Nondeterministic Syntactic Complexity

Robert S. R. Myers, Stefan Milius[1,⋆], and Henning Urbat[1]

[1]Friedrich-Alexander-Universität Erlangen-Nürnberg

Abstract We introduce a new measure on regular languages: their *nondeterministic syntactic complexity.* It is the least degree of any extension of the 'canonical boolean representation' of the syntactic monoid. Equivalently, it is the least number of states of any *subatomic* nondeterministic acceptor. It turns out that essentially all previous structural work on nondeterministic state-minimality computes this measure. Our approach rests on an algebraic interpretation of nondeterministic finite automata as deterministic finite automata endowed with semilattice structure. Crucially, the latter form a self-dual category.

1 Introduction

Regular languages admit a plethora of equivalent representations: finite automata, finite monoids, regular expressions, formulas of monadic second-order logic, and numerous others. In many cases, the most succinct representation is given by a *nondeterministic finite automaton (nfa).* Therefore, the investigation of state-minimal nfas is of both computational and mathematical interest. However, this turns out to be surprisingly intricate; in fact, the task of minimizing an nfa, or even of deciding whether a given nfa is minimal, is known to be PSPACE-complete [18]. One intuitive reason is that minimal nfas lack structure: a language may have many non-isomorphic minimal nondeterministic acceptors, and there are no clearly identified and easily verifiable mathematical properties distinguishing them from non-minimal ones. As a consequence, all known algorithms for minimizing nfas require some form of exhaustive search. This sharply contrasts the situation for minimal *deterministic finite automata (dfa)*: they can be characterized by a universal property making them unique up to isomorphism, which immediately leads to efficient minimization.

In the present paper, we work towards the goal of bringing more structure into the theory of nondeterministic state-minimality. To this end, we propose a novel algebraic perspective on nfas resting on *boolean representations* of monoids, i.e. morphisms $M \to \mathbf{JSL}(S, S)$ from a monoid M into the endomorphism monoid of a finite join-semilattice S. Our focus lies on quotient monoids of the free monoid Σ^* recognizing a given regular language $L \subseteq \Sigma^*$. The largest such

---

⋆ Supported by Deutsche Forschungsgemeinschaft (DFG) under projects MI 717/5-2 and MI 717/7-1, and as part of the Research and Training Group 2475 "Cybercrime and Forensic Computing" (393541319/GRK2475/1-2019)

monoid is Σ^* itself, while the smallest one is the *syntactic monoid* $\mathsf{syn}(L)$. For both of them, L induces a *canonical boolean representation*

$$\Sigma^* \to \mathbf{JSL}(\mathsf{SLD}(L), \mathsf{SLD}(L)) \qquad \text{and} \qquad \mathsf{syn}(L) \to \mathbf{JSL}(\mathsf{SLD}(L), \mathsf{SLD}(L))$$

on the semilattice $\mathsf{SLD}(L)$ of all finite unions of left derivatives of L. The first representation gives rise to an algebraic characterization of minimal nfas:

Theorem. The size of a state-minimal nfa for L equals the least degree of any extension of the canonical representation of Σ^* induced by L.

Here, the *degree* of a representation refers to the number of join-irreducibles of the underlying semilattice. In the light of this result, it is natural to ask for an analogous automata-theoretic perspective on the canonical representation of $\mathsf{syn}(L)$ and its extensions. For this purpose, we introduce the class of *subatomic* nfas, a generalization of *atomic* nfas earlier introduced by Brzozowski and Tamm [4]. In order to get a handle on them, we employ an algebraic framework that interprets nfas in terms of **JSL**-*dfas*, i.e. deterministic finite automata in the category of semilattices. In this setting, the semilattice $\mathsf{SLD}(L)$ used in the canonical representations naturally arises as the *minimal* **JSL**-dfa for the language L. We shall demonstrate that much of the structure theory of (sub-)atomic nfas reduces to the observation that the category of **JSL**-dfas is *self-dual*. Our main result gives an algebraic characterization of minimal subatomic nfas:

Theorem. The size of a state-minimal subatomic nfa for L equals the least degree of any extension of the canonical representation of $\mathsf{syn}(L)$.

We call the measure suggested by the above theorem the *nondeterministic syntactic complexity* of the language L. It turns out to be extremely natural: as illustrated in Section 5, essentially all existing work on the structure of state-minimal nfas implicitly identifies classes of languages whose nondeterministic state complexity equals their nondeterministic syntactic complexity, and thus is actually concerned with computing minimal subatomic acceptors.

2 Preliminaries

We start by introducing some notation and terminology used in the paper.

Semilattices. A *(join-)semilattice* is a poset $(S, \leq_S)$ in which every finite subset $X \subseteq S$ has a least upper bound, a.k.a. join, denoted by $\bigvee X$. A *morphism* of semilattices is a map preserving all finite joins. Let **JSL** denote the category of join-semilattices and their morphisms. We write $J(S)$ for the set of *join-irreducible* elements of a semilattice S, i.e. those $j \in S$ such that for all finite subsets $X \subseteq S$ with $j = \bigvee X$ one has $j \in X$. Let $2 = \{0, 1\}$ denote the two-element semilattice with $0 \leq 1$. Since $2 \cong (\mathcal{P}(1), \subseteq)$ is the free semilattice on a single generator, morphisms from 2 into a semilattice S correspond uniquely to elements of S. Similarly, a morphism $f\colon S \to 2$ corresponds uniquely to a *prime filter* $F = f^{-1}[1] \subseteq S$, i.e. an upwards closed subset such that $\bigvee X \in F$

implies $X \cap F \neq \emptyset$ for every finite subset $X \subseteq S$. If S is a subsemilattice of a semilattice T, every prime filter F of S can be extended to the prime filter $T \setminus (\downarrow(S \setminus F))$ of T, where $\downarrow X = \{\, t \in T \;:\; t \leq x \text{ for some } x \in X \,\}$ denotes the down-closure of a subset $X \subseteq T$. Equivalently, every morphism $f\colon S \to 2$ can be extended to a morphism $g\colon T \to 2$. In category-theoretic terminology, this means that the semilattice 2 forms an injective object of **JSL**. Finally, the category $\mathbf{JSL}_\mathsf{f}$ of finite semilattices is *self-dual* [20], as witnessed by equivalence functor $\mathbf{JSL}_\mathsf{f} \xrightarrow{\simeq} \mathbf{JSL}_\mathsf{f}^\mathsf{op}$ that sends a semilattice S to its *dual semilattice* S^op obtained by reversing the order, and a morphism $f\colon S \to T$ to the morphism $f^*\colon T^\mathsf{op} \to S^\mathsf{op}$ mapping $t \in T$ to the $\leq_S$-largest element $s \in S$ with $f(s) \leq_T t$. Note that f is *adjoint* to f^*: for $s \in S$ and $t \in T$ we have $f(s) \leq_T t$ iff $s \leq_S f^*(t)$.

Languages. A *language* is a subset L of Σ^*, the set of finite words over an alphabet Σ. We let $\overline{L} = \Sigma^* \setminus L$ denote the *complement* and $L^\mathsf{r} = \{w^\mathsf{r} : w \in L\}$ the *reverse*, where $w^\mathsf{r} = a_n \ldots a_1$ for $w = a_1 \ldots a_n$. The *left derivatives*, *right derivatives* and *two-sided derivatives* of L are, respectively, given by $u^{-1}L = \{w \in \Sigma^* : uw \in L\}$, $Lv^{-1} = \{w \in \Sigma^* : wv \in L\}$ and $u^{-1}Lv^{-1} = \{w \in \Sigma^* : uwv \in L\}$ for $u, v \in \Sigma^*$. More generally, for $U \subseteq \Sigma^*$ the language $U^{-1}L = \bigcup_{u \in U} u^{-1}L$ is called the *left quotient* of L w.r.t. U. We define the following sets of languages generated by L:

- $\mathsf{LD}(L) = \{u^{-1}L : u \in \Sigma^*\}$, the set of all left derivatives of L;
- $\mathsf{SLD}(L)$, its closure under finite union;
- $\mathsf{BLD}(L)$, its closure under all set-theoretic boolean operations;
- $\mathsf{BLRD}(L)$, its closure under all boolean operations and right derivatives.

In other words, $\mathsf{SLD}(L)$ is the $\cup$-semilattice of all left quotients of L, or equivalently, the $\cup$-subsemilattice of $\mathcal{P}(\Sigma^*)$ generated by all left derivatives. Moreover, $\mathsf{BLD}(L)$ and $\mathsf{BLRD}(L)$ form the boolean subalgebras of $\mathcal{P}(\Sigma^*)$ generated by all left derivatives and all two-sided derivatives, respectively.

3 Duality Theory of Semilattice Automata

In this section, we set up the algebraic framework in which nondeterministic automata can be studied. Since it involves considering several different types of automata, it is convenient to view them all as instances of a general categorical concept. For the rest of this paper, let Σ denote a fixed finite input alphabet.

Definition 3.1. Let $\mathscr{C}$ be a category and let $X, Y \in \mathscr{C}$ be two fixed objects. An *automaton* in $\mathscr{C}$ is a quadruple (S, δ, i, f) consisting of an object $S \in \mathscr{C}$ of *states*, a family $\delta = (\delta_a\colon S \to S)_{a \in \Sigma}$ of morphisms representing *transitions*, and two morphisms $i\colon X \to S$ and $f\colon S \to Y$ representing *initial* and *final* states.

$$\begin{tikzcd}
X \arrow[r, "i"] & S \arrow[loop above, "\delta_a"] \arrow[r, "f"] & Y
\end{tikzcd}$$

A *morphism* between automata (S, δ, i, f) and (S', δ', i', f') is given by a morphism $h\colon S \to S'$ in $\mathscr{C}$ preserving transitions, initial states and final states, i.e. making the following diagram commute for all $a \in \Sigma$:

$$\begin{tikzcd}
X \arrow[r, "i"] \arrow[dr, "i'"'] & S \arrow[r, "\delta_a"] \arrow[d, "h"'] & S \arrow[r, "f"] \arrow[d, "h"] & Y \\
& S' \arrow[r, "\delta'_a"'] & S' \arrow[ur, "f'"'] &
\end{tikzcd}$$

Let $\mathbf{Aut}(\mathscr{C})$ denote the category of automata in $\mathscr{C}$ and their morphisms.

Notation 3.2. We put $\delta_w := \delta_{a_n} \circ \cdots \circ \delta_{a_1}$ for $w = a_1 \dots a_n$ in Σ^*.

Example 3.3. (1) An automaton $D = (S, \delta, i, f)$ in **Set**, the category of sets and functions, with $X = 1$ and $Y = 2$, is precisely a classical *deterministic automaton.* It is called a *dfa* if S is finite. We identify the map $i\colon 1 \to S$ with an initial state $s_0 = i(*) \in S$, and the map $f\colon S \to 2$ with a set $F = f^{-1}[1] \subseteq S$ of final states. The language $L(D, s)$ *accepted* by a state $s \in S$ is the set of all words $w \in \Sigma^*$ such that $\delta_w(s) \in F$. The language $L(D)$ *accepted* by D is the language accepted by the state s_0.

(2) An automaton $N = (S, \delta, i, f)$ in **Rel**, the category of sets and relations, with $X = Y = 1$, is precisely a classical *nondeterministic automaton.* It is called an *nfa* if S is finite. We identify $i \subseteq 1 \times S$ with a set $I \subseteq S$ of initial states and $f \subseteq S \times 1$ with a set $F \subseteq S$ of final states. Thus, in our view an nfa may have multiple initial states. The language $L(N, R)$ *accepted* by a subset $R \subseteq S$ consists of all $w \in \Sigma^*$ such that $(r, s) \in \delta_w$ for some $r \in R$ and $s \in F$. The language $L(N)$ *accepted* by N is the language accepted by the set I.

(3) An automaton $A = (S, \delta, i, f)$ in **JSL** with $X = Y = 2$, shortly a **JSL**-*automaton*, is given by a semilattice S of states, a family $\delta = (\delta_a\colon S \to S)_{a \in \Sigma}$ of semilattice morphisms specifying transitions, an initial state $s_0 \in S$ (corresponding to $i\colon 2 \to S$), and a prime filter $F \subseteq S$ of final states (corresponding to $f\colon S \to 2$). It is called a **JSL**-*dfa* if S is finite. The language *accepted* by a state $s \in S$ or by the automaton A, resp., is defined as for deterministic automata.

Remark 3.4 (JSL-dfas vs. nfas). Dfas, nfas and **JSL**-dfas are expressively equivalent; they all accept precisely the regular languages. The interest of **JSL**-dfas is that they constitute an algebraic representation of nfas:

(1) Every **JSL**-dfa $A = (S, \delta, s_0, F)$ induces an equivalent nfa $J(A)$ on the set $J(S)$ of join-irreducibles of S. Given $s, t \in J(S)$ and $a \in \Sigma$, there is a transition $s \xrightarrow{a} t$ in $J(A)$ iff $t \leq \delta_a(s)$; the initial states are those $s \in J(S)$ with $s \leq s_0$, and the final states form the set $J(S) \cap F$.

(2) Conversely, for every nfa $N = (Q, \delta, I, F)$, the *subset construction* yields an equivalent **JSL**-dfa $\mathcal{P}(N)$ with states $\mathcal{P}(Q)$ (the $\cup$-semilattice of subsets of Q), transitions $\mathcal{P}\delta_a\colon \mathcal{P}(Q) \to \mathcal{P}(Q)$, $X \mapsto \delta_a[X]$, initial state $I \in \mathcal{P}(Q)$, and final states those subsets of Q containing some state from F. Note that $J(\mathcal{P}(Q)) \cong Q$.

It follows that the task of finding a state-minimal nfa for a given language is equivalent to finding a **JSL**-dfa with a minimum number of join-irreducibles [3]. This idea has recently been extended to a general coalgebraic framework [27,34].

Recall that the *minimal dfa* [5] for a regular language L, denoted by $\mathsf{dfa}(L)$, has states $\mathsf{LD}(L)$ (the set of left derivatives of L), transitions $K \xrightarrow{a} a^{-1}K$ for $K \in \mathsf{LD}(L)$ and $a \in \Sigma$, initial state $L = \varepsilon^{-1}L$, and final states those $K \in \mathsf{LD}(L)$ containing ε. Up to isomorphism, it can be characterized as the unique dfa accepting L that is *reachable* (i.e. every state is reachable from the initial state via transitions) and *simple* (i.e. any two distinct states accept distinct languages). We now develop the analogous concepts for **JSL**-automata; they are instances of the categorical theory of minimality due to Arbib and Manes [2] and Goguen [11]. Let us first observe that every language has two canonical infinite **JSL**-acceptors:

Definition 3.5. Let $L \subseteq \Sigma^*$ be a language.[1] SM Note!

(1) The *initial* **JSL**-*automaton* $\mathsf{Init}(L)$ for L has states $\mathcal{P}_\mathsf{f}(\Sigma^*)$ (the $\cup$-semilattice of finite subsets of Σ^*), initial state $\{\varepsilon\}$, final states all $X \in \mathcal{P}_\mathsf{f}(\Sigma^*)$ with $X \cap L \neq \emptyset$, and transitions $X \mapsto Xa = \{xa \;:\; x \in X\}$ for $X \in \mathcal{P}_\mathsf{f}(\Sigma^*)$ and $a \in \Sigma$.

(2) The *final* **JSL**-*automaton* $\mathsf{Fin}(L)$ for L has states $\mathcal{P}(\Sigma^*)$ (the $\cup$-semilattice of all languages), initial state L, final states all languages K containing ε, and transitions $K \mapsto a^{-1}K$ for $K \in \mathcal{P}(\Sigma^*)$ and $a \in \Sigma$.

