ARTICLE IN PRESS JLAMP:60

Journal of Logical and Algebraic Methods in Programming ••• (••••) •••–•••

Contents lists available at ScienceDirect

Journal of Logical and Algebraic Methods in Programming

www.elsevier.com/locate/jlamp

Algorithms for Kleene algebra with converse ☆

Paul Brunet, Damien Pous *,[1]

Plume team – LIP, CNRS, ENS de Lyon, Université de Lyon, UMR 5668, France

ARTICLE INFO

Article history:
Received 1 October 2014
Received in revised form 5 May 2015
Accepted 17 July 2015
Available online xxxx

Keywords:
Kleene algebra
Converse
Automata
Algorithm
Regular expressions

ABSTRACT

The equational theory generated by all algebras of binary relations with operations of union, composition, converse and reflexive transitive closure was studied by Bernátsky, Bloom, Ésik and Stefanescu in 1995. In particular, they obtained its decidability by using a particular automata construction.

We show that deciding this equational theory is PSpace-complete, by providing a PSpace algorithm (the problem is easily shown to be PSpace-hard). We obtain other algorithms that are time-efficient in practice, despite not being PSpace.

Our results use an alternative automata construction, inspired by the one from Bloom, Ésik and Stefanescu. We relate those two constructions by exhibiting a bisimulation between the resulting deterministic automata, and by showing how our construction results in more sharing between states, thus producing smaller automata.

1. Introduction

In many contexts in computer science and mathematics operations of union, sequence or product and iteration appear naturally. *Kleene Algebra*, introduced by John H. Conway under the name *regular algebra* [2], provide an algebraic framework allowing to express properties of these operators, by studying the equivalence of expressions built with these connectives. It is well known that the corresponding equational theory is decidable [3], and that it is complete for both language and relation models.

As expressive as it may be, one may wish to integrate other usual operations in such a setting. Theories obtained this way, by addition of a finite set of equations to the axioms of Kleene Algebra, are called *Extensions of Kleene Algebra*. We shall focus here on one of these extensions, where an operation of *converse* is added to Kleene Algebra. The converse of a word is its mirror image (the word obtained by reversing the order of the letters), and the converse $R^\vee$ of a relation R is its reciprocal ($x\, R^\vee\, y \triangleq y\, R\, x$). This natural operation can be expressed simply as a set of equations that we add to Kleene Algebra's axioms.

The question that arises once this theory is defined is its decidability: given two formal expressions built with the connectives product, sum, iteration and converse, can one decide automatically if they are equivalent, meaning that their equality can be proven using the axioms of the theory? Bloom, Ésik, Stefanescu and Bernátsky gave an affirmative answer to that question in two articles, [4] and [5], in 1995.

☆ Extended version of the abstract presented at RAMiCS '2014 [1].

* Corresponding author.
E-mail address: damien.pous@ens-lyon.fr (D. Pous).

[1] Work partially funded by the french projects PiCoq (ANR-09-BLAN-0169-01) and PACE (ANR-12IS02001).

http://dx.doi.org/10.1016/j.jlamp.2015.07.005

However, although the algorithm they define proves the decidability result, it is too costly (in terms of time and memory consumption) to be used in concrete applications. In this paper, beside some simplifications of the proofs given in [4], we give a new and more efficient algorithm to decide this problem, which we place in the complexity class PSPACE.

The equational theory of Kleene algebra cannot be finitely axiomatised [6]. Krob presented the first purely axiomatic (but infinite) presentation [7]. Several finite quasi-equational characterisations have been proposed [8,9,7,10,11]; here we follow the one from Kozen [10].

A Kleene Algebra is an algebraic structure $\langle K, +, \cdot, ^\star, \mathbb{0}, \mathbb{1}\rangle$ such that $\langle K, +, \cdot, \mathbb{0}, \mathbb{1}\rangle$ is an idempotent semiring, and the operation * satisfies the following axioms. (Here $a \leqslant b$ is a shorthand for $a + b = b$.)

$$\mathbb{1} + a \cdot a^\star \leqslant a^\star \tag{1a}$$

$$\mathbb{1} + a^\star \cdot a \leqslant a^\star \tag{1b}$$

$$b + a \cdot x \leqslant x \Rightarrow a^\star \cdot b \leqslant x \tag{1c}$$

$$b + x \cdot a \leqslant x \Rightarrow b \cdot a^\star \leqslant x \tag{1d}$$

The quasi-variety KA consists of the axioms of an idempotent semiring together with axioms and implications (1a) to (1d). Kleene Algebras are thus *models* of KA. We shall call *regular expressions over* X, written Reg_X, the expressions built from letters of X, the binary connectives $+$ and $\cdot$, the unary connective * and the two constants $\mathbb{0}$ and $\mathbb{1}$.

