \bigvee M\} \quad \text{for } M \subseteq GQ.$$

For every morphism $r : Q \rightarrow Q'$ in JSL_f let $Gr \subseteq GQ \times GQ'$ be the relation given by

$$Gr[x] = \{y \in GQ' : y \leq r(x)\}.$$

Remark A.1. The relation $R = Gr$ is continuous, see Definition 3.7. Indeed,

(a) $R[x]$ is closed in GQ' for every $x \in GQ$: given $y \in \overline{R[x]}$ we have

$$y \leq \bigvee R[x] = \bigvee_{z \leq r(x)} z \leq r(x),$$

thus $y \in R[x]$.

(b) $R[\overline{M}] \subseteq \overline{R[M]}$ for every $M \subseteq GQ$. To see this, observe that since r preserves joins,

$$\bigvee r[\overline{M}] \leq \bigvee r[M]$$

(indeed $z \in \overline{M}$ implies $r(z) \leq r[\bigvee M] = \bigvee r[M]$). Hence for every $y \in R[\overline{M}]$ we have $y \leq \bigvee r[\overline{M}] \leq \bigvee r[M]$ and therefore $y \in \overline{R[M]}$, as required.

In the following we use the notation $R : X \rightarrow X'$ for a continuous relation $R \subseteq X \times X'$ between closure spaces.

Example A.2. (1) For every closure space X the relation

$$I_X : X \rightarrow X, \quad I_X[x] = \overline{\{x\}},$$

is continuous. Indeed, $I_X[x] = \overline{\{x\}}$ is clearly closed, and $I_X[\overline{M}] \subseteq \overline{I_X[M]}$ because given $x \in \overline{M}$, we have

$$\overline{\{x\}} \subseteq \overline{M} \subseteq \overline{I_X[M]}.$$

(2) For every finite strict closure space X , the set

$$CX = \text{all closed subsets of } X$$

is a join-semilattice w.r.t. inclusion: the join of M_1 and M_2 is $\overline{M_1 \cup M_2}$. Observe that every join-irreducible member of CX has the form $\overline{\{x\}}$ for some $x \in X$. Indeed, for every closed subset $M \subseteq X$ we have

$$M = \bigvee_{x \in M} \overline{\{x\}}$$

so join-irreducibility implies $M = \overline{\{x\}}$ for some x . (However, the converse is wrong: the closure of a singleton need not be join-irreducible in CX .) Observe that the relation $I_{G(CX)}$, see (1), consists of all pairs $(\overline{\{x\}}, \overline{\{y\}})$ with $\overline{\{x\}}, \overline{\{y\}} \in G(CX)$ and $\overline{\{y\}} \subseteq \overline{\{x\}}$.

(3) The relation $\iota_X : X \rightarrow G(CX)$ defined by

$$\iota_X[x] = \{\overline{\{y\}} \in G(CX) : \overline{\{y\}} \subseteq \overline{\{x\}}\}$$

is continuous. Indeed, $\iota_X[x]$ is clearly closed. And $\iota_X[\overline{M}] \subseteq \overline{\iota_X[M]}$ holds since $\overline{\iota_X[M]}$ consists of all join-irreducible closed subsets of \overline{M} by the definition of the closure operator, and every element of $\iota_X[\overline{M}]$ is such a subset.

(4) Another continuous relation is

$$\varepsilon_X : G(CX) \rightarrow X, \quad \varepsilon_X[\overline{\{x\}}] = \overline{\{x\}}.$$

Indeed, clearly $\varepsilon_X[\overline{\{x\}}] = \overline{\{x\}}$ is closed. And $\varepsilon_X[\overline{M}] \subseteq \overline{\varepsilon_X[M]}$ because given $y \in \varepsilon_X[\overline{M}]$, we have $\overline{\{x\}} \in \overline{M}$ with $y \in \overline{\{x\}}$. Now $\overline{\{x\}} \in \overline{M}$ implies $\overline{\{x\}} \subseteq \bigvee_{CX} M = \bigcup M = \varepsilon_X[M]$. Therefore $y \in \varepsilon_X[M]$.