As suggested by the terminology, these automata form the initial and the final object in the category of **JSL**-automata accepting L:

Lemma 3.6 [2,11]. *For every* **JSL**-*automaton* $A = (S, \delta, s_0, F)$ *accepting the language* $L \subseteq \Sigma^*$, *there exist unique* **JSL**-*automata morphisms*

$$e_A \colon \mathsf{Init}(L) \to A \qquad \text{and} \qquad m_A \colon A \to \mathsf{Fin}(L).$$

The map e_A *sends* $\{w_1, \ldots, w_n\} \in \mathcal{P}_\mathsf{f}(\Sigma^*)$ *to the state* $\bigvee_{i=1}^n \delta_{w_i}(s_0)$, *and the map* m_A *sends a state* $s \in S$ *to* $L(A, s)$, *the language accepted by* s.

Definition 3.7. A **JSL**-automaton $A = (S, \delta, s_0, F)$ is called

(1) *reachable* if the unique morphism $e_A \colon \mathsf{Init}(L) \to A$ is surjective, i.e. every state is of the form $\bigvee_{i=1}^n \delta_{w_i}(s_0)$ for some $w_1, \ldots, w_n \in \Sigma^*$;

(2) *simple* if the unique morphism $m_A \colon A \to \mathsf{Fin}(L)$ in injective, i.e. any two distinct states accept distinct languages;

(3) *minimal* if it is both reachable and simple.

Remark 3.8. (1) The category **Aut**(**JSL**) has a factorization system given by surjective and injective morphisms. Thus, for every **JSL**-automata morphism $h \colon (S, \delta, i, f) \to (S', \delta', i', f')$ with image factorization $h = (S \overset{e}{\twoheadrightarrow} S'' \overset{m}{\rightarrowtail} S')$ in **JSL**, there exists a unique **JSL**-automaton structure $(S'', \delta'', i'', f'')$ on S'' making both e and m automata morphisms. We call e the *coimage* and m the *image* of h. *Subautomata* and *quotient automata* of **JSL**-automata are represented by injective and surjective morphisms, respectively.

---

[1] SM Note: Notation $\mathcal{P}_\mathsf{f}$ was not introduced.

(2) Every **JSL**-automaton A has a unique reachable subautomaton $\mathsf{reach}(A) \rightarrowtail A$, the *reachable part* of A. It is the smallest subautomaton of A and arises as the image of the unique morphism $e_A\colon \mathsf{Init}(L) \to A$. Thus,

$$A \text{ is reachable} \quad \text{iff} \quad A \cong \mathsf{reach}(A) \quad \text{iff} \quad A \text{ has no proper subautomaton.}$$

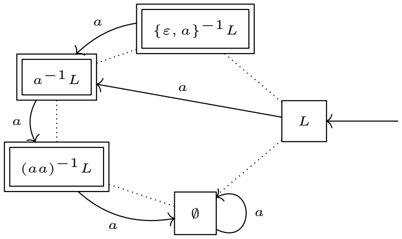
Let us emphasize that a state in $\mathsf{reach}(A)$ is not necessarily reachable when A is viewed as an ordinary dfa. For distinction, we thus call a state **JSL**-*reachable* if it lies in $\mathsf{reach}(A)$, and *dfa-reachable* if it is reachable in the usual sense.

(3) Dually, every **JSL**-automaton A has a unique simple quotient automaton $A \twoheadrightarrow \mathsf{simple}(A)$, the *simplification* of A. It is the smallest quotient automaton of A and arises as the coimage of the unique morphism $m_A\colon A \to \mathsf{Fin}(L)$. Thus,

$$A \text{ is simple} \quad \text{iff} \quad A \cong \mathsf{simple}(A) \quad \text{iff} \quad A \text{ has no proper quotient automaton.}$$

(4) Every language $L \subseteq \Sigma^*$ has a minimal **JSL**-automaton, unique up to isomorphism. It can be constructed as the image of the unique automata morphism $h_L\colon \mathsf{Init}(L) \to \mathsf{Fin}(L)$. Since h_L sends $\{w_1, \ldots, w_n\} \in \mathcal{P}_\mathsf{f}(\Sigma^*)$ to the language $\bigcup_{i=1}^n w_i^{-1}L$, the minimal automaton of L is the subautomaton $\mathsf{SLD}(L)$ of $\mathsf{Fin}(L)$ carried by the semilattice of finite unions of left derivatives of L.

Example 3.9. The minimal **JSL**-dfa accepting $L = \{a, aa\}$ is shown below, with the dashed lines representing the partial order.



Remark 3.10. The self-duality of $\mathbf{JSL_f}$ lifts to a self-duality of the category of **JSL**-dfas. The equivalence functor $\mathbf{Aut}(\mathbf{JSL_f}) \xrightarrow{\simeq} \mathbf{Aut}(\mathbf{JSL_f})^{\mathsf{op}}$ maps a **JSL**-dfa $A = (S,\, (\delta_a\colon S \to S)_{a\in\Sigma},\, i\colon 2 \to S,\, f\colon S \to 2)$ to its *dual automaton*

$$A^{\mathsf{op}} = (S^{\mathsf{op}},\, (\delta_a^*\colon S^{\mathsf{op}} \to S^{\mathsf{op}})_{a\in\Sigma},\, f^*: 2 \to S^{\mathsf{op}},\, i^*: S^{\mathsf{op}} \to 2),$$

using that $2^{\mathsf{op}} \cong 2$. Thus, the initial state of A^{op} is the $\leq_S$-largest non-final state of A, and its final states are those $s \in S$ with $s_0 \not\leq_S s$. Given $s, t \in S$ and $a \in \Sigma$, there is a transition $s \xrightarrow{a} t$ in A^{op} iff t is the $\leq_S$-largest state with $\delta_a(t) \leq_S s$.

The dualization of **JSL**-dfas can be seen as an algebraic generalization of the reversal operation on nfas. Recall that the *reverse* of an nfa N is the nfa N^r obtained by flipping all transitions and swapping initial and final states. If N accepts the language L, then N^r accepts the reverse language L^r.

Lemma 3.11. *For each nfa $N = (Q, \delta, I, F)$, we have the* **JSL**-*dfa isomorphism*

$$[\mathcal{P}(N)]^{\mathsf{op}} \xrightarrow{\cong} \mathcal{P}(N^\mathsf{r}), \qquad X \mapsto \overline{X} = Q \setminus X.$$

The following lemma summarizes some important properties of A^{op}:

Lemma 3.12. *Let $A = (S, \delta, i, f)$ be a* **JSL**-*dfa.*

(1) *For every $s \in S$, we have $L(A^{\mathsf{op}}, s) = \{\, w \in \Sigma^* \,:\, \delta_{w^{\mathsf{r}}}(s_0) \not\leq_S s \,\}$.*

(2) *If A accepts the language L, then A^{op} accepts the reverse language L^{r}.*

(3) *We have $[\mathsf{reach}(A)]^{\mathsf{op}} \cong \mathsf{simple}(A^{\mathsf{op}})$. Thus, A is reachable iff A^{op} is simple.*

Our next goal is to give, for every regular language L, dual characterizations of $\mathsf{SLD}(L)$, $\mathsf{BLD}(L)$ and $\mathsf{BLRD}(L)$, the **JSL**-subautomata of $\mathsf{Fin}(L)$ carried by all finite unions of left derivatives, boolean combinations of left derivatives and boolean combinations of two-sided derivatives, respectively. These results form the core of our duality-based approach to (sub-)atomic nfas in the next section. The minimal **JSL**-dfa $\mathsf{SLD}(L)$ admits the following dual description:

Proposition 3.13. *For every regular language L, the minimal* **JSL**-*dfas for L and L^{r} are dual. More precisely, we have the* **JSL**-*dfa isomorphism*

$$\mathrm{dr}_L\colon [\mathsf{SLD}(L^{\mathsf{r}})]^{\mathsf{op}} \xrightarrow{\cong} \mathsf{SLD}(L), \qquad K \mapsto (\overline{K^{\mathsf{r}}})^{-1}L.$$

Remark 3.14. (1) The isomorphism dr_L induces a bijection between the *left* and *right factors* of L, i.e. the inclusion-maximal left/right solutions of $X \cdot Y \subseteq L$. In fact, Conway [7] observed that the sets of left and right factors are given by $\{\overline{K^{\mathsf{r}}} \,:\, K \in \mathsf{SLD}(L^{\mathsf{r}})\}$ and $\{\overline{K} \,:\, K \in \mathsf{SLD}(L)\}$, respectively, and that these two sets are in bijective correspondence. Proposition 3.13 provides an explicit bijection and shows that it arises canonically via duality.

(2) The isomorphism dr_L is tightly connected to the *dependency relation* [14] of a regular language L, i.e. the binary relation given by

$$\mathcal{DR}_L \subseteq \mathsf{LD}(L) \times \mathsf{LD}(L^r), \qquad \mathcal{DR}_L(u^{-1}L, v^{-1}L^r) :\Longleftrightarrow\ uv^r \in L.$$

Its restriction $\mathcal{DR}_L^j := \mathcal{DR}_L \cap J(\mathsf{SLD}(L)) \times J(\mathsf{SLD}(L^r))$ to the $\cup$-irreducible left derivatives of L and L^{r} is called the *reduced dependency relation.* The following theorem shows that the semilattice of left quotients and the dependency relation are essentially the same concepts. In part (3), we use that the isomorphism dr_L restricts to a bijection between the $\cup$-irreducible derivatives of L^{r} and the meet-irreducible elements of the lattice $\mathsf{SLD}(L)$.

Theorem 3.15 (Dependency theorem).

(1) *We have the* **JSL**-*isomorphism*

$$\mathsf{SLD}(L) \xrightarrow{\cong} (\{\mathcal{DR}_L[X] : X \subseteq \mathsf{LD}(L)\}, \cup, \emptyset), \qquad K \mapsto \{v^{-1}L^{\mathsf{r}} : v \in K^{\mathsf{r}}\,\}.$$

Note that its codomain forms a subsemilattice of $\mathcal{P}(\mathsf{LD}(L^{\mathsf{r}}))$.

(2) *For all $u, v \in \Sigma^*$ we have $\mathcal{DR}_L(u^{-1}L, v^{-1}L^r) \iff u^{-1}L \nsubseteq \mathrm{dr}_L(v^{-1}L^r)$.*

(3) *The following diagram in* **Rel** *commutes:*

$$\begin{tikzcd}
J(\mathsf{SLD}(L^{\mathsf{r}})) \arrow[r, "\mathrm{dr}_L", "\cong"'] & M(\mathsf{SLD}(L)) \\
J(\mathsf{SLD}(L)) \arrow[u, "\mathcal{DR}_L^j"] \arrow[r, equal] & J(\mathsf{SLD}(L)) \arrow[u, "\nsubseteq"']
\end{tikzcd}$$

Let us now turn to a dual characterization of the **JSL**-dfa $\mathsf{BLD}(L)$:

Proposition 3.16. *For every regular language L, the* **JSL**-*dfa* $\mathsf{BLD}(L)$ *is dual to the subset construction of the minimal dfa for L^{r}:*

$$[\mathsf{BLD}(L)]^{\mathsf{op}} \;\cong\; \mathcal{P}(\mathsf{dfa}(L^{\mathsf{r}})).$$

The isomorphism maps $\{w_1^{-1}L^{\mathsf{r}}, \dots, w_n^{-1}L^{\mathsf{r}}\} \in \mathcal{P}(\mathsf{dfa}(L^{\mathsf{r}}))$ *to* $\bigcap_{i=1}^{n} \overline{\mathrm{At}(w_i^{\mathsf{r}})}$, *where* $\mathrm{At}(x)$ *is the unique atom (= join-irreducible) of* $\mathsf{BLD}(L)$ *containing x.*

To state the dual characterization of $\mathsf{BLRD}(L)$, we recall two standard concepts from algebraic language theory; for more details, see e.g. [28].

Remark 3.17. (1) The *transition monoid* of a deterministic automaton $D = (S, \delta, i, f)$ is the image $\mathsf{tm}(D) \subseteq \mathbf{Set}(S,S)$ of the monoid morphism

$$\Sigma^* \to \mathbf{Set}(S,S), \quad w \;\mapsto\; \delta_w.$$

Thus, $\mathsf{tm}(M)$ is carried by the set of extended transition maps δ_w ($w \in \Sigma^*$) with multiplication given by $\delta_v \bullet \delta_w = \delta_{vw}$ and unit $id_S = \delta_\varepsilon \colon S \to S$. We may view $\mathsf{tm}(D)$ itself as a deterministic automaton with initial state id_S, final states all δ_w such that w is accepted by D, and transitions given by $\delta_w \xrightarrow{a} \delta_{wa}$ for $w \in \Sigma^*$ and $a \in \Sigma$. Note that this automaton accepts the same language as D.

(2) The *syntactic monoid* $\mathsf{syn}(L)$ of a regular language $L \subseteq \Sigma^*$ is the transition monoid of its minimal dfa:

$$\mathsf{syn}(L) = \mathsf{tm}(\mathsf{dfa}(L)).$$

Equivalently, $\mathsf{syn}(L)$ is the quotient monoid of the free monoid Σ^* modulo the *syntactic congruence* of L, i.e the monoid congruence on Σ^* given by

$$v \equiv_L w \quad \text{iff} \quad \forall x, y \in \Sigma^* : xvy \in L \iff xwy \in L.$$

The associated surjective monoid morphism $\mu_L \colon \Sigma^* \twoheadrightarrow \mathsf{syn}(L)$, mapping $w \in \Sigma^*$ to its congruence class $[w]_L \in \mathsf{syn}(L)$, is called the *syntactic morphism.*

Proposition 3.18. *For every regular language L, the* **JSL**-*dfa* $\mathsf{BLRD}(L)$ *is dual to the subset construction of* $\mathsf{syn}(L^{\mathsf{r}})$, *viewed as a dfa:*

$$[\mathsf{BLRD}(L)]^{\mathsf{op}} \;\cong\; \mathcal{P}(\mathsf{syn}(L^{\mathsf{r}})).$$

The isomorphism maps $\{\, [w_1]_{L^{\mathsf{r}}}, \dots, [w_n]_{L^{\mathsf{r}}} \,\} \in \mathcal{P}(\mathsf{syn}(L^{\mathsf{r}}))$ *to* $\bigcap_{i=1}^{n} \overline{\mathrm{At}(w_i{}^{\mathsf{r}})}$, *with* $\mathrm{At}(x)$ *denoting the unique atom of* $\mathsf{BLRD}(L)$ *containing x.*

Our final duality result in this section concerns the *transition semiring* [30], a generalization of the transition monoid to **JSL**-automata. Note that the monoid $\mathbf{JSL}(S,S)$ of endomorphisms of a semilattice S forms an idempotent semiring with join defined pointwise: for any $f, g \colon S \to S$, the morphism $f \vee g \colon S \to S$

is given by $s \mapsto f(s) \vee g(s)$. The transition semiring of a **JSL**-automaton $A = (S, \delta, i, f)$ is the image $\mathsf{ts}(A) \subseteq \mathbf{JSL}(S, S)$ of the semiring morphism

$$\mathcal{P}_{\mathsf{f}}(\Sigma^*) \to \mathbf{JSL}(S, S), \quad \{w_1, \dots, w_n\} \;\mapsto\; \bigvee_{i=1}^{n} \delta_{w_i}.$$

Here $\mathcal{P}_{\mathsf{f}}(\Sigma^*)$ is the free idempotent semiring on Σ, with composition given by concatenation of languages and join given by union. Thus, $\mathsf{ts}(A)$ is the semiring carried by all morphisms $\bigvee_{i=1}^{n} \delta_{w_i}$ for $w_1, \dots, w_n \in \Sigma^*$, with join given as above and multiplication $\bigvee_j \delta_{v_j} \bullet \bigvee_i \delta_{w_i} = \bigvee_{i,j} \delta_{v_j w_i}$. We view $\mathsf{ts}(A)$ as a **JSL**-automaton with initial state $id_S = \delta_\varepsilon$, final states all $\bigvee_i \delta_{w_i}$ such that some w_i is accepted by A, and transitions $\bigvee_{i=1}^{n} \delta_{w_i} \xrightarrow{a} \bigvee_{i=1}^{n} \delta_{w_i a}$ for $w_1, \dots, w_n \in \Sigma^*$ and $a \in \Sigma$. This **JSL**-automaton is reachable and accepts the same language as A. It has the following dual characterization:

Notation 3.19. Given a simple **JSL**-automaton $A = (S, \delta, i, f)$, the subautomaton of $\mathsf{Fin}(L)$ obtained by closing S (viewed as a set of languages) under right derivatives is called the *right-derivative closure* of A and denoted $\mathsf{rdc}(A)$.