Two families of such algebras are of particular interest: languages (sets of finite words over a finite alphabet, with union as sum and concatenation as product) and relations (binary relations over an arbitrary set with union and composition). KA is complete for both these models [7,10], meaning that for any $e, f \in \mathrm{Reg}_X$, $\mathrm{KA} \vdash e = f$ if and only if e and f coincide under any language (resp. relational) interpretation. This last property will be written $e \equiv_{\mathrm{Lang}} f$ (resp. $e \equiv_{\mathrm{Rel}} f$).

More remarkably, if we denote by $[\![e]\!]$ the language denoted by an expression e, we have that for any $e, f \in \mathrm{Reg}_X$, $\mathrm{KA} \vdash e = f$ if and only if $[\![e]\!] = [\![f]\!]$. By Kleene's theorem (see [3]) the equality of two regular languages can be reduced to the equivalence of two finite automata, which is decidable. Hence, the theory KA is decidable.

Now let us add a unary operation of converse to regular expressions. We shall denote by $\mathrm{Reg}_X^\vee$ the set of regular expressions with converse over a finite alphabet X. While doing so, several questions arise:

1. Can the converse on languages and on relations be encoded in the same theory?
2. What axioms do we need to add to KA to model these operations?
3. Are the resulting theories complete for languages and relations?
4. Are these theories decidable?

There is a simple answer to the first question: no. Indeed the equation $a \leqslant a \cdot a^\vee \cdot a$ is valid for any relation a (because if $(x, y) \in a$, then $(x, y) \in a$, $(y, x) \in a^\vee$, and $(x, y) \in a$, so that $(x, y) \in a \cdot a^\vee \cdot a$). But this equation is not satisfied for all languages a (for instance, with the language $a = \{x\}$, $a \cdot a^\vee \cdot a = \{xxx\}$ and $x \notin \{xxx\}$). This means that there are two distinct theories corresponding to these two families of models. Let us begin by considering the case of languages.

Theorem 1 *(Completeness of KAC$^-$). (See [4].) A complete axiomatisation of the variety* $\mathrm{Lang}^\vee$ *of languages generated by concatenation, union, star, and converse consists of the axioms of KA together with axioms (2a) to (2d).*

$$(a + b)^\vee = a^\vee + b^\vee \tag{2a}$$

$$(a \cdot b)^\vee = b^\vee \cdot a^\vee \tag{2b}$$

$$(a^\star)^\vee = (a^\vee)^\star \tag{2c}$$

$$a^{\vee\vee} = a \tag{2d}$$

We call this theory KAC$^-$; it is decidable.

As before, we write $e \equiv_{\mathrm{Lang}^\vee} f$ if e and f have the same language interpretations (for a formal definition, see the "Notation" subsection below). To prove this result, one first associates to any expression $e \in \mathrm{Reg}_X^\vee$ an expression $\mathbf{e} \in \mathrm{Reg}_{\mathbf{X}}$, where $\mathbf{X}$ is an alphabet obtained by adding to X a disjoint copy of itself. Then, one proves that the following implications hold.

$$e \equiv_{\mathrm{Lang}^\vee} f \quad \Rightarrow \quad [\![\mathbf{e}]\!] = [\![\mathbf{f}]\!] \tag{3}$$

$$[\![\mathbf{e}]\!] = [\![\mathbf{f}]\!] \quad \Rightarrow \quad \mathrm{KAC}^- \vdash e = f \tag{4}$$

(That $\mathrm{KAC}^- \vdash e = f$ entails $e \equiv_{\mathrm{Lang}^\vee} f$ is obvious; decidability comes from that of regular languages equivalence.) We reformulate Bloom et al.'s proofs of these implications in elementary terms in Section 2.1.

As stated before, the equation $a \leqslant a \cdot a^\vee \cdot a$ provides a difference between languages with converse and relations with converse. It turns out that it is the only difference, in the sense that the following theorem holds: