Available online at www.sciencedirect.com

ScienceDirect

Electronic Notes in Theoretical Computer Science 308 (2014) 3–23

Electronic Notes in Theoretical Computer Science

www.elsevier.com/locate/entcs

On Continuous Nondeterminism and State Minimality

Jiří Adámek, Robert S. R. Myers, Henning Urbat

Institut für Theoretische Informatik
Technische Universität Braunschweig
Germany

Stefan Milius

Lehrstuhl für Theoretische Informatik
Friedrich-Alexander-Universität Erlangen-Nürnberg
Germany

Abstract

This paper is devoted to the study of nondeterministic closure automata, that is, nondeterministic finite automata (nfas) equipped with a strict closure operator on the set of states and continuous transition structure. We prove that for each regular language L there is a unique minimal nondeterministic closure automaton whose underlying nfa accepts L. Here minimality means no proper sub or quotient automata exist, just as it does in the case of minimal dfas. Moreover, in the important case where the closure operator of this machine is topological, its underlying nfa is shown to be state-minimal. The basis of these results is an equivalence between the categories of finite semilattices and finite strict closure spaces.

Keywords: Canonical Nondeterministic Automata, State Minimality, Closure Spaces, Semilattices

1 Introduction

Why are state-minimal deterministic finite automata (dfas) easy to construct, whilst no efficient minimization procedure for nondeterministic finite automata (nfas) is known? Let us start with the observation that minimal dfas are built inside the category Set_f of finite sets and functions and are characterized by having no proper subautomata (reachability) and no proper quotient automata (simplicity). Nfas can be regarded as dfas interpreted in the category Rel_f of finite sets and relations, and so one might hope to build minimal nfas in the same way as minimal dfas, but now in Rel_f. However, there is a significant difference: Set_f is both finitely complete and cocomplete, yet Rel_f does not have coequalizers, i.e., canonical quotients. The lack of such canonical constructions provides evidence for the lack of *canonical* state-minimal nfas.

http://dx.doi.org/10.1016/j.entcs.2014.10.002

This suggests the following strategy: form the cocompletion of Rel_f obtained by freely adding canonical quotients, which turns out to be the category JSL_f of finite join-semilattices (see Appendix), and build minimal automata in this larger category. Every nfa may be viewed as a dfa in JSL_f via the usual subset construction. In order to obtain more efficient presentations of nfas, avoiding the full power set of states, we make use of a categorical equivalence between JSL_f and the category Cl_f of finite strict closure spaces [2]. The objects of the latter are finite sets Z equipped with a strict closure operator (i.e., an extensive, monotone and idempotent map $\mathbf{cl}_Z : \mathcal{P}Z \to \mathcal{P}Z$ preserving the empty set), and the morphisms are continuous relations, see Definitions 2.9 and 2.10 below. For example, every finite topological space induces a finite strict closure space; these closures are called *topological*.

Just as nfas may be viewed as deterministic automata interpreted in Rel_f, *nondeterministic closure automata* (ncas) are deterministic automata interpreted in Cl_f: an nca is an nfa with a strict closure operator on its set of states, continuous transition relations, an open set of final states and a closed set of initial states. Since the category Cl_f has the same relevant properties as Set_f, we derive for each regular language $L \subseteq \Sigma^*$ the existence of a unique minimal nca $\mathcal{N}(L)$ whose underlying nfa (forgetting the closure operator) accepts L. It is minimal in the sense that it has no proper subautomata (reachability) and no proper quotient automata (simplicity), and can be constructed in a way very much analogous to Brzozowski's classical construction of the minimal dfa [6]: starting with any nca $\mathcal{N}$ accepting L, one has

$$\mathcal{N}(L) = \mathsf{reach} \circ \mathsf{rev} \circ \mathsf{reach} \circ \mathsf{rev}(\mathcal{N})$$

where reach and rev are continuous versions of the reachable subset construction and the reversal operation for nfas, respectively.

The states of $\mathcal{N}(L)$ are the *prime derivatives* of L, i.e., those non-empty derivatives $w^{-1}L = \{v \in \Sigma^* : wv \in L\}$ of L that do not arise as a union of other derivatives. The underlying nfa of $\mathcal{N}(L)$ accepts L, thus it is natural to ask when this nfa is *state-minimal*. Our main result is:

If the closure of $\mathcal{N}(L)$ is topological then the underlying nfa is state-minimal.

In other words, we identify a natural class of regular languages for which *canonical* state-minimal nondeterministic acceptors exist.

Related Work. Our paper is inspired by the work of Denis, Lemay and Terlutte [7] who define a canonical nondeterministic acceptor for each regular language L. In fact, the underlying nfa of our minimal nca $\mathcal{N}(L)$ is precisely their 'canonical residual finite state automaton', and our Brzozowski construction of $\mathcal{N}(L)$ in Section 3 generalizes their construction in [7, Theorem 5.2]. The main conceptual difference is that the latter works on the level of nfas, while our construction takes the continuous structure of nondeterministic closure automata into account. We hope to convince the reader that ncas provide the proper setting in which to study these canonical nfas and their construction.

We have introduced nondeterministic closure automata in [2] where we demon-