Proposition 3.20. *Let A be a reachable* **JSL**-*dfa. Then the transition semiring of A, viewed as a* **JSL**-*dfa, is dual to the right-derivative closure of A^{op}:*

$$[\mathsf{ts}(A)]^{\mathsf{op}} \cong \mathsf{rdc}(A^{\mathsf{op}}).$$

Note that both $[\mathsf{ts}(A)]^{\mathsf{op}}$ and $\mathsf{rdc}(A^{\mathsf{op}})$ are simple, hence subautomata of $\mathsf{Fin}(L)$. Thus, the isomorphism just expresses that their states accept the same languages.

4 Boolean Representations and Subatomic NFAs

Based upon the duality results of the previous section, we will now introduce our algebraic approach to nondeterministic state minimality. It rests on the concept of a representation of a monoid on a finite semilattice.

Definition 4.1 (Boolean representation). Let M be a monoid.

(1) A *boolean representation* of M is given by a finite semilattice S together with a monoid morphism $\rho\colon M \to \mathbf{JSL}(S, S)$. The *degree* of ρ is

$$\deg(\rho) := |J(S)|.$$

(2) Given boolean representations $\rho_i\colon M \to \mathbf{JSL}(S_i, S_i)$, $i = 1, 2$, an *equivariant map* $f\colon \rho_1 \to \rho_2$ is a **JSL**-morphism $f\colon S_1 \to S_2$ such that

$$f(\rho_1(m)(s)) = \rho_2(m)(f(s)) \text{ for all } m \in M \text{ and } s \in S_1.$$

If f is injective, we say that the representation ρ_2 *extends* ρ_1.

Remark 4.2. (1) The above representations are called *boolean* because semilattices are precisely semimodules over the boolean semiring $2 = \{0, 1\}$ with $1 + 1 = 1$. For more on representations over general semirings, see [16].

(2) The category of boolean representations of M coincides with the functor category $\mathbf{JSL}_\mathsf{f}^M$, viewing M as a one object category.

Definition 4.3 (Canonical representation). For every regular language L, the *canonical boolean representation* of the syntactic monoid $\mathsf{syn}(L)$ is given by

$$\kappa_L \colon \mathsf{syn}(L) \to \mathbf{JSL}(\mathsf{SLD}(L), \mathsf{SLD}(L)), \quad [w]_L \mapsto \lambda K.w^{-1}K.$$

It induces the *canonical boolean presentation* of the free monoid Σ^* given by

$$\kappa_L \circ \mu_L \colon \Sigma^* \to \mathbf{JSL}(\mathsf{SLD}(L), \mathsf{SLD}(L)), \quad w \mapsto \lambda K.w^{-1}K,$$

where $\mu_L \colon \Sigma^* \twoheadrightarrow \mathsf{syn}(L)$ is the syntactic morphism.

Example 4.4. We describe the canonical boolean representation κ_{L_n} for the language $L_n := (0+1)^*1(0+1)^n$, $n \in \mathbb{N}$. Let $S := 2_\bot^{n+1}$ be the semilattice of binary words of length $n+1$, ordered pointwise, with an additional bottom element $\bot$. Then $\mathsf{SLD}(L_n)$ is isomorphic to S, as witnessed by the isomorphism

$$f \colon S \xrightarrow{\cong} \mathsf{SLD}(L_n), \quad f(\bot) = \emptyset, \quad f(w) = w^{-1}L_n.$$

Thus, κ_{L_n} is isomorphic to the representation $\rho \colon \mathsf{syn}(L_n) \to \mathbf{JSL}(S, S)$ where:

(1) $\rho([0]_{L_n}) \colon S \to S$ performs a left-shift (distinct from left-rotate);

(2) $\rho([1]_{L_n}) \colon S \to S$ performs a left-shift and sets the last bit as 1.

Finally, $\deg(\kappa_{L_n}) = \deg(\rho) = 1 + |J(2^{n+1})| = n+2$ is the number of states of the usual minimal nfa for L.

Example 4.5. We describe the canonical boolean presentation κ_L for the language $L = a_1(a_2+a_3)+a_2(a_1+a_3)+a_3(a_1+a_2)$ over $\Sigma = \{a_1, a_2, a_3\}$. Consider the $\cup$-semilattice $M_3 = \{\emptyset, \{a_1,a_2\}, \{a_1,a_3\}, \{a_2,a_3\}, \Sigma\}$. Then $\mathsf{SLD}(L)$ is isomorphic to the product semilattice $2 \times M_3 \times 2$: the map[2] SM Note!

$$f \colon \mathsf{SLD}(L) \xrightarrow{\cong} 2 \times M_3 \times 2, \quad f(X) = (X \cap \Sigma^2, X \cap \Sigma, X \cap \{\varepsilon\}).$$

is an isomorphism. For $i = 1,2,3$ we define the following semilattice morphisms:

$$\begin{aligned} &\alpha_i \colon 2 \to M_3, && \alpha_i(1) = \Sigma \setminus \{a_i\}; \\ &\beta_i \colon M_3 \to 2, && \beta_i(S) = 1 \iff a_i \in S; \\ &\gamma \colon 2 \to 2 && \gamma(1) = 0; \\ &\delta \colon M_3 \times 2 \times 2 \to 2 \times M_3 \times 2, && \delta(x,y,z) = (z,x,y). \end{aligned}$$

Then κ_L is isomorphic to $\rho \colon \mathsf{syn}(L) \to \mathbf{JSL}(2 \times M_3 \times 2,\ 2 \times M_3 \times 2)$ where

$$\rho([a_i]_L) = (\,2 \times M_3 \times 2 \xrightarrow{\alpha_i \times \beta_i \times \gamma} M_3 \times 2 \times 2 \xrightarrow{\delta} 2 \times M_3 \times 2\,).$$

Thus, $\deg(\kappa_L) = \deg(\rho) = 1+3+1 = 5$. An analogous description of κ_L exists for any language L where each word has the same length.

---

[2] SM Note: It needs to be explained why the 1st and 3rd component are in 2 to make it easier for the reader.

The next theorem links minimal nfas and representations.

Definition 4.6. The *nondeterministic state complexity* $\mathrm{ns}(L)$ of a regular language L is the least number of states of any nfa accepting L.

Theorem 4.7. *For every regular language L, the nondeterministic state complexity* $\mathrm{ns}(L)$ *is the least degree of any boolean representation extending the canonical representation* $\kappa_L \circ \mu_L \colon \Sigma^* \to \mathbf{JSL}(\mathsf{SLD}(L), \mathsf{SLD}(L))$.

Proof (Sketch).

(1) Given a k-state nfa $N = (Q, \delta, I, F)$ accepting L, consider the subsemilattice $\mathsf{langs}(N) = \mathsf{simple}(\mathcal{P}(N))$ of $\mathcal{P}(\Sigma^*)$ on all languages accepted by subsets of Q. The embedding $\mathsf{SLD}(L) \rightarrowtail \mathsf{langs}(N)$ yields an extension of $\kappa_L \circ \mu_L$. Since the semilattice $\mathsf{langs}(N)$ is generated by the languages accepted by single states of N, this extension has degree at most k.

(2) Conversely, let $\rho\colon \Sigma^* \to \mathbf{JSL}(S, S)$ be a boolean representation of degree k extending $\kappa_L \circ \mu_L$, witnessed by an injective equivariant map $h\colon \mathsf{SLD}(L) \rightarrowtail S$. One can equip S with a **JSL**-dfa structure making h an automata morphism. Since morphisms preserve accepted languages, it follows that S accepts L. Then the nfa of join-irreducibles of S, see Remark 3.4, is a k-state nfa accepting L. □

As an application, let us return to the dependency relation $\mathcal{DR}_L$ introduced in Remark 3.14(2). Recall that a *biclique* of a relation $R \subseteq X \times Y$ (viewed as a bipartite graph) is a subset of the form $X' \times Y' \subseteq R$, where $X' \subseteq X$ and $Y' \subseteq Y$. A *biclique cover* of R is a set $\mathscr{C}$ of bicliques with $R = \bigcup \mathscr{C}$. The *bipartite dimension* $\dim(R)$ is the least cardinality of any biclique cover of R.

Theorem 4.8 (Gruber-Holzer [14]). *For every regular language L, we have*

$$\dim(\mathcal{DR}_L) \leq \mathrm{ns}(L).$$

We give a new algebraic proof of this result based on boolean representations.

Proof. (1) The task of computing biclique covers is well-known to be equivalent to the *set basis* problem. Given a family $C \subseteq \mathcal{P}(Y)$ of subsets of a finite set Y, a set basis for C is a family $B \subseteq \mathcal{P}(Y)$ such that each element of C can be expressed as a union of elements of B. A relation $R \subseteq X \times Y$ has a biclique cover of size k iff the family $C_R = \{R[x] : x \in X\} \subseteq \mathcal{P}(Y)$ of neighborhoods of nodes in X has a set basis of size k.

(2) Given an instance $C \subseteq \mathcal{P}(Y)$ of the set basis problem, consider the $\cup$-subsemilattice $\langle C \rangle \subseteq \mathcal{P}(Y)$ generated by C, i.e. the semilattice of all unions of sets in C. We claim that C has a set basis of size at most k iff there exists an extension of $\langle C \rangle$ of degree at most k, i.e. a monomorphism $\langle C \rangle \rightarrowtail S$ into some finite semilattice S with $|J(S)| \leq k$.

For the "only if" direction, suppose that $B \subseteq \mathcal{P}(Y)$ is a set basis of C of size at most k. The the embedding $\langle C \rangle \rightarrowtail \langle B \rangle$ gives an extension of $\langle C \rangle$ with the desired property: since the semilattice $\langle B\rangle$ has a set of generators with at most k elements, it has at most k join-irreducibles.

For the "if" direction, suppose that $m\colon \langle C\rangle \rightarrowtail S$ with $|J(S)| \leq k$ is given. Since the free semilattice $\mathcal{P}(Y)$ is an injective object of **JSL** [15, Corollary 2.9], there exists a morphism $f\colon S \to \mathcal{P}(Y)$ extending the embedding $\langle C\rangle \rightarrowtail \mathcal{P}(Y)$. Consider the image $S' \subseteq \mathcal{P}(Y)$ of f, leading to the commutative diagram below:

$$\begin{tikzcd}
\langle C\rangle \arrow[r, tail, "m"] \arrow[dr, tail, "\subseteq"'] & S \arrow[r, two heads, "e"] \arrow[d, "f"] & S' \arrow[dl, tail, "\subseteq"] \\
& \mathcal{P}(Y) &
\end{tikzcd}$$

We thus have $\langle C\rangle \subseteq S' \subseteq \mathcal{P}(Y)$. Every set of generators of the semilattice S' is a basis of C. Since the morphism e is surjective, we have $|J(S')| \leq |J(S)| \leq k$, i.e. S' has a set of generators with at most k elements.

(3) Let $C_{\mathcal{DR}_L} \subseteq \mathcal{P}(\mathsf{LD}(L^{\mathsf{r}}))$ be the instance of the set basis problem corresponding to the dependency relation $\mathcal{DR}_L \subseteq \mathsf{LD}(L) \times \mathsf{LD}(L^{\mathsf{r}})$. Note that $\langle C_{\mathcal{DR}_L}\rangle$ consists of all $\mathcal{DR}_L[X]$ for $X \subseteq \mathsf{LD}(L)$. Thus, Theorem 3.15(1) shows that $\langle C_{\mathcal{DR}_L}\rangle \cong \mathsf{SLD}(L)$. In particular, every extension of the canonical boolean representation of Σ^* yields an extension of the semilattice $\langle C_{\mathcal{DR}_L}\rangle$ of the same degree. Therefore, by part (1) and (2) and Theorem 4.7, we have $\dim(\mathcal{DR}_L) \leq \mathrm{ns}(L)$, as required.

Theorem 4.7 motivates the following definition, which can be considered the key concept of our paper:

Definition 4.9. The *nondeterministic syntactic complexity* $\mathrm{n}\mu(L)$ of a regular language L is the least degree of any boolean representation of $\mathsf{syn}(L)$ extending the canonical boolean representation $\kappa_L\colon \mathsf{syn}(L) \to \mathbf{JSL}(\mathsf{SLD}(L), \mathsf{SLD}(L))$.

Just like the degrees of boolean representations of Σ^* determine the state complexity of nfas, we will provide an automata-theoretic characterization of $\mathrm{n}\mu(L)$ in terms of *subatomic* nfas in Theorem 4.14 below.

Definition 4.10. An nfa accepting the language L is called

(1) *atomic* if each state accepts a language from $\mathsf{BLD}(L)$, and

(2) *subatomic* if each state accepts a language from $\mathsf{BLRD}(L)$.

The notion of an atomic nfa goes back to Brzozowski and Tamm [4], as does the following characterization.

Notation 4.11. For any nfa N, let $\mathsf{rsc}(N)$ denote the dfa obtained via the *reachable subset construction*, i.e. the dfa-reachable part of $\mathcal{P}(N)$.

Theorem 4.12. *An nfa N is atomic iff $\mathsf{rsc}(N^{\mathsf{r}})$ is a minimal dfa.*

We present a new conceptual proof, interpreting this theorem as an instance of the self-duality of **JSL**-dfas.

Proof (Sketch). Let L be the language accepted by N. We establish the theorem by showing each of the following statements to be equivalent to the next one:

(1) N is atomic.

(2) There exists a **JSL**-automata morphism from $\mathcal{P}(N)$ to $\mathsf{BLD}(L)$.

(3) There exists a **JSL**-automata morphism from $\mathsf{simple}(\mathcal{P}(N))$ to $\mathsf{BLD}(L)$.

(4) There exists a **JSL**-automata morphism from $\mathcal{P}(\mathsf{dfa}(L^{\mathsf{r}}))$ to $\mathsf{reach}(\mathcal{P}(N^{\mathsf{r}}))$.

(5) There exists a dfa morphism from $\mathsf{dfa}(L^{\mathsf{r}})$ to $\mathsf{reach}(\mathcal{P}(N^{\mathsf{r}}))$.

(6) There exists a dfa morphism from $\mathsf{dfa}(L^{\mathsf{r}})$ to $\mathsf{rsc}(N^{\mathsf{r}})$.

(7) $\mathsf{rsc}(N^{\mathsf{r}})$ is a minimal dfa.

The key step is (3)⇔(4), which follows via duality from Lemmas 3.11 and 3.12, and Proposition 3.16. All remaining equivalences follow from the definitions. □

The next theorem gives an analogous characterization of subatomic nfas. Again, the proof is based on duality.

