Ana Isabel Pereira de Moura

Pseudovarieties: idempotent-generated semigroups and representations of DA

U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Departamento de Matemática
Faculdade de Ciências da Universidade do Porto
2009

Ana Isabel Pereira de Moura

Pseudovarieties: idempotent-generated semigroups and representations of DA

Tese submetida à Faculdade de Ciências da Universidade do Porto para obtenção do grau de Doutor em Matemática

Dezembro de 2009

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Aos meus pais

Agradecimentos

Esta tese foi financiada com uma bolsa da FCT (Bolsa SFRH/BD/19720/2004) no âmbito do Programa Operacional "Ciência, Tecnologia, Inovação" (POCTI) e do Programa Operacional Sociedade da Informação (POSI) do Quadro Comunitário de Apoio III (2000/2006), comparticipado pelo Fundo Social Europeu e por fundos do Ministério da Ciência, Tecnologia e Ensino Superior (MCTES).

O meu primeiro e especial agradecimento é dirigido ao Professor Jorge Almeida, orientador dos meus trabalhos de doutoramento, pela motivação, pelo incentivo, pela inteira disponibilidade, por todas as sugestões, por me ter ensinado a investigar e a escrever matemática e, ainda, pela sua extraordinária paciência a ensinar-me a escrever matemática em inglês. Foi e continuará a ser muito gratificante trabalhar com ele.

Agradeço ao *Departamento de Matemática Pura* da *Faculdade de Ciências da Universidade do Porto*, actualmente *Departamento de Matemática*, e às pessoas que o integram (docentes, funcionários e estudantes), pelo acolhimento e pelas condições disponibilizadas para a preparação desta tese. Em particular, o meu muito obrigada aos colegas de Doutoramento e Pós-Doutoramento Ângela Cardoso, Helder Vilarinho, Marina Logares, Mário Bessa, e a Andreia Teixeira e André Souto de Ciência de Computadores, por todo o companheirismo demonstrado. Em especial, ao André Oliveira por todos os sucessos e percalços partilhados durante os longos meses no gabinete 3.89, e ao Alfredo Costa pela camaradagem, pelo apoio e pelo incentivo no momento certo.

O meu agradecimento ao *Centro de Matemática da Universidade do Porto* (CMUP) e ao projecto *ASA - Autómatos, Semigrupos e Aplicações* (PTDC/MAT/65481/2006), pelo financiamento prestado. Agradeço ainda aos membros do projecto ASA, em particular, ao Professor Manuel Delgado e ao Professor Pedro Silva, pelo prazer em trabalhar e conviver com eles.

O meu reconhecimento ao *Instituto Superior de Engenharia do Porto* e, em particular, ao *Departamento de Matemática*, pelo apoio prestado e pelos esforços envidados na renovação do meu contrato a tempo integral.

À *Universidade Estatal dos Urais*, em Ekaterinburg, e, muito especialmente, ao Professor Mikhail Volkov, pelo acolhimento e pela amizade durante os quatro meses da minha visita, спасибо!

Aos meus pais, por toda a estabilidade emocional e financeira que me proporcionaram, à minha avó Maria, à minha tia Ana, à minha prima Filipa e ao Luís, à Goretti, à restante família e amigos, por tudo aquilo que são, à Susana, por todos os amparos, muito obrigada!

E a ti, pai, por desde muito cedo me teres incutido o teu gosto pela Matemática: esta tese também é tua!

Resumo

Realçam-se nesta tese três problemas distintos na teoria de pseudovariedades de semigrupos finitos bem como diversas conexões entre eles.

Primeiramente é feito o estudo da pseudovariedade DA, cuja importância nos campos da ciência da computação e da teoria da complexidade é evidenciada no *"Diamonds are forever"* de Tesson e Thérien. Também os métodos profinitos desenvolvidos por vários investigadores têm revelado potencialidades para a resolução de problemas na teoria de semigrupos finitos. É feito o estudo do semigrupo pró-DA livre tendo como motivação o facto de ele codificar toda a informação acerca das propriedades combinatórias e algébricas dos semigrupos da pseudovariedade DA. Apresentam-se três representações das operações implícitas sobre DA: sob a forma de árvores etiquetadas de altura finita, sob a forma de árvores etiquetadas *quasi-ternárias*, e sob a forma de ordens lineares etiquetadas. Mostra-se que duas operações implícitas sobre DA são iguais se e só se tiverem a mesma representação, para qualquer uma das representações referidas. Comprimindo a representação em árvores etiquetadas *quasi-ternárias*, apresenta-se a denominada representação por DA-autómatos. Prova-se que, no caso da operação implícita sobre DA ser um ω-termo, o DA-autómato mínimo associado é finito, o que permite resolver o problema da palavra para ω-termos sobre DA. Apresenta-se ainda um algoritmo com complexidade polinomial para o cálculo do DA-autómato mínimo associado a um ω-termo. Com este intuito, são estendidas a esta pseudovariedade técnicas desenvolvidas por Almeida e Weil, e Almeida e Zeitoun aquando do estudo similar para a pseudovariedade R.

A segunda parte desta tese baseia-se no estudo do operador sobre pseudovariedades que constrói a pseudovariedade gerada pelos elementos que são gerados por idempotentes de uma pseudovariedade dada. Embora não se tenha conseguido determinar uma fronteira entre as pseudovariedades que são geradas por estes seus elementos e as que não são, apresentam-se alguns exemplos relevantes de pseudovariedades com esta propriedade, quer relacionando resultados já existentes na teoria de semigrupos, quer apresentando resultados originais envolvendo ainda métodos profinitos. É dada particular atenção às pseudovariedades J, R, L e DA.

Por último, inspirada num resultado de Tilson sobre a pseudovariedade dos semigrupos aperiódicos, é definida a propriedade da E-*localidade* sobre uma pseudovariedade. Considera-se E-*local* uma pseudovariedade V que satisfaz a condição seguinte: para um semigrupo finito, o subsemigrupo gerado pelos seus idempotentes pertence a V se e só se também pertencerem a V os subsemigrupos gerados pelos idempotentes de cada uma das suas $\mathcal{D}$-classes regulares. São apresentadas algumas condições necessárias ou suficientes para que uma pseudovariedade possua esta propriedade e é estendido o conceito às pseudoidentidades, identificando-se mais algumas condições necessárias ou suficientes para que uma pseudoidentidade seja deste tipo. Considera-se ainda um operador sobre pseudovariedades que constrói a menor pseudovariedade E-local que contém uma dada pseudovariedade e apresentam-se alguns exemplos.

Abstract

This thesis is concerned with three different problems in the theory of pseudovarieties of finite semigroups as well as several connections between them.

Firstly, the pseudovariety DA is investigated. Its importance in computational and complexity theory is evinced in the *"Diamonds are forever"* of Tesson and Thérien. The profinite methods developed by several researchers have also been shown to be very powerful in the solution of problems in finite semigroup theory. Since the free pro-DA semigroup encodes information about all algebraic and combinatorial properties of the semigroups in DA, the study of this object is developed. Three representations of implicit operations over DA are presented: the first one by means of finite-height labeled trees; the second one by means of *quasi-ternary* labeled trees; and the third one by means of labeled linear orderings. It is shown that two implicit operations over DA are equal if and only if they have the same representation, for any of the representations. *Wrapping* the *quasi-ternary* labeled trees we obtain the so-called representation by DA-automata. It is proved that the representation by the minimal DA-automaton is finite if and only if the implicit operation it represents is an ω-term, which solves the word problem for ω-terms over DA. To complete the result, an algorithm with polynomial-time complexity to compute the minimal DA-automaton associated to an ω-term is presented. For this purpose, the techniques developed by Almeida and Weil, and Almeida and Zeitoun to solve the analogous problem for the pseudovariety R are extended to this pseudovariety.

The second part of this thesis is based on the study of the operator on pseudovarieties that constructs the pseudovariety generated by the idempotent-generated elements of a given pseudovariety. Although the boundary between the pseudovarieties that are generated by these elements and those that are not has not yet been determined, several relevant examples of pseudovarieties with this property are presented, either by relating existing results in semigroup theory, or by presenting original results involving also profinite methods. Particular attention is paid to the pseudovarieties J, R, L and DA.

Finally, inspired by a property observed by Tilson in the pseudovariety of aperiodic semigroups, the property of E-*locality* on a pseudovariety is defined. A pseudovariety V is said to be E-*local* if it satisfies the following condition: for a finite semigroup, the subsemigroup generated by its idempotents belongs to V if and only if so do the subsemigroups generated by the idempotents in each of its regular $\mathcal{D}$-classes. Several necessary or sufficient conditions for a pseudovariety to have this property are presented and this concept is extended to pseudoidentities, where some necessary or sufficient additional conditions for a pseudoidentity to be of this type are identified. The operator on pseudovarieties which associates the smallest E-local pseudovariety that contains a given pseudovariety is also considered, and some examples are presented.

Résumé

On met en évidence dans cette thèse trois problèmes différents dans la théorie des pseudovariétés des semigroupes finis ainsi que plusieurs liaisons.

En premier, on fait l'étude de la pseudovariété DA, dont l'importance est soulignée dans le *"Diamonds are forever"* de Tesson et Thérien dans les domaines de l'informartique et de la théorie de la complexité. Les méthodes profinies développées par plusieurs chercheurs ont aussi montré un potenciel pour résoudre des problèmes dans la théorie des semigroupes finis. On étudie le sémigroupe pro-DA libre face à la motivation venant du fait qu'il contient toutes les informations sur les propriétés algébriques et combinatoires des semigroupes de la pseudovariété DA. On présente trois représentations des opérations implicites sur DA: sous la forme d'arbres étiquetés à hauteur finie, sous la forme d'arbres étiquetés *quasi-ternaires*, et sous la forme d'ordres linéaires étiquetés. On prouve que deux opérations implicites sur DA sont égaux si elles ont la même représentation, pour chacune des représentations. Si l'on comprime la représentation pour arbres étiquetés *quasi-ternaires*, on obtient une représentation par DA-automates. Au cas où l'opération implicite sur DA est un ω-terme, on prouve que le DA-automate minimum associé est fini et, par conséquent, le problème du mot pour ω-termes sur DA est résolu. On présente en plus un algorithme avec complexité polynomiale pour calculer le DA-automate minimum associé à un ω-terme. À ce but, on étend à la pseudovariété DA les techniques développées par Almeida et Weil, et par Almeida et Zeitoun dans leurs études similaires pour la pseudovariété R.

En second, cette thèse se rapporte à l'étude de l'opérateur qui associe à une pseudovariété donnée la pseudovariété engendrée par ses membres qui sont engendrés par leurs idempotents. Quoiqu'on n'arrive pas à déterminer la frontière entre les pseudovariétés qui ont cette propriété et celles qui ne l'ont pas, on présente quelques exemples pertinents de pseudovariétés avec cette propriété, soit en appliquant des résultats connus de la théorie des semigroupes, soit par des résultats nouveaux utilisant encore des méthodes profinies. On porte une attention particulière aux pseudovariétés J, R, L et DA.

Finalement, inspirée par un résultat de Tilson sur la pseudovariété des semigroupes apériodiques, on définit la propriété de la E-*localité* sur une pseudovariété. Une pseudovariété V est E-*local* si elle satisfait la condition suivante: pour un semigroupe fini, le sous-semigroupe engendré par ses idempotents appartient à V si et seulement si les sous-semigroupes engendrés par les idempotents de chacune de leurs $\mathcal{D}$-classes régulières appartiennent aussi à V. On présente des conditions nécessaires ou suffisantes pour qu'une pseudovariété ait cette propriété et on étend cette notion aux pseudoidentités. On identifie d'autres conditions nécessaires ou suffisantes pour q'une pseudoidentité soit de ce type. On considère aussi l'opérateur sur les pseudovariétés qui construit la plus petite pseudovariété E-local qui contient une certaine pseudovariété et on présente quelques exemples.

Contents

Introduction

The first significant result on finite semigroups is Suschkewitsch's structure theorem for finite simple semigroups [**45**]. This was generalized in 1940 by Rees [**37**] who gave a structure theorem for completely 0-simple semigroups in which the finiteness assumption is weakened to a chain condition on ideals. Many other structure results were obtained following the key introduction of Green's relations [**19**], often based on Rees' Theorem for the "local" structure meaning the structure of the principal factors. The algebraic theory of semigroups that evolved partly along these lines up to the nineteen sixties is presented in the two volumes of Clifford and Preston [**15, 16**]. Results specifically about finite semigroups started to be proved again around the same time.

In the middle of the nineteen fifties, with logic and computer science, the theory of finite automata and the study of languages recognized by them enter the scene. Kleene calls such languages *regular events* and in 1956 [**23**], motivated by modeling of human neural activity, he presents the fundamental theorem of the theory of finite automata: the regular languages are those which admit rational expressions in finite languages using only the Boolean finitary operations, the operation of concatenation and the star operation (the so called *rational languages*). Following this result, Schützenberger [**41**] establishes that the *star-free languages* are those whose syntactic semigroups, which can be automatically constructed from a rational expression [**33**], are finite and all its subgroups are trivial. Schützenberger thus begins the study of algebraic properties of syntactic semigroups in order to understand combinatorial properties of rational languages. Later, Brzozowski and Simon [**14**], and McNaughton [**27**] prove that a language is *locally testable* if and only if its syntactic semigroup is finite and all its submonoids are semilattices, and Simon [**43**] also shows that a rational language is *piecewise testable* if and only if its syntactic semigroup is $\mathcal{J}$-trivial.