Remark A.3. Every continuous relation R fulfills

$$x \in \overline{\{y\}} \Rightarrow R[x] \subseteq R[y].$$

Indeed, we have $R[x] \subseteq R[\overline{\{y\}}] \subseteq \overline{R[y]} = R[y]$.

Lemma A.4. *Let Q and Q' be finite join-semilattices. For every continuous relation $R : GQ \rightarrow GQ'$ there exists a unique JSL_f -morphism $r : Q \rightarrow Q'$ with $R = Gr$.*

Proof. For every element $x \in Q$ denote by M_x the set of all join-irreducible elements under x .

(1) Uniqueness. If $R = Gr$, then since $x = \bigvee M_x$ implies

$$r(x) = \bigvee r[M_x] = \bigvee R[M_x],$$

we have a formula for r :

$$r(x) = \bigvee R[M_x].$$

(2) Existence. We must prove that the above formula defines a JSL_f -homomorphism r with $R = Gr$. To prove that r preserves joins, observe that $r(x \vee y) = \bigvee_{z \leq x \vee y} \bigvee_{zRu} u$ where z ranges through GQ . From $z \leq (\bigvee M_x) \vee (\bigvee M_y)$ we get $z \in \overline{M_x \cup M_y}$. Thus each u lies in $R[\overline{M_x \cup M_y}] \subseteq \overline{R[M_x \cup M_y]}$, hence

$$r(x \vee y) = \bigvee \overline{R[M_x \cup M_y]} = \bigvee R[M_x \cup M_y].$$

From $R[M_x \cup M_y] = R[M_x] \cup R[M_y]$ we get $r(x \vee y) = r(x) \vee r(y)$, as required.

To prove $R = Gr$, consider any $y \in R[x]$: since $x \in M_x$, we have $y \leq \bigvee R[M_x]$, that is, $y \in Gr[x]$. Conversely, consider any $y \in Gr[x]$. We know that $y \leq \bigvee R[M_x]$, i.e., y lies in the closure $\overline{R[M_x]}$. By Remark A.3, for every $z \in M_x$ we have $R[z] \subseteq R[x]$, thus $R[M_x] = R[x]$, so $y \in \overline{R[x]} = R[x]$, as required. \square

Remark A.5. Let $Q \xrightarrow{r} Q' \xrightarrow{s} Q''$ be morphisms in \mathbf{JSL}_f and $GQ \xrightarrow{R} GQ' \xrightarrow{S} GQ''$ the corresponding continuous relations. Then the relation corresponding to the composite $sr : Q \rightarrow Q''$ is given by

$$G(sr)[x] = \overline{S \circ R[x]}$$

where \circ is the usual composition of relations. Indeed, given $x \in GQ$ we have $r(x) = \bigvee_{i=1}^n y_i$ where y_1, \dots, y_n are the join-irreducibles under $r(x)$. Then for all $x'' \in GQ''$ we conclude:

$$x'' \in G(sr)[x] \quad \text{iff} \quad x'' \leq s\left(\bigvee y_i\right) = \bigvee s(y_i).$$

Moreover, $Q \circ P[x]$ consists of all $y'' \in GQ''$ for which some y_i fulfills $y'' \leq s(y_i)$. Hence

$$x'' \in \overline{S \circ R[x]} \quad \text{iff} \quad x'' \leq \bigvee y'' = \bigvee s(y_i).$$

Recall from Definition 3.7 that the continuous composition of two continuous relations $X \xrightarrow{R} X' \xrightarrow{S} X''$ is the relation $S \bullet R : X \rightarrow X''$ with

$$S \bullet R[x] = \overline{S \circ R[x]}.$$

Corollary A.6. $G(sr) = Gs \bullet Gr$ for all composable morphisms in \mathbf{JSL}_f .