Theorem 4.13. *An nfa N accepting the language L is subatomic iff the transition monoid of* $\mathsf{rsc}(N^{\mathsf{r}})$ *is isomorphic to the syntactic monoid* $\mathsf{syn}(L^{\mathsf{r}})$.

Proof (Sketch). Each of the following statements is equivalent to the next one:

(1) N is subatomic.

(2) There exists a **JSL**-dfa morphism from $\mathcal{P}(N)$ to $\mathsf{BLRD}(L)$.

(3) There exists a **JSL**-dfa morphism from $\mathsf{rdc}(\mathsf{simple}(P(N)))$ to $\mathsf{BLRD}(L)$.

(4) There exists a **JSL**-dfa morphism from $\mathcal{P}(\mathsf{syn}(L^{\mathsf{r}}))$ to $\mathsf{ts}(\mathsf{reach}(\mathcal{P}(N^{\mathsf{r}})))$.

(5) There exists a dfa morphism from $\mathsf{syn}(L^{\mathsf{r}})$ to $\mathsf{ts}(\mathsf{reach}(\mathcal{P}(N^{\mathsf{r}})))$.

(6) There exists a dfa morphism from $\mathsf{syn}(L^{\mathsf{r}})$ to $\mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$.

(7) The monoids $\mathsf{syn}(L^{\mathsf{r}})$ and $\mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$ are isomorphic.

The equivalence (3)⇔(4) follows via duality from Lemma 3.11, Proposition 3.18 and Proposition 3.20. All remaining equivalences follow from the definitions. □

We are prepared to state the main result of our paper, an automata-theoretic characterization of the nondeterministic syntactic complexity:

Theorem 4.14. *For every regular language L, the nondeterministic syntactic complexity* $\mathrm{n}\mu(L)$ *is the least number of states of any subatomic nfa accepting L.*

Proof (Sketch).

(1) Let N be a k-state subatomic nfa accepting the language L. As in the proof of Theorem 4.7, we consider the semilattice $\mathsf{langs}(N) = \mathsf{simple}(\mathcal{P}(N))$. Then

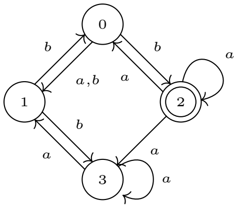
$$\rho\colon \mathsf{syn}(L) \to \mathbf{JSL}(\mathsf{langs}(N), \mathsf{langs}(N)), \quad [w]_L \mapsto \lambda K.w^{-1}K,$$

is a representation of $\mathsf{syn}(L)$ of degree at most k extending κ_L.

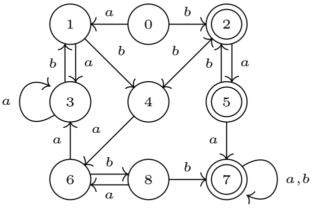
(2) Conversely, let $\rho\colon \mathsf{syn}(L) \to \mathbf{JSL}(S, S)$ be a boolean representation extending κ_L, and let $h\colon \mathsf{SLD}(Q) \rightarrowtail S$ be the embedding. As in the proof of Theorem 4.7, we can equip S with the structure of a **JSL**-dfa making h an automata morphism. Its nfa of join-irreducibles, see Remark 3.4, is a subatomic nfa accepting L with $\deg(\rho)$ states. ☐

We conclude this section with the observation that the state complexity of unrestricted nfas, subatomic nfas and atomic nfas generally differs:

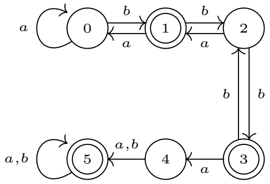
Example 4.15 (Subatomic more succinct than atomic). Consider the language L accepted by the nfa N shown below, along with the minimal dfas for L and L^{r}. All three automata have the initial state 0.



N



$\mathsf{dfa}(L)$



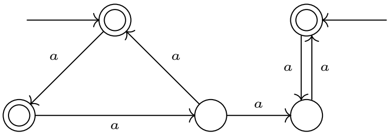
$\mathsf{dfa}(L^{\mathsf{r}})$

Brzozowski and Tamm [4] showed that there is no atomic nfa with four states accepting L. However, N is subatomic: one can verify that the transition monoids of $\mathsf{dfa}(L^{\mathsf{r}})$ and $\mathsf{rsc}(N^{\mathsf{r}})$ both have 22 elements. Since the former is the syntactic monoid of L^{r}, they are isomorphic, and so Theorem 4.13 applies.

Example 4.16 (Subatomic less succinct than general nfas). There is a regular language for which no state-minimal nfa is subatomic:

$$L \;:=\; \{\, a^n \;:\; n \in \mathbb{N},\, n \neq 5 \,\} \subseteq \{a\}^*.$$

It is accepted by the following nfa:



An exhaustive search shows that no subatomic nfa with five states accepts L. In fact, L is the unique (!) unary language with $\mathrm{ns}(L) \leq 5$ and $\mathrm{ns}(L) < \mathrm{n}\mu(L)$. Moreover, the above nfa and its reverse are the only state-minimal nfas for L.

5 Applications

While subatomic nfas are generally less succinct then unrestricted ones, all structural results concerning nondeterministic state complexity we have encountered in the literature are actually about nondeterministic syntactic complexity: they implicitly identify classes of languages where the two measures coincide. In the present section, we illustrate this in a few selected applications.

5.1 Unary languages

For unary languages $L \subseteq \{a\}^*$, two-sided derivatives are left derivatives. Thus, a unary nfa is atomic iff it is subatomic.

Example 5.1 (Cyclic unary languages). A unary language L is *cyclic* if its minimal dfa is a cycle [12]. We claim that

$$\mathrm{ns}(L) = \mathrm{n}\mu(L).$$

To see this, let $d := |\mathsf{LD}(L)|$ be the *period* (i.e. number of states) of the minimal dfa. By Fact 1 of [12] (originally from [17]) every state-minimal nfa N accepting L is a disjoint union of cyclic dfas whose periods divide d.[3] Then $|\mathsf{rsc}(N^{\mathsf{r}})| = d$:

- $|\mathsf{rsc}(N^{\mathsf{r}})| \geq d$ since $\mathsf{rsc}(N^{\mathsf{r}})$ is a dfa accepting $L = L^{\mathsf{r}}$ and d is the size of the minimal dfa for L.
- $|\mathsf{rsc}(N^{\mathsf{r}})| \leq d$ because after d steps, each cycle will be back in its initial state.

Thus N is atomic by Theorem 4.12 and hence subatomic.

As an application, we obtain the following result for (not necessarily unary) languages whose syntactic monoid is a cyclic group:

Theorem 5.2. *If L is a regular language such that $\mathsf{syn}(L)$ is cyclic, then*

$$\mathrm{ns}(L) = \mathrm{n}\mu(L).$$

Proof (Sketch). Suppose that $\mathsf{syn}(L) = \mathsf{tm}(\mathsf{dfa}(L))$ is cyclic. Then there exists $w_0 \in \Sigma^*$ such that the map $\lambda X.w_0^{-1}X \colon \mathsf{LD}(L) \to \mathsf{LD}(L)$ generates $\mathsf{tm}(\mathsf{dfa}(L))$. Fix an alphabet $\Sigma_0 = \{a_0\}$ disjoint from Σ and consider the unary language

$$L_0 := \{\, a_0^n \;:\; n \in \mathbb{N},\, w_0^n \in L \,\} \subseteq \Sigma_0^*.$$

Let $g : \Sigma_0^* \to \Sigma^*$ be the monoid morphism where $g(a_0) := w_0$. Then we have the **JSL**-isomorphism

$$f \colon \mathsf{SLD}(L_0) \xrightarrow{\cong} \mathsf{SLD}(L), \quad f(X^{-1}L_0) := [g[X]]^{-1}L.$$

For each $a \in \Sigma$ choose $n_a \in \mathbb{N}$ such that $a^{-1}K = (w_0^{n_a})^{-1}K$ for all $K \in \mathsf{LD}(L)$. The respective transition endomorphisms of the **JSL**-automata $\mathsf{SLD}(L_0)$ and $\mathsf{SLD}(L)$ determine each other in the sense that the following diagrams commute:

$$\begin{tikzcd}
\mathsf{SLD}(L_0) \arrow[r, "f", "\cong"'] \arrow[d, "a_0^{-1}(-)"'] & \mathsf{SLD}(L) \arrow[d, "w_0^{-1}(-)"] \\
\mathsf{SLD}(L_0) \arrow[r, "\cong", "f"'] & \mathsf{SLD}(L)
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathsf{SLD}(L_0) \arrow[r, "f", "\cong"'] \arrow[d, "(a_0^{n_a})^{-1}(-)"'] & \mathsf{SLD}(L) \arrow[d, "a^{-1}(-)"] \\
\mathsf{SLD}(L_0) \arrow[r, "\cong", "f"'] & \mathsf{SLD}(L)
\end{tikzcd}$$

Then $\mathrm{ns}(L) = \mathrm{ns}(L_0)$ by Theorem 4.7 and $\mathrm{n}\mu(L) = \mathrm{n}\mu(L_0)$ by Theorem 4.14. Moreover, by Example 5.1 we know that $\mathrm{ns}(L_0) = \mathrm{n}\mu(L_0)$, so the claim follows.

---

[3] In [12] nfas are restricted to have a single initial state and so are distinguished from unions of dfas; the latter are valid nfas from our perspective.

Example 5.3 ($\mathrm{n}\mu(L)$ no larger than Chrobak normal form). A unary nfa is in *Chrobak normal form* [6, 9] if it has a single initial state and at most one state with multiple successors, all of which lie in disjoint cycles. We claim that for any nfa N in Chrobak normal form accepting the language L, we have

$$\mathrm{n}\mu(L) \leq |N|,$$

where $|N|$ denotes the number of states of N. To see this, observe that each state of N up to and including the single nondeterministic choice accepts some $X \in \mathsf{LD}(L)$. The state where the nondeterministic choice occurs accepts a cyclic unary language $u^{-1}L$. By Example 5.1, we may replace this state and all subsequent ones by an atomic nfa accepting $u^{-1}L$, without increasing the number of states. The resulting nfa is atomic.

Since every unary nfa on n states can be transformed into an nfa in Chrobak normal form with $O(n^2)$ states [6, Lemma 4.3], we get:

Corollary 5.4. *If L is a unary regular language, then* $\mathrm{n}\mu(L) = O(\mathrm{ns}(L)^2)$.

5.2 Languages with a canonical state-minimal nfa

There are several natural classes of regular languages for which *canonical* state-minimal nondeterministic acceptors have been identified. We show that these acceptors are actually subatomic. In our arguments, we frequently consider the *length* of a finite semilattice S, i.e. the maximum length n of any ascending chain $s_0 < s_1 < \ldots < s_n$ in S. Note that since every element is uniquely determined by the set of join-irreducibles below it, the length of S is at most $|J(S)|$.

Example 5.5 (Bideterministic and biseparable languages).

(1) A language is called *bideterministic* if it is accepted by a dfa whose reverse is also a dfa. In this case, the minimal dfa is a minimal nfa [29, 33]. Bideterministic languages of the form $L \subseteq \{0,1\}^n$ are precisely the well-studied *rectangular codes* [22, 31]. We show that for every bideterministic language L,

$$\mathrm{ns}(L) = \mathrm{n}\mu(L) = |\mathsf{LD}(L)|.$$

To this end, we first note that by [31, Theorem 3.1] a language $L \subseteq \Sigma^*$ is bideterministic iff the left derivatives of L are pairwise disjoint. This implies that $\mathsf{SLD}(L)$ is a boolean algebra with atoms $\mathsf{LD}(L)$. Since the length of a boolean algebra equals the number of atoms (= join-irreducibles), we conclude that for every finite semilattice extension $\mathsf{SLD}(L) \rightarrowtail S$, the semilattice S has length at least $|\mathsf{LD}(L)|$. Thus, $|\mathsf{LD}(L)| \leq |J(S)|$, so any representation ρ extending κ_L or $\kappa_L \circ \mu_L$ satisfies $|\mathsf{LD}(L)| \leq \deg(\rho)$. Hence, $\mathrm{ns}(L) = \mathrm{n}\mu(L) = |\mathsf{LD}(L)|$ by Theorem 4.7 and 4.14. In particular, the minimal dfa of L is a minimal nfa.

(2) A language L is *biseparable* if $\mathsf{SLD}(L)$ is a boolean algebra [23].[4] For every biseparable language L, the *canonical residual automaton* [8], i.e. the nfa N_L

---

[4] Actually [23] defines biseparability as a property of nfas, and characterizes biseparable nfas as those accepting a language L for which no $\cup$-irreducible left derivative is contained in the union of other $\cup$-irreducible left derivatives. This is equivalent to the lattice $\mathsf{SLD}(L)$ being boolean, i.e. to L being 'biseparable' in our sense.

of join-irreducibles of the minimal **JSL**-dfa $\mathsf{SLD}(L)$, is a state-minimal nfa; it is subatomic because every state of N_L accepts a derivative of L. This follows exactly as in (1): our argument only used that $\mathsf{SLD}(L)$ is a boolean algebra.

Example 5.6 (Maximal reachability). A folklore result asserts that if N is an nfa whose accepted language L satisfies $|\mathsf{LD}(L)| = 2^{|N|}$, then N is state-minimal. Since $\mathsf{LD}(L)$ forms the set of states of the minimal dfa for L and $\mathsf{rsc}(N)$ accepts L, we have $\mathsf{rsc}(N) = \mathcal{P}(N)$. It follows the **JSL**-dfa $\mathcal{P}(N)$ is reachable and simple, hence isomorphic to the minimal **JSL**-dfa $\mathsf{SLD}(L)$. This proves that $\mathsf{SLD}(L)$ is a boolean algebra, i.e. L is a biseparable language. We conclude from Example 5.5(2) that $\mathrm{ns}(L) = \mathrm{n}\mu(L) = |N|$ and N_L is a subatomic minimal nfa.

Example 5.7 (BiRFSA and topological languages). So far $\mathsf{SLD}(L)$ has been a boolean algebra. But the argument in Example 5.5 also applies when $\mathsf{SLD}(L)$ is a distributive lattice, noting that the length of a finite distributive lattice is equal to the number of its join-irreducibles [13, Corollary 2.14]. Languages with this property are called *topological* [1]. It thus follows as in Example 5.5(2) that for any topological language L, the canonical residual automaton N_L is subatomic and a state-minimal nfa. Thus, $\mathrm{ns}(L) = \mathrm{n}\mu(L) = |J(\mathsf{SLD}(L))|$.

There is another class of languages where N_L is known to be a state-minimal nfa, the *biRFSA* languages [23]. A language L is called biRFSA if N_L is isomorphic to $(N_{L^\mathsf{r}})^\mathsf{r}$. Surprisingly, these languages are exactly the topological ones:

(1) *Suppose that L is topological.* Recall that N_L is the nfa of join-irreducibles of the minimal **JSL**-dfa. Thus, it has states $J(\mathsf{SLD}(L))$ and transitions given by $X \xrightarrow{a} Y$ iff $Y \subseteq a^{-1}X$ for $a \in \Sigma$. Moreover, a join-irreducible j is initial iff $j \subseteq L$ and final iff $\varepsilon \in j$. Since the lattice $\mathsf{SLD}(L)$ is distributive, we have a canonical bijection between its join- and meet-irreducibles:

$$\tau\colon J(\mathsf{SLD}(L)) \xrightarrow{\cong} M(\mathsf{SLD}(L)), \quad \tau(j) = \bigcup\{X \in \mathsf{SLD}(L) : j \nsubseteq X\}.$$

Let θ be the unique map making the following diagram commute, where dr_L is the restriction of the isomorphism of Proposition 3.13:

$$\begin{tikzcd} & J(\mathsf{SLD}(L)) \arrow[dl, "\theta"', "\cong"] \arrow[dr, "\tau", "\cong"'] & \\ J(\mathsf{SLD}(L^\mathsf{r})) \arrow[rr, "\cong", "\mathrm{dr}_L"'] & & M(\mathsf{SLD}(L)) \end{tikzcd}$$

One can show θ to be an nfa isomorphism from N_L to $(N_{L^\mathsf{r}})^\mathsf{r}$. Thus, L is biRFSA.

(2) *Suppose that L is biRFSA.* Then we have a surjective **JSL**-morphism

$$[\mathcal{P}(J(\mathsf{SLD}(L)))]^{\mathsf{op}} \cong \mathcal{P}(J(\mathsf{SLD}(L^\mathsf{r}))) \xrightarrow{e_{L^\mathsf{r}}} \mathsf{SLD}(L^\mathsf{r}) \cong [\mathsf{SLD}(L)]^{\mathsf{op}},$$

where the first isomorphism follows from $N_L \cong (N_{L^\mathsf{r}})^\mathsf{r}$ and Lemma 3.11, the second isomorphism is given by Proposition 3.13, and e_{L^r} sends $X \subseteq J(\mathsf{SLD}(L^\mathsf{r}))$ to $\bigcup X$. The dual of this morphism is the injective **JSL**-morphism

$$m_L\colon \mathsf{SLD}(L) \rightarrowtail \mathcal{P}(J(\mathsf{SLD}(L)))$$

sending $K \in \mathsf{SLD}(L)$ to the set of all $j \in J(\mathsf{SLD}(L))$ with $j \subseteq K$. Note that $e_L \circ m_L = id_{\mathsf{SLD}(Q)}$, showing that $\mathsf{SLD}(L)$ is a retract of $\mathcal{P}(J(\mathsf{SLD}(L)))$. Since **JSL**-retracts of finite distributive lattices are distributive, see e.g. [26, Lemma 2.2.3.15], it follows that $\mathsf{SLD}(L)$ is distributive. Thus, L is topological.

Example 5.8 (Extremal languages). Call a language *extremal* if $\mathsf{SLD}(L)$ has length $|J(\mathsf{SLD}(L))|$ i.e. we have an *extremal lattice* in the sense of Markowsky [24]. Again, the argument of Example 5.5 applies and we get $\mathrm{ns}(L) = \mathrm{n}\mu(L) = |J(\mathsf{SLD}(L))|$. Topological languages are extremal since every distributive lattice is an extremal lattice, although extremal languages need not be topological. Both classes are naturally characterized in terms of the reduced dependency relation:

(1) L is topological iff $\mathcal{DR}_L^j$ is essentially an order relation $\leq_P \subseteq P \times P$ of a finite poset [25, Example 2.2.12].

(2) L is extremal iff $\mathcal{DR}_L^j$ is *upper triangularizable* [24, Theorem 11].

The latter means the adjacency matrix of the bipartite graph $\mathcal{DR}_L^j$ can be put in upper-triangular form by permuting rows and columns. An order relation is upper triangularizable because it may be extended to a linear order.

6 Conclusion and Future Work

Motivated by the duality theory of deterministic finite automata over semilattices, we introduced a natural class of nondeterministic finite automata called *subatomic nfas* and studied their state complexity in terms of boolean representations of syntactic monoids. Furthermore, we demonstrated that a large body of previous work on state minimization of general nfas actually constructs minimal subatomic ones. There are several directions for future work.

As illustrated by Theorem 4.8, the dependency relation $\mathcal{DR}_L$ forms a useful tool for proving lower bounds on nfas. It is also a key element of the Kameda-Weiner algorithm [21,32] for minimizing nfas, which rests on computing biclique covers of $\mathcal{DR}_L$. We aim to give an algebraic interpretation of dependency relations based on the representation of finite semilattices by contexts [19], which can be augmented to a categorical equivalence between $\mathbf{JSL}_\mathsf{f}$ and a suitable category of bipartite graphs [26]. Under this equivalence, **JSL**-dfas correspond to *dependency automata*; in particular, the minimal **JSL**-dfa $\mathsf{SLD}(L)$ corresponds to a dependency automaton whose underlying bipartite graph is precisely the dependency relation $\mathcal{DR}_L$. We expect that this observation can lead to a fresh algebraic perspective on the Kameda-Weiner algorithm, as well as a generalization of it computing minimal (sub-)atomic nfas.

On a related note, we also intend to investigate the complexity of the minimization problem for (sub-)atomic nfas. While minimizing general nfas is PSPACE-complete, even if the input automaton is a dfa, we conjecture that the additional structure present in (sub-)atomic acceptors will simplify their minimization to an NP-complete task. First evidence in this direction is provided by the work of Geldenhuys, van der Merve, and van Zijl [10] who showed that minimal atomic nfas can be efficiently computed in practice using SAT solvers.

References

1. Adámek, J., Myers, R.S., Urbat, H., Milius, S.: On continuous nondeterminism and state minimality. In: Proc. 30th Conference on the Mathematical Foundations of Programming Semantics (MFPS XXX). vol. 308, pp. 3 – 23 (2014)
2. Arbib, M.A., Manes, E.G.: Adjoint machines, state-behavior machines, and duality. Journal of Pure and Applied Algebra 6(3), 313 – 344 (1975)
3. Arbib, M.A., Manes, E.G.: Fuzzy machines in a category. Bulletin of the Australian Mathematical Society 13(2), 169–210 (1975)
4. Brzozowski, J., Tamm, H.: Theory of átomata. Theoretical Computer Science 539, 13–27 (2014)
5. Brzozowski, J.A.: Derivatives of regular expressions. J. ACM 11(4), 481–494 (Oct 1964)
6. Chrobak, M.: Finite automata and unary languages. Theoretical Computer Science 47, 149–158 (1986)
7. Conway, J.H.: Regular Algebra and Finite Machines. Printed in GB by William Clowes & Sons Ltd (1971)
8. Denis, F., Lemay, A., Terlutte, A.: Residual finite state automata. In: Ferreira, A., Reichel, H. (eds.) STACS 2001: 18th Annual Symposium on Theoretical Aspects of Computer Science Dresden, Germany, February 15–17, 2001 Proceedings. pp. 144–157. Springer Berlin Heidelberg, Berlin, Heidelberg (2001)
9. Gawrychowski, P.: Chrobak normal form revisited, with applications. In: Bouchou-Markhoff, B., Caron, P., Champarnaud, J.M., Maurel, D. (eds.) Implementation and Application of Automata. pp. 142–153. Springer Berlin Heidelberg, Berlin, Heidelberg (2011)
10. Geldenhuys, J., van der Merwe, B., van Zijl, L.: Reducing nondeterministic finite automata with sat solvers. In: Yli-Jyrä, A., Kornai, A., Sakarovitch, J., Watson, B. (eds.) Finite-State Methods and Natural Language Processing. pp. 81–92. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
11. Goguen, J.A.: Discrete-time machines in closed monoidal categories. I. J. Comput. Syst. Sci. 10(1), 1–43 (1975)
12. Gramlich, G.: Probabilistic and nondeterministic unary automata. In: Proc. of Math. Foundations of Computer Science, Springer, LNCS 2747, 2003. pp. 460–469. Springer (2003)
13. Grätzer, G.: General Lattice Theory. Birkhäuser Verlag, 2. edn. (1998)
14. Gruber, H., Holzer, M.: Finding lower bounds for nondeterministic state complexity is hard. In: Ibarra, O.H., Dang, Z. (eds.) Developments in Language Theory: 10th International Conference, DLT 2006, Santa Barbara, CA, USA, June 26-29, 2006. Proceedings. pp. 363–374. Springer Berlin Heidelberg, Berlin, Heidelberg (2006)
15. Horn, A., Kimura, N.: The category of semilattices. Algebra Univ. 1, 26–38 (1971)
16. Izhakian, Z., Rhodes, J., Steinberg, B.: Representation theory of finite semigroups over semirings. Journal of Algebra 336(1), 139 – 157 (2011)
17. Jiang, T., McDowell, E., Ravikumar, B.: The structure and complexity of minimal nfa's over a unary alphabet. International Journal of Foundations of Computer Science 02(02), 163–182 (1991)
18. Jiang, T., Ravikumar, B.: Minimal nfa problems are hard. SIAM Journal on Computing 22(6), 1117–1141 (1993)
19. Jipsen, P.: Categories of algebraic contexts equivalent to idempotent semirings and domain semirings. In: Kahl, W., Griffin, T.G. (eds.) Relational and Algebraic Methods in Computer Science. pp. 195–206. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)

20. Johnstone, P.T.: Stone spaces. Cambridge University Press (1982)
21. Kameda, T., Weiner, P.: On the state minimization of nondeterministic finite automata. IEEE Transactions on Computers C-19(7), 617–627 (1970)
22. Kschischang, F.R.: The trellis structure of maximal fixed-cost codes. IEEE Transactions on Information Theory 42(6), 1828–1838 (1996)
23. Latteux, M., Roos, Y., Terlutte, A.: Minimal nfa and birfsa languages. RAIRO - Theoretical Informatics and Applications 43(2), 221–237 (2009)
24. Markowsky, G.: Primes, irreducibles and extremal lattices. Order 9, 265–290 (09 1992)
25. Myers, R.S.R.: Nondeterministic automata and JSL-dfas. CoRR abs/2007.06031 (2020), `https://arxiv.org/abs/2007.06031`
26. Myers, R.S.R.: Representing semilattices as relations. CoRR abs/2007.10277 (2020), `https://arxiv.org/abs/2007.10277`
27. Myers, R.S.R., Adámek, J., Milius, S., Urbat, H.: Coalgebraic constructions of canonical nondeterministic automata. Theoretical Computer Science 604, 81 – 101 (2015)
28. Pin, J.É.: Mathematical foundations of automata theory (September 2020), available at `http://www.liafa.jussieu.fr/~jep/PDF/MPRI/MPRI.pdf`
29. Pin, J.E.: On reversible automata. In: Simon, I. (ed.) LATIN '92. pp. 401–416. Springer Berlin Heidelberg, Berlin, Heidelberg (1992)
30. Polák, L.: Syntactic semiring of a language. In: Sgall, J., Pultr, A., Kolman, P. (eds.) Mathematical Foundations of Computer Science 2001: 26th International Symposium, MFCS 2001 Mariánské Lázne, Czech Republic, August 27–31, 2001 Proceedings. pp. 611–620. Springer Berlin Heidelberg, Berlin, Heidelberg (2001)
31. Shankar, P., Dasgupta, A., Deshmukh, K., Rajan, B.: On viewing block codes as finite automata. Theoretical Computer Science 290(3), 1775 – 1797 (2003)
32. Tamm, H.: New interpretation and generalization of the Kameda-Weiner method. In: Chatzigiannakis, I., Mitzenmacher, M., Rabani, Y., Sangiorgi, D. (eds.) 43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy. LIPIcs, vol. 55, pp. 116:1–116:12. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2016)
33. Tamm, H., Ukkonen, E.: Bideterministic automata and minimal representations of regular languages. Theoretical Computer Science 328(1), 135–149 (2004)
34. van Heerdt, G., Moerman, J., Sammartino, M., Silva, A.: A (co)algebraic theory of succinct automata. Journal of Logical and Algebraic Methods in Programming 105, 112 – 125 (2019)

A Appendix

This Appendix provides full proofs and additional details on the examples omitted for space reasons.

Proof of Lemma 3.11

Let $N = (Q, \delta, I, F)$. We claim that the semilattice isomorphism

$$h\colon [\mathcal{P}(Q)]^{\mathsf{op}} \xrightarrow{\cong} \mathcal{P}(Q), \quad X \mapsto \overline{X} = Q \setminus X,$$

gives an isomorphism of **JSL**-dfas from $[\mathcal{P}(N)]^{\mathsf{op}}$ to $\mathcal{P}(N^{\mathsf{r}})$.

Preservation of the initial state. The initial state of $[\mathcal{P}(N)^{\mathsf{op}}]$ is $\overline{F}$, the largest non-final state of $\mathcal{P}(N)$. Thus h maps it to F, the initial state of $\mathcal{P}(N^{\mathsf{r}})$.

Preservation of final states. By definition, a state X is final in $[\mathcal{P}(N)]^{\mathsf{op}}$ iff $I \nsubseteq X$. This is equivalent to $h(X) \cap I \neq \emptyset$, i.e. to $h(X)$ being final in $\mathcal{P}(N^{\mathsf{r}})$.

Preservation of transitions. Let $X, Y \in \mathcal{P}(Q)$ and $a \in \Sigma$ such that $X \xrightarrow{a} Y$ is a transition in $[\mathcal{P}(N)]^{\mathsf{op}}$. By definition, Y is the set of all $q \in Q$ with $\delta_a[q] \subseteq X$. Thus, $\overline{Y}$ is the set of all $q \in Q$ such $\delta_a[q] \cap \overline{X} \neq \emptyset$. This means that $\overline{X} \xrightarrow{a} \overline{Y}$ is a transition in $\mathcal{P}(N^{\mathsf{r}})$.

Proof of Lemma 3.12

(1) Let $g\colon S \to 2$ be the semilattice morphism corresponding to the prime filter $G = \{x \in S : x \not\leq_S s\}$. Then, for any word $w = a_1 \dots a_n$ in Σ^*, we have $\delta_{w^{\mathsf{r}}}(s_0) \not\leq_S s$ iff the morphism

$$2 \xrightarrow{i} S \xrightarrow{\delta_{a_n}} S \cdots S \xrightarrow{\delta_{a_1}} S \xrightarrow{g} 2$$

is equal to $id\colon 2 \to 2$. This is the case iff the dual morphism

$$2 \xrightarrow{g^*} S^{\mathsf{op}} \xrightarrow{\delta^*_{a_1}} S^{\mathsf{op}} \cdots S^{\mathsf{op}} \xrightarrow{\delta^*_{a_n}} S^{\mathsf{op}} \xrightarrow{i^*} 2$$

is equal to $id\colon 2 \to 2$. Since g^* maps 1 to s, this means precisely that the state s of A^{op} accepts w.

(2) follows from part (1) by choosing s to be the initial state of A^{op}, i.e. the largest non-final state of A.

(3) follows via duality: the smallest subautomaton $\mathsf{reach}(A)$ of A dualizes to the smallest quotient automaton $\mathsf{simple}(A^{\mathsf{op}})$ of A^{op}.