According to Satoh, Yama and Tokizawa [**40**] there exist about 1.8 billion of non-isomorphic nor anti-isomorphic semigroups of order 8. Does the classification of finite semigroups up to isomorphism is apparently unfeasible. In 1976, Eilenberg [**17**] introduces the concept of *varieties of rational languages*, of which the classes of languages of the above results are examples, and he establishes a one-to-one correspondence between these varieties and certain classes of finite semigroups, called *pseudovarieties*, which consist of classes of finite semigroups that are closed under the formation of subsemigroups, homomorphic images, and finite direct products. Since then the interest in the theory of pseudovarieties has been growing steadily.

On the other hand, in the middle of the nineteen sixties, Krohn and Rhodes [**24**] introduce the notion of division and state the *Prime Decomposition Theorem*: every finite semigroup divides a wreath product in which the factors are, alternately, finite (permutation) groups and finite aperiodic (transformation) semigroups. The least number of group factors in such a decomposition is said to be the *(group) complexity* of the semigroup [**24**]. The question then arises of how to build an algorithm that computes the group complexity of any finite semigroup; even the existence of such an algorithm is still unknown and this is a major open problem in the theory of finite semigroups. A detailed description of the Krohn-Rhodes complexity theory in various stages of development is given by Rhodes and Tilson in [**10**], by Tilson in the last two chapters of Eilenberg [**17**] and, most recently, by Rhodes and Steinberg [**39**].

The results in Universal Algebra of finite structures led to important developments in the theory of finite semigroups. Already the founder of Universal Algebra, Garrett Birkhoff, considered in 1937 the notion of *profinite algebra* namely of a projective limit of finite discrete algebras [**13**]. Unlike the finite world, in profinite world there exist free objects. For a pseudovariety (of semigroups) V, we denote by $\overline{\Omega}_A\mathsf{V}$ the free pro-V semigroup, that is the projective limit of all semigroups of V generated by an alphabet A. This semigroup encodes information about all algebraic and combinatorial properties of the elements of V, which strongly motivates the study of these relatively free objects. Until now very little is known about them. A *pseudoidentity* is a formal equality of elements of a free profinite semigroup. Reiterman [**38**] proves that the pseudovarieties are exactly the classes defined by pseudoidentities. Since the nineteen nineties, with the development of profinite methods by various authors led essentially by Almeida (see [**1, 3, 4**]), the theory of finite semigroups has made remarkable progress. This new approach attempts to answer questions in the theory of finite semigroups, for example by solving *word problems* for free objects and by finding *bases of pseudoidentities* for pseudovarieties.

A typical problem in the theory of finite semigroups is called the *membership problem* for a pseudovariety, which consists in determining whether a given finite semigroup belongs to it. A pseudovariety is said to be *decidable* if so is its membership problem. Many important pseudovarieties can be built from others by applying natural operators. It is known that the decidability of the membership problem is not preserved under many such operators on pseudovarieties (cf. [**12**]), which leads to a case by case investigation of decidability of pseudovarieties obtained by the application of such operators. In an attempt to establish the decidability of semidirect products of pseudovarieties of semigroups, Almeida and Steinberg [**6**] introduce the notion of *tameness*, a property that they then thought might lead to a substantial reduction of the Krohn-Rhodes problem. To prove σ-tameness (tameness relatively to a given signature σ) of a pseudovariety we have to solve two problems: the σ-word problem and another one that consists in proving that, if a system of equations with rational constraints has a solution in any semigroup of the pseudovariety, then it also has a solution in σ-terms. From the results of Ash [**11**], it follows that the pseudovariety of all finite groups is tame relatively to the signature consisting of multiplication and pseudoinversion. This result was rediscovered independently as a deep result in model theory [**20**], which further motivates the study of the property of tameness in other pseudovarieties. Several researchers have considered and established various forms of tameness for specifics pseudovarieties.

In 1976, Schützenberger [**42**] shows the importance of the pseudovariety DS, consisting of all finite semigroups whose regular $\mathcal{D}$-classes are subsemigroups. He describes the varieties of languages corresponding to various subpseudovarieties of DS in terms of closure operators for restricted forms of the concatenation product. The class DA of all finite semigroups whose regular elements are idempotents deserves specially attention. More recently, in their *"Diamonds are forever"*, Tesson and Thérien [**46**] show that languages whose syntactic monoid lies in DA have remarkable characterizations, algebraic, combinatorial, logical and even automata-theoretical, that lead to the solution of problems in computational and complexity theory. We recommend [**46**] for examples of such applications. This strongly motivates the study of the pseudovariety DA.

This thesis is written in the form of four papers, presented as appendices. There is also an extra appendix contained some computer programs. In the first two papers we investigate the pseudovariety DA, with the study of its free profinite object. In Paper 1 we give three representations of the free pro-DA semigroup on a finite alphabet: the first one by means of finite-height labeled trees; the second one by means of *quasi-ternary* labeled trees; and the third one by means of labeled linear orderings. To obtain such results, we extend to DA some techniques used by Almeida and Weil [**7**], and Almeida and Zeitoun [**8, 9**] introduced for the description of free profinite semigroups over R. We also develop the representation by *quasi-ternary* labeled trees, resulting from *wrapping* DA-trees. We obtain a representation by DA-automata, which is not necessarily finite. In Paper 2, we show that the representation by the minimal DA-automaton is finite if and only if the element it represents is an ω-term. This allows us to compute the minimal DA-automaton of a given ω-term, and the paper presents a polynomial-time algorithm that performs such a computation. It is then possible to decide whether two ω-terms are equal over all elements of the pseudovariety DA by testing the equality of the corresponding minimal DA-automata and, therefore, the word problem for ω-terms over DA is solved in polynomial-time. In Appendix 5, we present the complete programming of our algorithm in Python to compute and to visualize the minimal DA-automaton of an ω-term.

In order to acquire and consolidate knowledge in the theory of finite semigroups, in particular, in the theory of pseudovarieties and its relationship with the varieties of rational languages, the work for the preparation of this thesis began with the detailed study of Pin's book [**36**]. One of the exercises led to the characterization of the so called E-*local pseudovarieties*, the exercise consisting in establishing this property for the pseudovariety of aperiodic semigroups. This result was first observed by Tilson [**47**] when he established a method for calculating the complexity of a finite semigroup with a maximum of two non-zero $\mathcal{D}$-classes. We say that a pseudovariety V is E-*local* if it satisfies the following property: for a finite semigroup, the subsemigroup generated by its idempotents belongs to V if and only if so do the subsemigroups generated by the idempotents in each of its regular $\mathcal{D}$-classes. In Paper 4, we present several necessary or sufficient conditions for a pseudovariety to be E-local and we extend this concept to pseudoidentities, calling E-*local* a pseudoidentity which defines an E-local pseudovariety. Various characterizations of pseudoidentities with this property can be easily deduced, and we enlarge this study by presenting some more necessary or sufficient conditions for pseudoidentities to be of this type. Finally, we introduce a new operator which associates to a pseudovariety the smallest E-local pseudovariety containing it. For the results obtained in Paper 4, we use properties of idempotent-generated subsemigroups (in particular, the work of Fitz-Gerald [**18**]) and *blocks* of such subsemigroups.

It is in the attempt to characterize E-local pseudovarieties that we engaged in the study of idempotent-generated semigroups. And one question that immediately arises is how to determine which pseudovarieties are generated by their idempotent-generated members, which is the subject of Paper 3, done in collaboration with Almeida. We apply, in parallel, three different approaches to the problem of proving that certain pseudovarieties have that property. On the one hand, there are the works of Petrich and Reilly [**35**], Pastijn and Yan [**31, 32**], and Petrich [**34**] using embeddings of semigroups into idempotent-generated Rees matrix semigroups, which allow us to conclude that the pseudovarieties $\bar{\mathsf{H}}$, CR and CS have the property in question. On the other hand, several authors have studied the property of a certain transformation semigroup be idempotent-generated. In particular, we use the following results: every finite semigroup embeds in a finite regular idempotent-generated semigroup [**21**]; the semigroup of contractive full transformations on a set and the semigroup of contractive and order-preserving full transformations on a set are idempotent generated [**26, 22**]. Combining these last two results with the representation theorems for $\mathcal{R}$-trivial monoids and $\mathcal{J}$-trivial monoids due to Pin [**36**] and Straubing [**44**], respectively, we obtain that the pseudovarieties R, L and J also have the property in question. In the third approach to this question, we use profinite methods to obtain division properties of the semigroups of J, R, L, and also DA into idempotent-generated semigroups in the same pseudovariety. This new approach provides a significant improvement in terms of the reduction of the *generator rank* and *idempotent-generator rank* of the idempotent-generated semigroup, and such an analysis is also carried out at the end of Paper 3.

Several results obtained in Paper 3 are crucial to reach some of the conclusions in Paper 4. Moreover, it turns out that, chronologically the work for this thesis began with the problem concerned in Paper 4, from where we were led to the main question of Paper 3, which in turn raised the questions on the pseudovariety DA addressed in Papers 1 and 2.

Because this thesis is composed of several papers, the chapters which constitute them, and which appear in the form of appendices, are essentially self-contained, each having its own introduction and bibliography, which might repeat some of the content referred in this introduction.

All papers which compose this thesis have been submitted to international journals and are available as preprints from *Centro de Matemática da Universidade do Porto* (see [**28, 30, 5, 29**]).

Finally, we suggest as main references to the theory of finite semigroups, pseudovarieties, and profinite semigroups, the books of Eilenberg [**17**], Lallement [**25**], Pin [**36**], Almeida [**1**], and Rhodes and Steinberg [**39**].

Conclusion

From all work developed in this thesis, we emphasize the following observations, conclusions and problems.

Representations of the free profinite object over DA and the word problem for ω-terms over DA

For an implicit operation $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$, Almeida [**2**], and Trotter and Weil [**48**] define the *central basic factorization of* w, $\mathsf{CBF}(w)$, as a factorization of one of the following forms:

(i) *standard form:* $w = \alpha a\gamma b\beta$ with $a, b \in A$, $\alpha, \beta, \gamma \in \overline{\Omega}_A\mathsf{DA}$, $a \notin c(\alpha)$, $b \notin c(\beta)$ and $c(\alpha a) = c(b\beta) = c(w)$;

(ii) *overlapped form:* $w = \alpha b\gamma a\beta$ with $a, b \in A$, $\alpha, \beta, \gamma \in \overline{\Omega}_A\mathsf{DA}$, $a \notin c(\alpha b\gamma)$, $b \notin c(\gamma a\beta)$ and $c(\alpha b\gamma a) = c(b\gamma a\beta) = c(w)$;

(iii) *degenerate form:* $w = \alpha a\beta$ with $a \in A$, $\alpha, \beta \in \overline{\Omega}_A\mathsf{DA}$, $a \notin c(\alpha)$, $a \notin c(\beta)$ and $c(\alpha a) = c(a\beta) = c(w)$.

This central basic factorization of $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$ exists and is unique by Almeida [**2**]. Moreover, we may iterate this factorization. We do this in two different ways: in the first one we iterate the factorization while the content of the central factor (if it exists) does not decrease and in the second one we iterate until the central factor (if it exists) becomes 1. Results from Almeida [**2, 1**], and Almeida and Weil [**7**] allow us to conclude that, in case the iterated central basic factorization $\mathsf{ICBF}(w)$ is an infinite product, then it converges. We obtain two different iterated central basic factorizations that we denote by $\mathsf{I}_1\mathsf{CBF}(w)$ and $\mathsf{I}_2\mathsf{CBF}(w)$, respectively. Finally, we apply successively the corresponding factorization to each factor in $\mathsf{ICBF}(w)$, which is a finite process because the contents of the factors involved strictly decrease.

From $\mathsf{I}_1\mathsf{CBF}(w)$ we may represent an implicit operation w on DA by a tree with the following form. It has a root and its direct progeny has ordered vertices in a one-to-one correspondence with the factors of $\mathsf{I}_1\mathsf{CBF}(w)$. The odd vertices, reading from left to right and from right to left in case the progeny is infinite, correspond to the factors α_i, β_i and γ of $\mathsf{I}_1\mathsf{CBF}(w)$. We iterate this process on these vertices. The other vertices, called leaves, correspond to the distinguished letters and are labeled by them. We obtain a representation by a labeled tree of finite height. The set of these trees is denoted by $T_1(A)$ and is in bijection with the elements of $\overline{\Omega}_A\mathsf{DA}$.

The second representation is by *quasi-ternary* labeled trees and consists of the following. The tree corresponding to an implicit operation w has a root labeled with the pair of distinguished letters of $\mathsf{CBF}(w)$ (or with just one letter, in case of the degenerate form). The direct progeny has ordered vertices in a one-to-one correspondence with the factors α, β and γ of $\mathsf{CBF}(w)$. We iterate this process on each of these vertices. This type of representation is in the set $T_2(A)$ that we proved that is also in bijection with $\overline{\Omega}_A\mathsf{DA}$.

If, in the latter representation, we identify vertices with the same attached subtrees, we obtain the so-called representation by minimal DA-automata. We prove that two implicit operations over DA are equal if and only if they have isomorphic minimal DA-automata.

For every tree in $T_1(A)$, we consider the labeled linear ordering obtained by ordering the set of leaves from left to right. This linear ordering belongs to $\mathbf{rLO}^*(A)$ and we prove that the sets $T_1(A)$ and $\mathbf{rLO}^*(A)$ are in bijection. We may have an analogous representation by labeled linear orderings that is in bijection with the set $T_2(A)$. In fact, we observe how to relate all these representations of an implicit operation over DA.

If we work with an ω-term w, we prove that the set of factors containing w and closed under taking factors of the central basic factorization is finite. It follows that the minimal DA-automaton associated to an ω-term is also finite. We present an algorithm that effectively computes a finite DA-automaton for every ω-term and, using existing tools, we minimize it. So, the word problem for ω-terms over DA is solved. The complexity of this computation does not exceed $\mathcal{O}((nK)^4)$. We do not know whether this upper bound for the complexity of the word problem is optimal.

Idempotent-generated semigroups and pseudovarieties

Let $\_{\mathsf{E}}$ be the operator that constructs the pseudovariety V_E generated by all idempotent-generated elements of a given pseudovariety V. It is an idempotent increasing operator such that, given a pseudovariety V, the equations $\mathsf{X}_\mathsf{E} = \mathsf{V}_\mathsf{E}$ and $\mathsf{EX} = \mathsf{EV}$ in the variable X are equivalent and the class of its solutions is the interval $[\mathsf{V}_\mathsf{E}, \mathsf{EV}]$.

The main problem addressed in this paper is to determine which pseudovarieties satisfy the equality $\mathsf{V} = \mathsf{V}_\mathsf{E}$. Although some of the results obtained are not new, we present new approaches that are of interest. We use the embedding into the idempotent-generated Rees matrix semigroup of Petrich [**34**] and we observe that every pseudovariety of the form $\bar{\mathsf{H}}$ satisfies the equality. Considering idempotent-generated subsemigroups of Petrich's Rees matrix semigroup, we prove that the pseudovarieties CS and CR also satisfy the equality.

Using profinite methods, we obtain the following theorem:

Theorem 1. *Let V be a pseudovariety such that, for every n, there exists m such that $\overline{\Omega}_n\mathsf{V}$ embeds in $\overline{\langle X\rangle}$ for some $X \subseteq E(\overline{\Omega}_m\mathsf{V})$. Then $\mathsf{V}_\mathsf{E} = \mathsf{V}$.*

Moreover, the representations of the free profinite semigroups over J, R and DA obtained, respectively, by Almeida [**1**], Almeida and Weil [**7**], and in the first part of this thesis enable us to establish the following embedding from $\overline{\Omega}_n\mathsf{V}$ into an idempotent-generated subsemigroup of $\overline{\Omega}_{n+1}\mathsf{V}$, for $\mathsf{V} \in \{\mathsf{J}, \mathsf{R}, \mathsf{L}, \mathsf{DA}\}$:

$$\begin{array}{rccc} \mu_{\mathsf{V}}: & \overline{\Omega}_n\mathsf{V} & \to & \overline{\Omega}_{n+1}\mathsf{V} \\ & x_i & \mapsto & x_i^{\omega}y^{\omega}, \end{array}$$

which, combined with the above theorem, proves that the equality $\mathsf{V} = \mathsf{V}_{\mathsf{E}}$ holds for the pseudovarieties J, R, L and DA.

Finally, we prove that every pseudovariety in the interval $[\mathsf{J}, \mathsf{DS}]$ has infinite *generator rank* and *idempotent generator rank*. We also determine a lower bound for the idempotent generator rank of the subpseudovarieties generated by all n-generated members of any pseudovariety in the interval $[\mathsf{J}, \mathsf{DA}]$. The above embedding allows us to conclude that this lower bound is the exact value in the case of the pseudovarieties J, R, L and DA.

It remains an open problem to define the boundary between the pseudovarieties which satisfy the equality $\mathsf{V} = \mathsf{V}_{\mathsf{E}}$ and the ones that do not satisfy it. The particular case of the pseudovariety DS motivates the research for a corresponding representation result, which at present is yet inaccessible.

E-local pseudovarieties

We say that a pseudovariety V is E-*local* if it satisfies the following property: given $S \in \mathsf{S}$, $\langle E(S)\rangle \in \mathsf{V}$ if and only if $\langle E(D)\rangle \in \mathsf{V}$, for every regular $\mathcal{D}$-class D of S.

An easy observation, that follows from properties of subsemigroups generated by subsets of idempotents of a finite semigroup S, is that, if V is an E-local pseudovariety, then any pseudovariety in the interval $[\mathsf{V}_{\mathsf{E}}, \mathsf{EV}]$ is also E-local.

Combining the technique of Fitz-Gerald [**18**] that consists in writing a product of idempotents of $\langle E(S)\rangle$ as a product of idempotents of $\langle E(D)\rangle$, for a regular $\mathcal{D}$-class D of S, with some easy results involving blocks of subsemigroups generated by subsets of idempotents of S, enable us to reach several results.

Firstly, the families of pseudovarieties BV, DV and $\bar{\mathsf{H}}$ are E-local, where V and H are, respectively, an arbitrary pseudovariety and an arbitrary pseudovariety of groups. It follows immediately that the pseudovarieties J, R, L and A satisfy the same property. Moreover, with the equality $\mathsf{J}_{\mathsf{E}} = \mathsf{J}$ obtained in this thesis, we prove that J is the smallest pseudovariety having this property. These results allow us to readily determine if a pseudovariety V is E-local, for many pseudovarieties V.

Theorem 3.10 of Paper 4 presents several sufficient conditions for a pseudovariety to be E-local in terms of special interactions of the operators $\mathsf{E}\_$ and $\mathsf{B}\_$ on a given pseudovariety. It is also showed that these conditions are even necessary in case the pseudovariety is contained in EDS.

Since there are pseudovarieties which are not E-local, we consider the operator $\_^{\mathsf{E}}$ that associates to a pseudovariety V the smallest E-local pseudovariety V^{E} containing it. In particular, we obtain the following examples: $(\mathsf{CS})^{\mathsf{E}} = (\mathsf{DS})_{\mathsf{E}}$, $(\mathsf{CR})^{\mathsf{E}} = (\mathsf{DS})_{\mathsf{E}}$ and $(\mathsf{DS})_{\mathsf{E}} \subseteq (\mathsf{LG})^{\mathsf{E}} \subseteq \mathsf{DS}$, which provides additional motivation for the calculation of $(\mathsf{DS})_{\mathsf{E}}$.

Finally, we extend this property to pseudoidentities and we call E-*local* a pseudoidentity which defines an E-local pseudovariety. Besides the characterizations of E-local pseudoidentities that immediately follow from Theorem 3.10 of Paper 4, we also present several necessary or sufficient conditions for a pseudoidentity to be E-local.

On one hand, we use the equality $\mathsf{V} = \mathsf{V}_{\mathsf{E}}$ for the pseudovarieties J, R, L and DA, and we relate the E-locality of pseudoidentities of the form $u = v$, where $\mathrm{first}(u) \neq \mathrm{first}(v)$ or $\mathrm{last}(u) \neq \mathrm{last}(v)$, with the condition $\mathsf{V} \subseteq [\![u = v]\!]$, where V is one of such pseudovarieties.