Example A.7. (1) The relations ι_X and ε_X of Example A.2 clearly fulfill

$$\iota_X \bullet \varepsilon_X = I_{G(CX)} \quad \text{and} \quad \varepsilon_X \bullet \iota_X = I_X.$$

(2) We also have

$$\varepsilon_X \bullet G(id_{CX}) \bullet \iota_X = I_X$$

because $G(id_{CX}) \bullet \iota_X = I_{G(CX)} \bullet \iota_X = \iota_X$.

Lemma A.8. *The continuous composite of continuous relations is always continuous.*

Proof. Given continuous relations $X \xrightarrow{S} X' \xrightarrow{R} X''$ then $S \bullet R[x] = \overline{S \circ R[x]}$ is clearly closed. To prove $S \bullet R[\overline{M}] \subseteq \overline{S \bullet R[M]}$, first observe that $S \circ R$ has the corresponding property:

$$S \circ R[\overline{M}] = S[R[\overline{M}]] \subseteq S[\overline{R[M]}] \subseteq \overline{S \circ R[M]} \subseteq \overline{S \bullet R[M]}.$$

Next y lies in $S \bullet R[\overline{M}]$ iff $y \in \overline{S \circ R[x]}$ for some $x \in \overline{M}$. Hence

$$y \in \overline{S \circ R[x]} \subseteq \overline{S \circ R[\overline{M}]} \subseteq \overline{S \bullet R[M]} = \overline{S \bullet R[M]}. \quad \square$$

Recall that $\overline{\mathbf{JSL}}$ denotes the category of finite strict closure spaces and continuous relations. The composition is defined to be the continuous composition, and the identity morphisms are the relations I_X of Example A.2(1).

Theorem A.9. $\overline{\text{JSL}}$ is a well-defined category equivalent to JSL_f .

Proof. (1) We first prove that category $\overline{\text{JSL}}_0$ of all closure spaces GQ ($Q \in \text{JSL}_f$) and continuous relations with the above composition and identity morphisms is well-defined and equivalent to JSL_f . Indeed, the functor

$$G_0 : \text{JSL}_f \rightarrow \overline{\text{JSL}}_0$$

defined on objects by $Q \mapsto GQ$ and on morphisms by $r \mapsto Gr$ is an equivalence functor. That $\overline{\text{JSL}}_0$ is well-defined and G_0 is a functor follows from Remark A.5, Corollary A.6 and the fact that $I_{GX} = G_0(\text{id}_X)$. Moreover, G_0 is full and faithful by Lemma A.4 and surjective on objects by definition. Hence G_0 is an equivalence.

(2) The statement about $\overline{\text{JSL}}$ now follows. Recall the relation $\iota_X : X \rightarrow G(CX)$ from Example A.2 which by Example A.7 is invertible: $\varepsilon_X \bullet \iota_X = I_X$ and $\iota_X \bullet \varepsilon_X = I_{G(CX)}$. Thus, for every pair X, Y of finite strict closure spaces we have a bijection between all continuous relations $R : X \rightarrow Y$ and all continuous relations $R' : G(CX) \rightarrow G(CY)$, given by $R \mapsto \iota_Y \bullet R \bullet \varepsilon_X$:

$$\begin{array}{ccc} X & \xrightarrow{R} & Y \\ \iota_X \downarrow \uparrow \varepsilon_X & & \varepsilon_Y \uparrow \downarrow \iota_Y \\ G(CX) & \xrightarrow{R'} & G(CY) \end{array}$$

Then Lemma A.4 establishes a bijection G between homomorphisms $r : CX \rightarrow CY$ of semilattices and continuous relations $R : X \rightarrow Y$. By Remark A.5 this bijection preserves composition, thus, $\overline{\text{JSL}}$ is indeed a well-defined category with identity morphisms

$$\varepsilon_X \bullet G(\text{id}_{CX}) \bullet \iota_X = I_X,$$

and the isomorphisms ι_X prove that $\overline{\text{JSL}}$ is equivalent to $\overline{\text{JSL}}_0$. \square