Proof of Proposition 3.13

By Lemma 3.12, the dual of a minimal **JSL**-dfa accepting L^{r} is a minimal **JSL**-dfa accepting L. Thus, by the uniqueness of minimal automata, the unique **JSL**-automata morphism from $[\mathsf{SLD}(L^{\mathsf{r}})]^{\mathsf{op}}$ to $\mathsf{SLD}(L)$, mapping the state K of $[\mathsf{SLD}(L^{\mathsf{r}})]^{\mathsf{op}}$ to the language $L([\mathsf{SLD}(L^{\mathsf{r}})]^{\mathsf{op}}, K)$ it accepts, is an isomorphism. It only remains to verify that this language is equal to $(\overline{K^{\mathsf{r}}})^{-1}L$. To this end, we compute for all $w \in \Sigma^*$:

$$\begin{aligned} w \in L([\mathsf{SLD}(L^{\mathsf{r}})]^{\mathsf{op}}, K) &\iff (w^{\mathsf{r}})^{-1}L^{\mathsf{r}} \nsubseteq K && \text{by Lemma 3.12(1)} \\ &\iff \exists x \in \overline{K} : w^{\mathsf{r}}x \in L^{\mathsf{r}} \\ &\iff \exists y \in \overline{K^{\mathsf{r}}} : yw \in L \\ &\iff w \in (\overline{K^{\mathsf{r}}})^{-1}L \end{aligned}$$

Proof of Theorem 3.15

(1) We need to show that

$$\alpha \colon \mathsf{SLD}(L) \to (\{\mathcal{DR}_L[X] : X \subseteq \mathsf{LD}(L)\}, \cup, \emptyset), \qquad \alpha(K) := \{v^{-1}L^{\mathsf{r}} : v \in K^{\mathsf{r}}\},$$

is an isomorphism. To this end, let $K \in \mathsf{SLD}(L)$, say $K = K_1 \cup \cdots \cup K_n$ for $K_i \in \mathsf{LD}(L)$. We show that

$$\alpha(K) = \mathcal{DR}_L(\{K_1, \ldots, K_n\}),$$

which immediately implies that α is a well-defined isomorphism of semilattices. To this end, we compute for all $v \in \Sigma^*$:

$$\begin{aligned} v^{-1}L^{\mathsf{r}} \in \alpha(K) &\iff v \in K^{\mathsf{r}} \\ &\iff \exists i : v \in {K_i}^{\mathsf{r}} \\ &\iff \exists i : v^{\mathsf{r}} \in K_i \\ &\iff \exists i : v^{-1}L^{\mathsf{r}} \in \mathcal{DR}_L[K_i] \\ &\iff v^{-1}L^{\mathsf{r}} \in \mathcal{DR}_L(\{K_1, \ldots, K_n\}) \end{aligned}$$

(2) Let us first note that the isomorphism dr_L from Proposition 3.13 has the following alternative description:

$$\mathrm{dr}_L(U^{-1}L^{\mathsf{r}}) = \bigcup\{K \in \mathsf{LD}(L) : K \cap U^{\mathsf{r}} = \emptyset\} \qquad \text{for every } U \subseteq \Sigma^*. \qquad \text{(A.1)}$$

In fact, for every $w \in \Sigma^*$ we compute:

$$\begin{aligned} w \in \mathrm{dr}_L(U^{-1}L^{\mathsf{r}}) &\iff w \in \overline{(U^{-1}L^{\mathsf{r}})^{\mathsf{r}}}^{-1}L && \text{def. } \mathrm{dr}_L \\ &\iff \exists v \in \overline{U^{-1}L^{\mathsf{r}}} : v^{\mathsf{r}}w \in L \\ &\iff \exists v \in \Sigma^* : [v^{\mathsf{r}}w \in L \wedge \forall u \in U : uv \notin L^{\mathsf{r}}] \end{aligned}$$

$$\begin{aligned} &\iff \exists y \in \Sigma^* : [yw \in L \wedge \forall u \in U : yu^{\mathsf{r}} \notin L] \\ &\iff \exists y \in \Sigma^* : [w \in y^{-1}L \wedge y^{-1}L \cap U^{\mathsf{r}} = \emptyset] \\ &\iff w \in \bigcup\{K \in \mathsf{LD}(L) : K \cap U^{\mathsf{r}} = \emptyset\}. \end{aligned}$$

It thus follows for all $u, v \in \Sigma^*$:

$$\begin{aligned} u^{-1}L \nsubseteq \mathrm{dr}_L(v^{-1}L^{\mathsf{r}}) &\iff u^{-1}L \notin \{K \in \mathsf{LD}(L) : v^{\mathsf{r}} \notin K\} \\ &\iff v^{\mathsf{r}} \in u^{-1}L \\ &\iff \mathcal{DR}_L(u^{-1}L, v^{-1}L^{\mathsf{r}}). \end{aligned}$$

(3) follows immediately from (2), restricted to $\mathcal{DR}_L^j$.

Proof of Proposition 3.16

(1) Let $\mathrm{At}(x)$ denote the unique atom of $\mathsf{BLD}(L)$ containing the word $x \in \Sigma^*$. For any $v, w \in \Sigma^*$ we have $v^{-1}L^{\mathsf{r}} = w^{-1}L^{\mathsf{r}}$ iff $\mathrm{At}(v^{\mathsf{r}}) = \mathrm{At}(w^{\mathsf{r}})$. In fact,

$$\begin{aligned} & v^{-1}L^{\mathsf{r}} = w^{-1}L^{\mathsf{r}} \\ \text{iff}\quad & \forall x \in \Sigma^* : vx \in L^{\mathsf{r}} \iff wx \in L^{\mathsf{r}} \\ \text{iff}\quad & \forall y \in \Sigma^* : v^{\mathsf{r}} \in y^{-1}L \iff w^{\mathsf{r}} \in y^{-1}L \\ \text{iff}\quad & \mathrm{At}(v^{\mathsf{r}}) = \mathrm{At}(w^{\mathsf{r}}). \end{aligned}$$

In the final step, we use that the boolean algebra $\mathsf{BLD}(L)$ is generated by the left derivatives of L, so two words belong to the same atom iff they belong to the same left derivatives.

(2) It follows that the map $h\colon \mathcal{P}(\mathsf{dfa}(L^{\mathsf{r}})) \to [\mathsf{BLD}(L)]^{\mathsf{op}}$ defined by

$$\{\, w_1^{-1}L^{\mathsf{r}}, \ldots, w_n^{-1}L^{\mathsf{r}} \,\} \quad \mapsto \quad \bigcap_{i=1}^{n} \overline{\mathrm{At}(w_i{}^{\mathsf{r}})}$$

gives a well-defined isomorphism of semilattices. It remains to prove that it is an automata morphism.

Preservation of the initial state. The initial state $\{L^{\mathsf{r}}\}$ of $\mathcal{P}(\mathsf{dfa}(L^{\mathsf{r}}))$ is mapped to $\overline{\mathrm{At}(\varepsilon)}$. This is the largest non-final state of $\mathsf{BLD}(L)$, i.e. the initial state of $[\mathsf{BLD}(L)]^{\mathsf{op}}$.

Preservation of final states. Recall that the final states of $[\mathsf{BLD}(L)]^{\mathsf{op}}$ are those languages in $\mathsf{BLD}(L)$ not containing L. Thus,

$$\begin{aligned} & \{\, w_1^{-1}L^{\mathsf{r}}, \ldots, w_n^{-1}L^{\mathsf{r}} \,\} \text{ final in } \mathsf{BLD}(L) \\ \text{iff}\quad & w_i \in L^{\mathsf{r}} \text{ for some } i \\ \text{iff}\quad & w_i{}^{\mathsf{r}} \in L \text{ for some } i \\ \text{iff}\quad & \mathrm{At}(w_i{}^{\mathsf{r}}) \subseteq L \text{ for some } i \end{aligned}$$

$$\begin{aligned} &\text{iff} \quad L \not\subseteq \overline{\mathrm{At}(w_i{}^{\mathsf{r}})} \text{ for some } i \\ &\text{iff} \quad L \not\subseteq \bigcap_{i=1}^{n} \overline{\mathrm{At}(w_i{}^{\mathsf{r}})} \\ &\text{iff} \quad \bigcap_{i=1}^{n} \overline{\mathrm{At}(w_i{}^{\mathsf{r}})} \text{ final in } [\mathsf{BLD}(L)]^{\mathsf{op}} \end{aligned}$$

Preservation of transitions. Since the semilattice $\mathcal{P}(\mathsf{dfa}(L^{\mathsf{r}}))$ is generated by the left derivatives of L^{r}, it suffices to prove that for each $w \in \Sigma^*$ and $a \in \Sigma$ we have the transition

$$h(\{w^{-1}L^{\mathsf{r}}\}) \xrightarrow{a} h(\{a^{-1}w^{-1}L^{\mathsf{r}}\}),$$

i.e.

$$\overline{\mathrm{At}(w^{\mathsf{r}})} \xrightarrow{a} \overline{\mathrm{At}(aw^{\mathsf{r}})}$$

in $[\mathsf{BLD}(L)]^{\mathsf{op}}$. But this is immediate because $a^{-1}\mathrm{At}(aw^{\mathsf{r}}) \supseteq \mathrm{At}(w^{\mathsf{r}})$.

Proof of Proposition 3.18

The proof is much analogous to the one of Proposition 3.16.

(1) Let $\mathrm{At}(x)$ denote the atom of $\mathsf{BLRD}(L)$ containing the word $x \in \Sigma^*$. For any two words $v, w \in \Sigma^*$ we have $v \equiv_{L^{\mathsf{r}}} w$ iff $\mathrm{At}(v^{\mathsf{r}}) = \mathrm{At}(w^{\mathsf{r}})$. In fact,

$$\begin{aligned} &v \equiv_{L^{\mathsf{r}}} w \\ \text{iff} \quad &\forall x, y \in \Sigma^* : v \in x^{-1}L^{\mathsf{r}}y^{-1} \iff w \in x^{-1}L^{\mathsf{r}}y^{-1} \\ \text{iff} \quad &\forall s, t \in \Sigma^* : v^{\mathsf{r}} \in s^{-1}Lt^{-1} \iff w^{\mathsf{r}} \in s^{-1}Lt^{-1} \\ \text{iff} \quad &\mathrm{At}(v^{\mathsf{r}}) = \mathrm{At}(w^{\mathsf{r}}). \end{aligned}$$

In the final step, we use that the boolean algebra $\mathsf{BLRD}(L)$ is generated by the two-sided derivatives of L, so two words belong to the same atom iff they belong to the same two-sides derivatives.

(2) It follows that the map $h\colon \mathcal{P}(\mathsf{syn}(L^{\mathsf{r}})) \to [\mathsf{BLRD}(L)]^{\mathsf{op}}$ defined by

$$\{\, [w_1]_{L^{\mathsf{r}}}, \dots, [w_n]_{L^{\mathsf{r}}} \,\} \quad \mapsto \quad \bigcap_{i=1}^{n} \overline{\mathrm{At}(w_i{}^{\mathsf{r}})}$$

gives a well-defined isomorphism of semilattices. It remains to prove that it is an automata morphism.

Preservation of the initial state. The initial state $\{[\varepsilon]_{L^{\mathsf{r}}}\}$ of $\mathcal{P}(\mathsf{syn}(L^{\mathsf{r}}))$ is mapped to $\overline{\mathrm{At}(\varepsilon)}$. This is the largest non-final state of $\mathsf{BLRD}(L)$, i.e. the initial state of $[\mathsf{BLRD}(L)]^{\mathsf{op}}$.

Preservation of final states. The final states of $[\mathsf{BLRD}(L)]^{\mathsf{op}}$ are those languages in $\mathsf{BLRD}(L)$ not containing L. Thus,

$$\begin{aligned}
&\{\,[w_1]_{L^{\mathsf{r}}},\dots,[w_n]_{L^{\mathsf{r}}}\,\} \text{ final in } \mathsf{BLRD}(L)\\
\text{iff}\quad & w_i \in L^{\mathsf{r}} \text{ for some } i\\
\text{iff}\quad & {w_i}^{\mathsf{r}} \in L \text{ for some } i\\
\text{iff}\quad & \mathrm{At}({w_i}^{\mathsf{r}}) \subseteq L \text{ for some } i\\
\text{iff}\quad & L \nsubseteq \overline{\mathrm{At}({w_i}^{\mathsf{r}})} \text{ for some } i\\
\text{iff}\quad & L \nsubseteq \bigcap_{i=1}^{n} \overline{\mathrm{At}({w_i}^{\mathsf{r}})}\\
\text{iff}\quad & \bigcap_{i=1}^{n} \overline{\mathrm{At}({w_i}^{\mathsf{r}})} \text{ final in } [\mathsf{BLRD}(L)]^{\mathsf{op}}
\end{aligned}$$

Preservation of transitions. Since the semilattice $\mathcal{P}(\mathsf{syn}(L^{\mathsf{r}}))$ is generated by the elements of $\mathsf{syn}(L)$, it suffices to prove that for each $w \in \Sigma^*$ and $a \in \Sigma$ we have the transition

$$h(\{[w]_{L^{\mathsf{r}}}\}) \xrightarrow{a} h(\{[wa]_{L^{\mathsf{r}}}\}),$$

i.e.

$$\overline{\mathrm{At}(w^{\mathsf{r}})} \xrightarrow{a} \overline{\mathrm{At}(aw^{\mathsf{r}})}$$

in $[\mathsf{BLRD}(L)]^{\mathsf{op}}$. But this is immediate because $a^{-1}\mathrm{At}(aw^{\mathsf{r}}) \supseteq \mathrm{At}(w^{\mathsf{r}})$.

Proof of Proposition 3.20

Let $A = (S, \delta, s_0, F)$. For any $K \subseteq \Sigma^*$ we put $\delta_K := \bigvee_{w\in K} \delta_w$.

(1) We first show that

$$L([\mathsf{ts}(A)]^{\mathsf{op}}, \delta_K) \;=\; \bigcup_{v\in\Sigma^*} L(A^{\mathsf{op}}, \delta_{vK}(s_0))(v^{\mathsf{r}})^{-1} \qquad \text{for each } K \subseteq \Sigma^*. \qquad (\mathrm{A}.2)$$

To see this, we compute for all $u \in \Sigma^*$:

$$\begin{aligned}
& u \in L([\mathsf{ts}(A)]^{\mathsf{op}}, \delta_K) & \\
\text{iff}\quad & \delta_{u^{\mathsf{r}}} \nleq \delta_K & \text{by Lemma 3.12(1)}\\
\text{iff}\quad & \exists v \in \Sigma^* : \delta_{u^{\mathsf{r}}}(\delta_v(s_0)) \nleq_S \delta_K(\delta_v(s_0)) & \text{since } A \text{ is reachable}\\
\text{iff}\quad & \exists v \in \Sigma^* : \delta_{vu^{\mathsf{r}}}(s_0) \nleq_S \delta_{vK}(s_0) & \\
\text{iff}\quad & \exists v \in \Sigma^* : uv^{\mathsf{r}} \in L(A^{\mathsf{op}}, \delta_{vK}(s_0)) & \text{by Lemma 3.12(1)}\\
\text{iff}\quad & \exists v \in \Sigma^* : u \in L(A^{\mathsf{op}}, \delta_{vK}(s_0))(v^{\mathsf{r}})^{-1}. &
\end{aligned}$$

(2) For any $w \in \Sigma^*$, consider the two semilattice morphisms

$$\begin{aligned} \gamma_w \colon \mathsf{ts}(A) &\to \mathsf{ts}(A), \quad f \mapsto \delta_w \circ f, \\ \varphi_w \colon \mathsf{ts}(A) &\to \mathsf{ts}(A), \quad f \mapsto f \circ \delta_w. \end{aligned}$$

along with their dual morphisms $\gamma_w^*, \varphi_w^* \colon [\mathsf{ts}(A)]^{\mathsf{op}} \to [\mathsf{ts}(A)]^{\mathsf{op}}$. We claim that

$$L([\mathsf{ts}(A)]^{\mathsf{op}}, \delta_K)(w^{\mathsf{r}})^{-1} = L([\mathsf{ts}(A)]^{\mathsf{op}}, \varphi_w^*(\delta_K)) \qquad \text{for each } K \subseteq \Sigma^*. \tag{A.3}$$