On the other hand, we use again the result from Fitz-Gerald [**18**] to obtain another sufficient condition for a pseudoidentity $u = v$ to be E-local: $u, v \in \overline{\langle X \rangle}$ with all elements of $X \subseteq \overline{\Omega}_A\mathsf{S}$ lying in a same regular $\mathcal{D}$-class of $\overline{\Omega}_A\mathsf{S}$. We do not know if every E-local pseudovariety is defined by a set of pseudoidentities of this form although we can show that many examples of E-local pseudovarieties do enjoy this property.

Bibliography

[1] J. Almeida, *Finite Semigroups and Universal Algebra*, World Scientific, Singapore, 1994. English translation.

[2] ————, *A syntactical proof of locality of* DA, Int. J. Algebra and Comp. **6** (1996) 165–177.

[3] ————, *Finite semigroups: an introduction to a unified theory of pseudovarieties*, in Semigroups, Algorithms, Automata and Languages, J.-E. P. G. M. S. Gomes and P. V. Silva, eds., Singapore, 2002, World Scientific, 3–64.

[4] ————, *Profinite semigroups and applications*, in Structural Theory of Automata, Semigroups, and Universal Algebra, V. B. Kudryavtsev and I. G. Rosenberg, eds., vol. 207 of NATO Science Series II: Mathematics, Physics and Chemistry, New York, 2005, Springer, 1–45. Proceedings of the NATO Advanced Study Institute on Structural Theory of Automata, Semigroups and Universal Algebra, Montréal, Québec, Canada, 7-18 July 2003.

[5] J. Almeida and A. Moura, *Idempotent-generated semigroups and pseudovarieties*, Tech. Rep. CMUP 2009-42, Univ. of Porto, 2009. `http://cmup.fc.up.pt/cmup/v2/include/filedb.php?id=292&table=publicacoes&field=file`.

[6] J. Almeida and B. Steinberg, *Syntactic and global semigroup theory: A synthesis approach*, Birkhauser, 2000.

[7] J. Almeida and P. Weil, *Free profinite* R*-trivial monoids*, Int. J. Algebra and Comput. **7** (1997) 625–671.

[8] J. Almeida and M. Zeitoun, *The equational theory of ω-terms for finite* R*-trivial semigroups*, in Proc. of the Workshop on Semigroups and Languages, Lisbon 2002, V. H. F. I. M. Araújo M. J. J. Branco and G. M. S. Gomes, eds., World Scientific, 2004, 1–22.

[9] ————, *An automata-theoretic approach to the word problem for ω-terms over* R, Theor. Comp. Sci. **370** (2007) 131–169.

[10] M. Arbib, *Algebraic Theory of Machines, Languages and Semigroups*, Academic Press, New York, 1968.

[11] C. J. Ash, *Inevitable graphs: a proof of the type $\{II\}$ conjecture and some related decision procedures*, Int. J. Algebra and Comput. **1** (1991) 127–146.

[12] K. Auinger and B. Steinberg, *On the extension problem for partial permutations*, Proc. Amer. Math. Soc. **131** (2003) 2693–2703.

[13] G. Birkhoff, *Moore-Smith convergence in general topology*, Annals of Math. **38** (1937) 39–56.

[14] J. A. Brzozowski and I. Simon, *Characterizations of locally testable events*, Discrete Math. **4** (1973) 243–271.

[15] A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups*, vol. I, Amer. Math. Soc., Providence, R.I., 1961.

[16] ————, *The Algebraic Theory of Semigroups*, vol. II, Amer. Math. Soc., Providence, R.I., 1967.

[17] S. Eilenberg, *Automata, Languages and Machines*, vol. B, Academic Press, New York, 1976.

[18] D. G. Fitz-Gerald, *On inverses of products of idempotents in regular semigrups*, J. Aust. Math. Soc. **13** (1972) 335–337.

[19] J. Green, *On the structure of semigroups*, Annals of Math. **54** (1951) 163–172.

[20] B. Herwig and D. Lascar, *Extending partial automorphisms and the profinite topology on free groups*, Trans. Amer. Math. Soc. **352** (2000) 1985–2021.

[21] J. M. Howie, *The subsemigroup generated by the idempotents of a full transformation semigroup*, J. London Math. Soc. **41** (1966) 707–716.

[22] ————, *Products of idempotents in certain semigroups of order-preserving transformations*, Proc. Edinburgh Math. Soc. **17** (1971) 223–236.

[23] S. C. Kleene, *Representations of events in nerve nets and finite automata*, in Automata Studies, C. E. Shannon, ed., vol. 34, Princeton, N.J., 1956, Princeton University Press, 3–41.

[24] K. Krohn and J. Rhodes, *Algebraic theory of machines. I. Prime decomposition theorem for finite semigroups and machines*, Trans. Amer. Math. Soc. **116** (1965) 450–464.

[25] G. Lallement, *Semigroups and Combinatorial Applications*, Wiley, New York, 1979.

[26] A. Laradji and A. Umar, *On certain finite semigroups of order-decreasing transformations I*, Semigroup Forum **69** (2004) 184–200.

[27] R. McNaughton, *Algebraic decision procedures for local testability*, Math. Syst. Theory **8** (1974) 60–76.

[28] A. Moura, *Representations of the free profinite object over* DA, Tech. Rep. CMUP 2009-26, Univ. of Porto, 2009. `http://cmup.fc.up.pt/cmup/v2/include/filedb.php?id=275&table=publicacoes&field=file`.

[29] ————, E*-local pseudovarieties*, Tech. Rep. CMUP 2009-45, Univ. of Porto, 2009. `http://cmup.fc.up.pt/cmup/v2/include/filedb.php?id=296&table=publicacoes&field=file`.

[30] ————, *The word problem for* ω*-terms over* DA, Tech. Rep. CMUP 2009-27, Univ. of Porto, 2009. `http://cmup.fc.up.pt/cmup/v2/include/filedb.php?id=276&table=publicacoes&field=file`.

[31] F. Pastijn, *Embbeding semigroups in semibands*, Semigroup Forum **14** (1977) 247–263.

[32] F. Pastijn and X. Yan, *Varieties of semigroups and varieties of completely regular semigroups closed for certain extensions*, J. Algebra **163** (1994) 777–794.

[33] D. Perrin, *Handbook of Theoretical Computer Science*, vol. B: Formal Models and Semantics, Elsevier, Amsterdam, Jan van Leeuwen, ed., 1990.

[34] M. Petrich, *Embbeding semigroups into idempotent generated ones*, Monatsh. Math. **141** (2004) 315–322.

[35] M. Petrich and N. R. Reilly, *Completely Regular Semigroups*, Canadian Math. Soc. Series of Monographs and Advanced Texts 23, Wiley, New York, 1999.

[36] J.-E. Pin, *Varieties of Formal Languages*, Plenum, London, 1986. English translation.

[37] D. Rees, *On semi-groups*, Proc. Cambridge Phil. Soc. **36** (1940) 387–340.

[38] J. Reiterman, *The Birkhoff theorem for finite algebras*, Algebra Universalis **14** (1982) 1–10.

[39] J. Rhodes and B. Steinberg, *The q-theory of finite semigroups*, Springer, New York, 2009.

[40] S. Satoh, K. Yama, and T. M., *Semigroups of order 8*, SF **49** (1994) 7–30.

[41] M. P. Schützenberger, *On finite monoids having only trivial subgroups*, Inform. and Control **8** (1965) 190–194.

[42] ————, *Sur le produit de concaténation non ambigu*, Semigroup Forum **13** (1976) 47–75.

[43] I. Simon, *Piecewise testable events*, in Proc. 2nd GI Conf., vol. 33 of Lect. Notes in Comput. Sci., Berlin, 1975, Springer, 214–222.

[44] H. Straubing, *On finite J-trivial monoids*, Semigroup Forum **19** (1980) 107–110.

[45] A. K. Suschekewitsch, *Über die endlichen Gruppen ohne das Gesetz der eineutigen Umkehrbarkeit*, Math. Ann. **99** (1928) 30–50.

[46] P. Tesson and D. Thérien, *Diamonds are forever: the variety* DA, in Semigroups, Algorithms, Automata and Languages, World Scientific, 2002, 475–499.

[47] B. Tilson, *Complexity of two J-class semigroups*, Advances Math. **11** (1973) 215–237.

[48] P. Trotter and P. Weil, *The lattice of pseudovarieties of idempotent semigroups and a non-regular analogue*, Algebra Univers. **37** (1997) 491–526.

Paper 1

REPRESENTATIONS OF THE FREE PROFINITE OBJECT OVER DA

ANA MOURA

Abstract. In this paper, we extend to DA some techniques developed by Almeida and Weil, and Almeida and Zeitoun for the pseudovariety R to obtain representations of the implicit operations on DA: by labeled trees of finite height, by *quasi-ternary* labeled trees, and by labeled linear orderings. We prove that two implicit operations are equal over DA if and only if they have the same representation, for any of the three representations. We end the paper by relating these representations.

1. Introduction

The importance of the study of pseudovarieties of finite monoids became evident with Eilenberg [10] in the middle of the 1970's, who established the correspondence between varieties of rational languages and those classes of finite monoids. Some years later, Reiterman [11] showed that every pseudovariety of finite monoids is defined by some set of finitary pseudoidentities, which are equalities between implicit operations. As implicit operations over a pseudovariety of monoids contain information on the structure of the finite monoids in the pseudovariety, it became important to develop the study of the set of implicit operations over a pseudovariety V on a finite alphabet A, $\overline{\Omega}_A\mathsf{V}$, which has the structure of a pro-V monoid.

Schützenberger [13] noted the interest of the study of the pseudovariety DS and Almeida and Weil [7] stated that for this pseudovariety and its subpseudovarieties it should be easy to make a description of the free profinite object. In fact, Almeida [1] factorized each element of the free profinite monoid over J in terms of component projections and idempotents and Azevedo [9] proved that a similar kind of factorization could be implemented to any subpseudovariety of DS, although it has not yet been discovered a canonical form of such factorizations in this pseudovariety. Almeida and Weil [7] gave two complementary descriptions of the monoid of implicit operations on R, one by labeled ordinals and the other by labeled infinite trees of finite depth. They did a similar study for the pseudovariety DRG. On the other hand, in their recent work, Almeida, Costa and Zeitoun [5, 6] presented structural properties of the free profinite semigroup over A.

In their *"Diamonds are forever"*, Tesson and Thérien [14] showed that languages whose syntactic monoid lies in DA have powerful characterizations, from combinatorial ones, to logical and even automata-theoretical ones. This characterizations

2000 *Mathematics Subject Classification.* 20M05; 20M07; 20M35.

Key words and phrases. Finite monoid; pseudovariety; profinite monoid; implicit operation; aperiodic; regular D-class.

are useful to solve problems in computational and complexity theory and the authors gave examples of such applications. They emphasize that these problems are efficiently solvable when the syntactic monoid, besides being aperiodic, is in DA.

Thus, it becomes interesting to characterize the free profinite object over DA. In this paper, we present three representations of the free pro-DA object on a finite alphabet extending techniques developed for the pseudovariety R: the first one by means of finite-height labeled trees, using the ideas of Almeida and Weil [7], the second one by means of *quasi-ternary* labeled trees, based on the work of Almeida and Zeitoun [8] and the last one by means of labeled linear orderings, extending the description done by Almeida and Weil [7]. In connection with the second representation, we also exhibit a representation by *wrapped* automata, which turns out to be useful for recent work of the author.

The paper is organized as follows. In Section 2, we recall the basics of the theory of pseudovarieties of monoids, pro-V monoids and some notions on automata and linear orderings. In Section 3, we use the central basic factorization of an implicit operation on DA and we present two forms of iteration of that factorization. We prove the convergence of the infinite product resulting from the iterated factorizations and we end it with a characterization of the idempotents in terms of the type of iterated central basic factorization. In Section 4, we present the representations of the implicit operations over this pseudovariety that we had announced above. We end the section by relating the various representations.

2. Preliminaries

We briefly recall some basics of the theory of pseudovarieties of monoids, profinite monoids, automata and linear orderings and we introduce some related notation. We recommend [2, 4] for a better understanding of these concepts and [12] as a reference on linear orderings.

In this paper, A is a finite set called *alphabet* and its elements are called *letters*. We denote by A^* (respectively by A^+) the free monoid (respectively the free semigroup) generated by A, whose elements are called *words*. The *empty word* is denoted by 1. The *length* of a word u is denoted by $|u|$ and the *cardinality* of A is denoted by $|A|$. The *content* of a word u is the smallest subset B of A such that $u \in B^*$. In particular, the content of the empty word is $\emptyset$. Finally, a word $u = a_1 \cdots a_n$, with $a_i \in A$, for all i, is a *subword* of v if there exist words $v_0,\ v_1, \ldots, v_n \in A^*$ such that $v = v_0 a_1 v_1 \cdots a_n v_n$.

Given a semigroup S, we denote by S^1 the monoid defined as follows: if S is itself a monoid, then $S^1 = S$; otherwise, $S^1 = S \cup \{1\}$, where 1 is an element that does not belong to S and the multiplication in S^1 is the (unique) extension to the multiplication in S in which 1 acts as a neutral element. For $s \in S$, we denote by s^ω the unique idempotent in the subsemigroup generated by s and we set $s^{\omega+1} = s^\omega s$.