To see this, we compute as follows for all $u \in \Sigma^*$, where $\leq$ is the order of the semilattice $\mathbf{JSL}(S, S)$:

$$\begin{array}{lll} & u \in L([\mathsf{ts}(A)]^{\mathsf{op}}, \varphi_w^*(\delta_K)) & \\ \text{iff} & id_S \not\leq (\gamma_{u^{\mathsf{r}}})^*(\varphi_w^*(\delta_K)) & \text{def. } L(-,-) \\ \text{iff} & id_S \not\leq (\varphi_w \circ \gamma_{u^{\mathsf{r}}})^*(\delta_K) & \\ \text{iff} & \varphi_w \circ \gamma_{u^{\mathsf{r}}}(id_S) \not\leq \delta_K & \text{by adjointness} \\ \text{iff} & \delta_{wu^{\mathsf{r}}} \not\leq \delta_K & \\ \text{iff} & \gamma_{wu^{\mathsf{r}}}(id_S) \not\leq \delta_K & \\ \text{iff} & id_S \not\leq (\gamma_{wu^{\mathsf{r}}})^*(\delta_K) & \text{by adjointness} \\ \text{iff} & uw^{\mathsf{r}} \in L([\mathsf{ts}(A)]^{\mathsf{op}}, \delta_K) & \text{def. } L(-,-) \\ \text{iff} & u \in L([\mathsf{ts}(A)]^{\mathsf{op}}, \delta_K)(w^{\mathsf{r}})^{-1} & \end{array}$$

(3) We are ready to prove the proposition. Since both $[\mathsf{ts}(A)]^{\mathsf{op}}$ and $\mathsf{rdc}(A^{\mathsf{op}})$ are simple $\mathbf{JSL}$-dfas, and thus can be viewed as subautomata of $\mathsf{Fin}(L)$, it suffices to show that they contain the same languages. The inclusion $[\mathsf{ts}(A)]^{\mathsf{op}} \subseteq \mathsf{rdc}(A^{\mathsf{op}})$ follows from (A.2). For the reverse inclusion, since $[\mathsf{ts}(A)]^{\mathsf{op}}$ is closed under right derivatives by (A.3), we only need to prove that $A^{\mathsf{op}} \subseteq [\mathsf{ts}(A)]^{\mathsf{op}}$. To this end, we show that, for any $s \in S$,

$$L(A^{\mathsf{op}}, s) = L([\mathsf{ts}(A)]^{\mathsf{op}}, \delta_K), \quad \text{where} \quad K = \{\, w \in \Sigma^* \;:\; \delta_w(s_0) \leq_S s \,\}.$$

For the proof, we first note that for all $u \in \Sigma^*$,

$$\delta_u(s_0) \leq_S s \qquad \Longleftrightarrow \qquad \forall v \in \Sigma^* : \delta_{vu}(s_0) \leq_S \delta_{vK}(s_0). \tag{A.4}$$

In fact, "$\Leftarrow$" follows by taking $v = \varepsilon$; we have $s = \delta_K(s_0)$ because A is reachable. For "$\Rightarrow$", suppose that $\delta_u(s_0) \leq_S s$. Then $u \in K$ and therefore

$$\delta_{vu}(s_0) \leq_S \bigvee_{w \in K} \delta_{vw}(s_0) = \delta_{vK}(s_0)$$

We now compute

$$\begin{array}{lll} & u \in L(A^{\mathsf{op}}, s) & \\ \text{iff} & \delta_{u^{\mathsf{r}}}(s_0) \not\leq_S s & \text{by Lemma 3.12(1)} \\ \text{iff} & \exists v \in \Sigma^* : \delta_{vu^{\mathsf{r}}}(s_0) \not\leq_S \delta_{vK}(s_0) & \text{by (A.4)} \end{array}$$

$$\begin{aligned}
&\text{iff} \quad \exists v \in \Sigma^* : uv^{\mathsf{r}} \in L(A^{\mathsf{op}}, \delta_{vK}(s_0)) && \text{by Lemma 3.12(1)}\\
&\text{iff} \quad \exists v \in \Sigma^* : u \in L(A^{\mathsf{op}}, \delta_{vK}(s_0))(v^{\mathsf{r}})^{-1} && \\
&\text{iff} \quad u \in L([\mathsf{ts}(A)]^{\mathsf{op}}, \delta_K) && \text{by (A.2)}
\end{aligned}$$

This concludes the proof.

Proof of Theorem 4.7

Let $d(L)$ denote the least degree of any boolean representation extending the canonical representation $\kappa_L \circ \mu_L$.

(1) A boolean presentation of Σ^* is given by a finite semilattice lattice S together with a family of semilattice morphisms $\delta = (\delta_a \colon S \to S)_{a\in\Sigma}$. An equivariant map between boolean presentations (S, δ) and (S', δ') is a semilattice morphism $h\colon S \to S'$ with $\delta'_a \circ h = h \circ \delta_a$ for all $a \in \Sigma$. If S carries a **JSL**-automata structure (S, δ, i, f) and h is a monic, there exists an automata structure on S' making h an automata morphism: put $i' := h \circ i$, and choose $f'\colon S' \to 2$ to be any semilattice morphism with $f' = h \circ f$. Such an f' exists because the semilattice 2 is an injective object of **JSL**.

$$\begin{tikzcd}
2 \arrow[r, "i"] \arrow[dr, "i'"'] & S \arrow[r, "\delta_a"] \arrow[d, "h"] & S \arrow[r, "f"] \arrow[d, "h"'] & 2 \\
& S' \arrow[r, "\delta'_a"'] & S' \arrow[ur, "f'"'] &
\end{tikzcd}$$

(2) To prove $d(L) \leq \mathrm{ns}(L)$, suppose that N is an nfa accepting the language L. Consider the **JSL**-subautomaton $\mathsf{langs}(N) = \mathsf{simple}(\mathcal{P}(N))$ of $\mathsf{Fin}(L)$ carried by the semilattice of all languages accepted by subsets of N. Note that $\mathsf{SLD}(L)$ is a subautomaton of $\mathsf{langs}(N)$: every finite union $\bigcup_i w_i^{-1}L$ of left derivatives of L is accepted by the set of all states of N reachable on input w_i for some i. Thus, the inclusion map $\mathsf{SLD}(L) \rightarrowtail \mathsf{langs}(N)$ defines an extension of the canonical representation $\kappa_L \circ \mu_L$. Since the semilattice $\mathsf{langs}(N)$ is generated by the set of languages accepted by single states of N, it follows that the degree of this representation is at most the number of states of N.

(3) To prove $\mathrm{ns}(L) \leq d(L)$, suppose that (S, δ) is a boolean representation of Σ^* of degree k extending $\kappa_L \circ \mu_L$, witnessed by an injective equivariant map $h\colon \mathsf{SLD}(L) \rightarrowtail S$. By part (1), we can equip S with a **JSL**-dfa structure making h an automata morphism. Since morphisms preserve accepted languages, it follows that S accepts L. The automaton S has k join-irreducibles, so Remark 3.4 shows that there exists an nfa on k states accepting L.

Proof of Theorem 4.12

Remark A.1. The subset construction, restricted to dfas, gives rise to a left adjoint $\mathcal{P}\colon \mathbf{Aut}(\mathbf{Set}_{\mathsf{f}}) \to \mathbf{Aut}(\mathbf{JSL}_{\mathsf{f}})$ between the categories of dfas and **JSL**-

dfas. Thus, for any dfa D and any **JSL**-dfa A, there is a bijective correspondence between dfa morphisms from D to A and **JSL**-dfa morphisms from $\mathcal{P}(D)$ to A.

Our proof of Theorem 4.12 is essentially an instance of the self-duality of **JSL**-dfas. Let L be the language accepted by N. We establish the theorem by showing that each of the following statements is equivalent to the next one:

(1) N is atomic.

(2) There exists a **JSL**-automata morphism from $\mathcal{P}(N)$ to $\mathsf{BLD}(L)$.

(3) There exists a **JSL**-automata morphism from $\mathsf{simple}(\mathcal{P}(N))$ to $\mathsf{BLD}(L)$.

(4) There exists a **JSL**-automata morphism from $\mathcal{P}(\mathsf{dfa}(L^\mathsf{r}))$ to $\mathsf{reach}(\mathcal{P}(N^\mathsf{r}))$.

(5) There exists a dfa morphism from $\mathsf{dfa}(L^\mathsf{r})$ to $\mathsf{reach}(\mathcal{P}(N^\mathsf{r}))$.

(6) There exists a dfa morphism from $\mathsf{dfa}(L^\mathsf{r})$ to $\mathsf{rsc}(N^\mathsf{r})$.

(7) $\mathsf{rsc}(N^\mathsf{r})$ is a minimal dfa.

Ad (1)⇔(2). The unique automata morphism $m_{\mathcal{P}(N)}\colon \mathcal{P}(N) \to \mathsf{Fin}(L)$ maps every state of $\mathcal{P}(N)$ to the language it accepts. Thus, N is atomic iff $m_{\mathcal{P}(N)}$ factorizes through the subautomaton $\mathsf{BLD}(L)$ of $\mathsf{Fin}(L)$.

Ad (2)⇔(3). Since $\mathsf{BLD}(L)$ is simple, being a subautomaton of $\mathsf{Fin}(L)$, every automata morphism from $\mathcal{P}(N)$ to $\mathsf{BLD}(L)$ factorizes through $\mathsf{simple}(\mathcal{P}(N))$. The converse is obvious since $\mathsf{simple}(\mathcal{P}(N))$ is a quotient automaton of $\mathcal{P}(N)$.

Ad (3)⇔(4). This follows via duality from Lemma 3.11, Lemma 3.12 and Proposition 3.16.

Ad (4)⇔(5). This follows from Remark A.1.

Ad (5)⇔(6). Note that $\mathsf{rsc}(\mathcal{P}(N^\mathsf{r}))$ forms the dfa-reachable part of the automaton $\mathsf{reach}(\mathcal{P}(N^\mathsf{r}))$. Thus, since $\mathsf{dfa}(L^\mathsf{r})$ is a reachable dfa, every dfa morphism from $\mathsf{dfa}(L^\mathsf{r})$ to $\mathsf{reach}(\mathcal{P}(N^\mathsf{r}))$ factorizes through $\mathsf{rsc}(\mathcal{P}(N^\mathsf{r}))$.

Ad (6)⇔(7). Every dfa morphism from $\mathsf{dfa}(L^\mathsf{r})$ to $\mathsf{rsc}(N^\mathsf{r})$ is an isomorphism: it is injective because $\mathsf{dfa}(L^\mathsf{r})$ is a simple dfa and surjective because $\mathsf{rsc}(N^\mathsf{r})$ is a reachable dfa. Conversely, if $\mathsf{rsc}(N^\mathsf{r})$ is a minimal dfa, then it is isomorphic to $\mathsf{dfa}(L^\mathsf{r})$ by the uniqueness of minimal dfas.

Proof of Theorem 4.13

Let us first recall the concept of algebraic language recognition [28].

Remark A.2. A finite monoid M is said to *recognize* the language $L \subseteq \Sigma^*$ if there exists a monoid morphism $h\colon \Sigma^* \to M$ and a subset $P \subseteq M$ with $L = h^{-1}[P]$. Regular languages are exactly the languages recognizable by finite monoids. In fact, we have the following connections between monoids and dfas:

(1) If L is recognized by a finite monoid M via $h\colon \Sigma^* \to M$ and $P \subseteq M$, then M can be viewed as dfa accepting L, with transitions $m \xrightarrow{a} m \bullet h(a)$ for $m \in M$ and $a \in \Sigma$, initial state 1_M, and final states P.

(2) Conversely, if L is accepted by a dfa $D = (S, \delta, s_0, F)$, then the transition monoid $\mathsf{tm}(D)$ recognizes L via the morphism $h\colon \Sigma^* \twoheadrightarrow \mathsf{tm}(D)$, $w \mapsto \delta_w$, and $P = \{\delta_w \colon w \in L\}$. In particular, the syntactic monoid recognizes L via the syntactic morphism $\mu_L\colon \Sigma^* \twoheadrightarrow \mathsf{syn}(L)$. It can be characterized as the least quotient monoid of Σ^* recognizing L: for any surjective monoid morphism $h\colon \Sigma^* \twoheadrightarrow M$ recognizing L, there is a unique morphism $g\colon M \twoheadrightarrow \mathsf{syn}(L)$ with $\mu_L = g \circ h$:

$$\begin{tikzcd} & \Sigma^* \arrow[dl, two heads, "h"'] \arrow[dr, two heads, "\mu_L"] & \\ M \arrow[rr, dashed, two heads, "g"'] & & \mathsf{syn}(L) \end{tikzcd}$$

(3) Finally, there is a tight connection between morphisms of monoids and dfas. Suppose that two surjective monoid morphisms $h_i\colon \Sigma^* \twoheadrightarrow M_i$ and subsets $P_i \subseteq M_i$ for $i = 1, 2$ are given. As in part (1), we view M_1 and M_2 as dfas. Then every dfa morphism $g\colon M_1 \to M_2$ makes the triangle below commute:

$$\begin{tikzcd} & \Sigma^* \arrow[dl, two heads, "h_1"'] \arrow[dr, two heads, "h_2"] & \\ M_1 \arrow[rr, dashed, two heads, "g"'] & & M_2 \end{tikzcd}$$

In fact, M_1 and M_2 accept the same language L and Σ^* can be seen as the initial dfa accepting L when equipped with $L \subseteq \Sigma^*$ as the set of final states. From the surjectivity of h_1 it easily follows that g is a monoid morphism. Conversely, every monoid morphism g making the above triangle commute and satisfying $g[P_1] = P_2$ is a dfa morphism.

Remark A.3. For any **JSL**-dfa A, the dfa-reachable part of $\mathsf{ts}(\mathsf{reach}(A))$ is $\mathsf{tm}(A_r)$, where A_r denotes the dfa-reachable part of A. In fact, letting $\mathsf{reach}(A) = (S, \delta, s_0, F)$ and $A_r = (S_r, \delta_r, s_{0,r}, F_r)$, we have that A_r is a sub-dfa of $\mathsf{reach}(A)$. Then the map $(\delta_r)_w \mapsto \delta_w$ gives a well-defined injective dfa morphism from $\mathsf{tm}(A_r)$ to $\mathsf{ts}(\mathsf{reach}(A))$, using that the semilattice S is generated by the subset $S_r \subseteq S$. Thus, $\mathsf{tm}(A_r)$ is a sub-dfa of $\mathsf{ts}(\mathsf{reach}(A))$. Since it is reachable, it it isomorphic to the dfa-reachable part of $\mathsf{ts}(\mathsf{reach}(A))$.

With these preparations, we are ready to prove Theorem 4.13. Again, the argument crucially rests on the self-duality of **JSL**-dfas. We show that each of the following statements is equivalent to the next one:

(1) N is subatomic.

(2) There exists a **JSL**-dfa morphism from $\mathcal{P}(N)$ to $\mathsf{BLRD}(L)$.

(3) There exists a **JSL**-dfa morphism from $\mathsf{rdc}(\mathsf{simple}(P(N)))$ to $\mathsf{BLRD}(L)$.

(4) There exists a **JSL**-dfa morphism from $\mathcal{P}(\mathsf{syn}(L^{\mathsf{r}}))$ to $\mathsf{ts}(\mathsf{reach}(\mathcal{P}(N^{\mathsf{r}})))$.

(5) There exists a dfa morphism from $\mathsf{syn}(L^{\mathsf{r}})$ to $\mathsf{ts}(\mathsf{reach}(\mathcal{P}(N^{\mathsf{r}})))$.

(6) There exists a dfa morphism from $\mathsf{syn}(L^{\mathsf{r}})$ to $\mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$.

(7) The monoids $\mathsf{syn}(L^{\mathsf{r}})$ and $\mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$ are isomorphic.

Ad (1)⇔(2). The unique automata morphism $m_{\mathcal{P}(N)}\colon \mathcal{P}(N) \to \mathsf{Fin}(L)$ maps every state of $\mathcal{P}(N)$ to the language it accepts. Thus, N is subatomic iff $m_{\mathcal{P}(N)}$ factorizes through the subautomaton $\mathsf{BLRD}(L)$ of $\mathsf{Fin}(L)$.

Ad (2)⇔(3). This is clear since $\mathsf{BLRD}(L)$ is closed under right derivatives.

Ad (3)⇔(4). This follows via duality from Lemma 3.11, Proposition 3.18 and Proposition 3.20.

Ad (4)⇔(5). This follows from Remark A.1.

Ad (5)⇔(6). Putting $A = \mathcal{P}(N^{\mathsf{r}})$ in Remark A.3, we see that $\mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$ is the dfa-reachable part of $\mathsf{ts}(\mathsf{reach}(\mathcal{P}(N^{\mathsf{r}})))$. Since $\mathsf{syn}(L^{\mathsf{r}})$ is reachable as a dfa, it follows that every dfa morphism into $\mathsf{ts}(\mathsf{reach}(\mathcal{P}(N^{\mathsf{r}})))$ factorizes through $\mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$.

Ad (6)⇒(7). Let $q_{N^{\mathsf{r}}}\colon \Sigma^* \twoheadrightarrow \mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$ denote the canonical monoid morphism mapping $w \in \Sigma^*$ to the transition morphism δ_w of the dfa $\mathsf{rsc}(N^{\mathsf{r}})$. Note that the dfa structure of $\mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$ is precisely the one induced by $q_{N^{\mathsf{r}}}$. Thus, given a dfa morphism $h\colon \mathsf{syn}(L^{\mathsf{r}}) \to \mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$ we know that the following diagram commutes by initiality, see Remark A.2(3):

$$\begin{tikzcd}
& \Sigma^* \arrow[dl, two heads, "\mu_{L^{\mathsf{r}}}"'] \arrow[dr, two heads, "q_{N^{\mathsf{r}}}"] & \\
\mathsf{syn}(L^{\mathsf{r}}) \arrow[rr, dashed, two heads, "h"'] & & \mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))
\end{tikzcd} \tag{A.5}$$

Then h is necessarily a monoid morphism because $\mu_{L^{\mathsf{r}}}$ is surjective. Since $q_{N^{\mathsf{r}}}$ recognizes the language L^{r}, we get a unique monoid morphism $g\colon \mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}})) \to \mathsf{syn}(L^{\mathsf{r}})$ with $g \circ q_{N^{\mathsf{r}}} = \mu_{L^{\mathsf{r}}}$. It follows that h is an isomorphism with $h^{-1} = g$.

Ad (7)⇒(6). Suppose that the monoids $\mathsf{syn}(L^{\mathsf{r}})$ and $\mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$ are isomorphic. Let again $g\colon \mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}})) \to \mathsf{syn}(L^{\mathsf{r}})$ be the unique monoid morphism with $g \circ q_{N^{\mathsf{r}}} = \mu_{L^{\mathsf{r}}}$. Then g is surjective because $\mu_{L^{\mathsf{r}}}$ is. Since $\mathsf{syn}(L^{\mathsf{r}})$ and $\mathsf{tm}(\mathsf{rsc}(L^{\mathsf{r}}))$ have the same number of elements, it follows that g is also injective, i.e. an isomorphism of monoids. Then Remark A.2(3) shows that its inverse $g^{-1}\colon \mathsf{syn}(L^{\mathsf{r}}) \to \mathsf{tm}(\mathsf{rsc}(N^{\mathsf{r}}))$ is a dfa morphism.

Proof of Theorem 4.14

Let $a(L)$ denote the least number of states of any subatomic nfa accepting L. We are to prove $a(L) = \mathrm{n}\mu(L)$.

(1) To prove $\mathrm{n}\mu(L) \leq a(L)$, suppose that N is a subatomic nfa accepting the language L. Consider the subsemilattice $\mathsf{langs}(N) = \mathsf{simple}(\mathcal{P}(N))$ of $\mathsf{Fin}(L)$ of all languages accepted by subsets of N. We claim that

$$\rho\colon \mathsf{syn}(L) \to \mathbf{JSL}(\mathsf{langs}(N), \mathsf{langs}(N)), \quad [w]_L \mapsto \lambda K.w^{-1}K$$

is a boolean representation of $\mathsf{syn}(L)$ extending the canonical one. This is obvious once we prove ρ to be a well-defined map, i.e.

$$v \equiv_L w \quad \text{implies} \quad v^{-1}K = w^{-1}K$$

for $v, w \in \Sigma^*$ and $K \in \mathsf{langs}(N)$. Since $K \in \mathsf{BLRD}(L)$, the boolean algebra generated by all two-sided derivatives of L, and derivatives commute with all set-theoretic boolean operations, we can assume w.l.o.g. that $K = s^{-1}Lt^{-1}$ for some $s, t \in \Sigma^*$. Then, for all $x \in \Sigma^*$,

$$\begin{aligned} x \in v^{-1}K &\iff x \in v^{-1}s^{-1}Lt^{-1} \\ &\iff svxt \in L \\ &\iff swxt \in L && \text{since } v \equiv_L w \\ &\iff x \in w^{-1}s^{-1}Lt^{-1} \\ &\iff x \in w^{-1}K \end{aligned}$$

proving that $v^{-1} = w^{-1}L$, as required. Since the semilattice $\mathsf{langs}(N)$ is generated by the set of languages accepted by single states of N, it follows that $\deg(\rho)$ is at most the number of states of N.

(2) To prove $a(L) \leq \mathrm{n}\mu(L)$, let $\rho\colon \mathsf{syn}(L) \to \mathbf{JSL}(S, S)$ be a boolean representation of $\mathsf{syn}(L)$ extending the canonical one. Then $\rho \circ \mu_L \colon \Sigma^* \to \mathbf{JSL}(S, S)$ extends the canonical presentation $\kappa_L \circ \mu_L$ of Σ^*, and so like in proof of Theorem 4.7 we can equip S with the structure of a $\mathbf{JSL}$-dfa $A = (S, \delta, i, f)$ accepting L. Its extended transition morphism for $w \in \Sigma^*$ is given by

$$\delta_w \colon S \to S, \quad s \mapsto \rho([w]_L)(s).$$

In particular, $v \equiv_L w$ implies $\delta_v = \delta_w$, which shows that every state of A accepts a union of syntactic congruence classes of L. Since

$$[w]_L = \bigcap_{xwy \in L} x^{-1}Ly^{-1} \cap \bigcap_{xwy \notin L} \overline{x^{-1}Ly^{-1}},$$

it follows that all languages accepted by states of A lie in $\mathsf{BLRD}(L)$. Therefore, the nfa N of join-irreducibles of A (see Remark 3.4) is a subatomic nfa with $\deg(\rho)$ states accepting L.

Proof of Theorem 5.2

(1) Suppose that $\mathsf{syn}(L) = \mathsf{tm}(\mathsf{dfa}(L))$ is cyclic. Then there exists $w_0 \in \Sigma^*$ such that the map $\lambda X.w_0^{-1}X \colon \mathsf{LD}(L) \to \mathsf{LD}(L)$ generates $\mathsf{tm}(\mathsf{dfa}(L))$. We claim that, for all $K, M \subseteq \Sigma^*$

$$K^{-1}L = M^{-1}L \quad \text{iff} \quad [\forall n \in \mathbb{N} : w_0^n \in K^{-1}L \iff w_0^n \in M^{-1}L]. \tag{A.6}$$

The "only if" direction is trivial. For the converse, suppose that $K^{-1}L \neq M^{-1}L$. W.l.o.g. we may assume that there exists $w \in K^{-1}L \setminus M^{-1}L$. Choose $i_1, \ldots i_k$

and $j_1, \dots, j_m$ such that $K^{-1}L = \bigcup_{p=1}^{k} (w_0^{i_p})^{-1}L$ and $M^{-1}L = \bigcup_{r=1}^{m} (w_0^{j_r})^{-1}L$. Moreover, choose $n \in \mathbb{N}$ such that $w^{-1}L = (w_0^n)^{-1}L$. Then we have $w \in (w_0^{i_p})^{-1}L$ for some p and thus $w_0^{i_p} \in w^{-1}L = (w_0^n)^{-1}L$, using that $\mathsf{tm}(\mathsf{dfa}(L))$ is a commutative monoid. Thus, $w_0^n \in (w_0^{i_p})^{-1}L \subseteq K^{-1}L$. On the other hand, we have $w \notin (w_0^{j_r})^{-1}L$ for all r, so the same argument shows that $(w_0)^n \notin M^{-1}L$.

(2) Fix an alphabet $\Sigma_0 = \{a_0\}$ disjoint from Σ and consider the unary language

$$L_0 := \{\, a_0^n \;:\; n \in \mathbb{N},\, w_0^n \in L \,\} \subseteq \Sigma_0^*.$$

Let $g : \Sigma_0^* \to \Sigma^*$ be the monoid morphism where $g(a_0) := w_0$. We claim that the following map is a **JSL**-isomorphism:

$$f \colon \mathsf{SLD}(L_0) \xrightarrow{\cong} \mathsf{SLD}(L), \quad f(X^{-1}L_0) := g[X]^{-1}L.$$

To see that f is well-defined and injective, we prove for all $X, Y \subseteq \Sigma_0^*$:

$$X^{-1}L_0 = Y^{-1}L_0 \quad \text{iff} \quad g[X]^{-1}L = g[Y]^{-1}L.$$

In fact, we have

$$\begin{aligned}
& X^{-1}L_0 = Y^{-1}L_0 \\
\text{iff}\quad & \forall n \in \mathbb{N} : a_0^n \in X^{-1}L_0 \iff a_0^n \in Y^{-1}L_0 \\
\text{iff}\quad & \forall n \in \mathbb{N} : [\exists a_0^k \in X : a_0^{n+k} \in L_0] \iff [\exists a_0^m \in Y : a_0^{n+m} \in L_0] \\
\text{iff}\quad & \forall n \in \mathbb{N} : [\exists a_0^k \in X : w_0^{n+k} \in L] \iff [\exists a_0^m \in Y : w_0^{n+m} \in L] \\
\text{iff}\quad & \forall n \in \mathbb{N} : [\exists a_0^k \in X : w_0^n \in (g(a_0)^k)^{-1}L] \Leftrightarrow [\exists a_0^m \in Y : w_0^n \in (g(a_0)^m)^{-1}L] \\
\text{iff}\quad & \forall n \in \mathbb{N} : w_0^n \in g[X]^{-1}L \iff w_0^n \in g[Y]^{-1}L \\
\text{iff}\quad & g[X]^{-1}L = g[Y]^{-1}L
\end{aligned}$$

where the final step uses (A.6). This proves f to be well-defined and injective. Moreover, it immediately follows from the definition that f is surjective and preserves finite unions.

(3) For each $a \in \Sigma$ choose $n_a \in \mathbb{N}$ such that $a^{-1}K = (w_0^{n_a})^{-1}K$ for all $K \in \mathsf{LD}(L)$. The respective transition endomorphisms of the **JSL**-automata $\mathsf{SLD}(L_0)$ and $\mathsf{SLD}(L)$ determine each other in the sense that the following diagrams commute:

$$\begin{tikzcd}
\mathsf{SLD}(L_0) \arrow[r, "f"] \arrow[r, "\cong"'] \arrow[d, "a_0^{-1}(-)"'] & \mathsf{SLD}(L) \arrow[d, "w_0^{-1}(-)"] \\
\mathsf{SLD}(L_0) \arrow[r, "\cong"] \arrow[r, "f"'] & \mathsf{SLD}(L)
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathsf{SLD}(L_0) \arrow[r, "f"] \arrow[r, "\cong"'] \arrow[d, "(a_0^{n_a})^{-1}(-)"'] & \mathsf{SLD}(L) \arrow[d, "a^{-1}(-)"] \\
\mathsf{SLD}(L_0) \arrow[r, "\cong"] \arrow[r, "f"'] & \mathsf{SLD}(L)
\end{tikzcd}$$

It follows that extensions of the canonical representations κ_L and $\kappa_L \circ \mu_L$ correspond uniquely to extensions of the canonical representations κ_{L_0} and $\kappa_{L_0} \circ \mu_{L_0}$, respectively. Therefore, $\mathrm{ns}(L) = \mathrm{ns}(L_0)$ by Theorem 4.7 and $\mathrm{n}\mu(L) = \mathrm{n}\mu(L_0)$ by Theorem 4.14. Moreover, from Example 5.1 we know that $\mathrm{ns}(L_0) = \mathrm{n}\mu(L_0)$, and so $\mathrm{ns}(L) = \mathrm{n}\mu(L)$ as claimed.

Details for Example 5.7

We prove that the map θ gives an nfa isomorphism from N_L to $(N_{L^{\mathsf{r}}})^{\mathsf{r}}$. Note first that if $\theta(u^{-1}L) = v^{-1}L^{\mathsf{r}}$, we have

$$u^{-1}L \subseteq X \iff v^{\mathsf{r}} \in X \qquad \text{for } X \in \mathsf{LD}(L).$$

In fact,

$$\begin{aligned} u^{-1}L \subseteq X &\iff X \not\subseteq \tau(u^{-1}L) && \text{def. } \tau \\ &\iff X \not\subseteq \mathrm{dr}_L(v^{-1}L^{\mathsf{r}}) && \tau = \mathrm{dr}_L \circ \theta \\ &\iff \mathcal{DR}_L(X, v^{-1}L^{\mathsf{r}}) && \text{by Theorem 3.15} \\ &\iff v^{\mathsf{r}} \in X && \text{def. } \mathcal{DR}_L \end{aligned}$$

With this preparation, we verify that θ satisfies the properties of an nfa morphism:

Preservation of initial and final states. Let $u^{-1}L \in J(\mathsf{SLD}(L))$ and $\theta(u^{-1}L) = v^{-1}L^{\mathsf{r}}$. Then

$$u^{-1}L \subseteq L \iff v^{\mathsf{r}} \in L \iff v \in L^{\mathsf{r}} \iff \varepsilon \in v^{-1}L^{\mathsf{r}}.$$

A symmetric argument, exchanging the roles of L and L^{r}, shows that

$$\varepsilon \in u^{-1}L \iff v^{-1}L^{\mathsf{r}} \subseteq L^{\mathsf{r}}.$$

Thus, the state $u^{-1}L$ is initial/final in N_L iff $v^{-1}L^{\mathsf{r}}$ is initial/final in $(N_{L^{\mathsf{r}}})^{\mathsf{r}}$.

Preservation of transitions. Let $u^{-1}L, \overline{u}^{-1}L \in J(\mathsf{SLD}(L))$ and $\theta(u^{-1}L) = v^{-1}L^{\mathsf{r}}$, $\theta(\overline{u}^{-1}L) = \overline{v}^{-1}L^{\mathsf{r}}$. For each $a \in \Sigma$, we need to show that there is a transition $u^{-1}L \xrightarrow{a} \overline{u}^{-1}L$ in N_L iff there is a transition $v^{-1}L^{\mathsf{r}} \xrightarrow{a} \overline{v}^{-1}L^{\mathsf{r}}$ in $(N_{L^{\mathsf{r}}})^{\mathsf{r}}$. In fact:

$$\begin{aligned} \overline{u}^{-1}L \subseteq (ua)^{-1}L &\iff \overline{v}^{\mathsf{r}} \in (ua)^{-1}L \\ &\iff ua\overline{v}^{\mathsf{r}} \in L \\ &\iff \overline{v}au^{\mathsf{r}} \in L^{\mathsf{r}} \\ &\iff u^{\mathsf{r}} \in (\overline{v}a)^{-1}L^{\mathsf{r}} \\ &\iff v^{-1}L^{r} \subseteq (\overline{v}a)^{-1}L^{\mathsf{r}}. \end{aligned}$$