A class of finite monoids that is closed under taking submonoids, homomorphic images and finite direct products is called a *pseudovariety* and generally denoted by V. For example, M is the pseudovariety of all finite monoids, R is the pseudovariety of all $\mathcal{R}$-trivial monoids, where a monoid S is $\mathcal{R}$-trivial if, for all $s, t \in S$, $s \mathcal{R} t$ implies $s = t$. In this paper, we are interested in DA, the pseudovariety of monoids whose regular $\mathcal{D}$-classes are aperiodic semigroups. Note that a semigroup S is *aperiodic* if $s^\omega = s^{\omega+1}$, for all $s \in S$, and a monoid S is in DA if and only if, for all $s,\ t \in S$, we have $(st)^\omega (ts)^\omega (st)^\omega = (st)^\omega$ and $s^\omega = s^{\omega+1}$.

A *topological monoid* is a monoid equipped with a topology for which the multiplication in the monoid is a continuous function. We view a finite monoid as a topological monoid with respect to the discrete topology. A topological monoid S is a *profinite monoid* (respectively a *pro-*V *monoid*) if it is a compact monoid which is *residually finite* (respectively *residually in* V), which means that, whenever $s, t \in S$ and $s \neq t$, there exists a continuous homomorphism $\varphi : S \to F$ into a finite monoid (respectively into a member of V) such that $\varphi(s) \neq \varphi(t)$. It is well known that profinite monoids are 0*-dimensional*, which means that the topology has an open basis consisting of clopen sets (which is equivalent to being a *totally disconnected* monoid).

Given an alphabet A and a pseudovariety V, the *free pro-*V *monoid on* A, denoted by $\overline{\Omega}_A\mathsf{V}$, is the unique (up to isomorphism of topological monoids) pro-V monoid such that, for every mapping $\mu : A \to T$ into a pro-V monoid T, there is a unique continuous homomorphism $\hat{\mu} : \overline{\Omega}_A\mathsf{V} \to T$ such that $\hat{\mu} \circ \iota = \mu$, where $\iota : A \to \overline{\Omega}_A\mathsf{V}$ is the natural generating function. The elements of $\overline{\Omega}_A\mathsf{V}$ are called *implicit operations on* V or *pseudowords*. For a pseudovariety V containing Sl, the *content function* is the unique continuous homomorphism $c : \overline{\Omega}_A\mathsf{V} \to \mathcal{P}(A)$ such that $c\iota(a) = \{a\}$, for all $a \in A$.

A *pseudoidentity* is an equality of the form $u = v$, with $u, v \in \overline{\Omega}_A\mathsf{M}$, and $|A|$ is called the *arity* of the pseudoidentity. We say that a pseudoidentity is *valid* in a profinite monoid T, and we write $T \models u = v$, if $\varphi(u) = \varphi(v)$ for every continuous homomorphism $\varphi : \overline{\Omega}_A\mathsf{M} \to T$. Reiterman's Theorem [11] says that every pseudovariety is defined by some set of finitary (A is finite) pseudoidentities. That the class of all finite monoids which verify all the elements of a set of pseudoidentities is a pseudovariety follows immediately from the fact that the validity of a pseudoidentity in a finite monoid is preserved under taking homomorphic images, submonoids and finite direct products. For example, the pseudovariety DA is defined by the set of pseudoidentities $\{(xy)^\omega(yx)^\omega(xy)^\omega = (yx)^\omega,\ x^\omega = x^{\omega+1}\}$.

A *deterministic automaton* over an alphabet A is a tuple $\mathcal{A} = (V, \to, q, F)$, where V is the set of states, $q \in V$ is the initial state, $F \subseteq V$ is the set of final states and $\to : V \times A \to V$ is its transition function. We denote by $v.a$ the state reached from v by reading the letter a, if this state exists, and we denote by $v.L$ the set of states reached from v by reading some word of L.

Finally, we suppose that the reader is acquainted with the basic notions of linear orderings. In this paper, we use two different linear orderings of the set of natural numbers: the usual ordering, $R_{\mathbb{N}}$, and the backwards ordering, $R^*_{\mathbb{N}}$. We also use suborderings of these orderings and operations on linear orderings. We denote by ω, ω^* and $\mathbf{n}$ the order type of $\langle \mathbb{N}, R_{\mathbb{N}}\rangle$, $\langle \mathbb{N}, R^*_{\mathbb{N}}\rangle$ and $\langle P, R\rangle$, which is a subordering of $\langle \mathbb{N}, R_{\mathbb{N}}\rangle$ with $|P| = n$, respectively.

3. Factorization of implicit operations and convergence of infinite products in pro-DA monoids

Let $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$. We define the *central basic factorization of* w (see Almeida [3] or Trotter and Weil [15]) as a factorization of one of the following forms:

(i) **standard form:** $w = \alpha a \gamma b \beta$ with $a, b \in A$, $\alpha, \beta, \gamma \in \overline{\Omega}_A\mathsf{DA}$, $a \notin c(\alpha)$, $b \notin c(\beta)$ and $c(\alpha a) = c(b\beta) = c(w)$;

(ii) **overlapped form:** $w = \alpha b \gamma a \beta$ with $a, b \in A$, $\alpha, \beta, \gamma \in \overline{\Omega}_A\mathsf{DA}$, $a \notin c(\alpha b \gamma)$, $b \notin c(\gamma a \beta)$ and $c(\alpha b \gamma a) = c(b \gamma a \beta) = c(w)$;

(iii) **degenerate form:** $w = \alpha a \beta$ with $a \in A$, $\alpha, \beta \in \overline{\Omega}_A\mathsf{DA}$, $a \notin c(\alpha)$, $a \notin c(\beta)$ and $c(\alpha a) = c(a\beta) = c(w)$.

By the following theorem, the central basic factorization of $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$ exists and is unique:

Proposition 3.1 (Almeida [3]). *Let $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$. Then w has a unique central basic factorization. In other words, if w, $w' \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$, the equality $w = w'$ is valid in DA, and Φ and Ψ are central basic factorizations of w and w', respectively, then the two factorizations are both of the same type ((i),(ii) or (iii)) and the equalities of factors in corresponding positions are valid in DA.*

We denote the central basic factorization of w by the tuple $\mathsf{CBF}(w) = (\alpha, a, \gamma, b, \beta)$ or by the triple $\mathsf{CBF}(w) = (\alpha, a, \beta)$, as it is of the standard or overlapped form, or of the degenerate form, respectively.

For what follows, we define two different types of iteration of this factorization: in the first one we iterate the factorization while the content of the central factor (if it exists) does not decrease and in the second one we iterate until the central factor (if it exists) becomes 1. We proceed to explain this in detail.

Let $\gamma_0 = w$. If $c(\gamma_k) = c(w)$, we consider the central basic factorization of γ_k which, in the case of being of the standard form, is $\gamma_k = \alpha_{k+1}a_{k+1}\gamma_{k+1}b_{k+1}\beta_{k+1}$. The $(k+1)$-iteration of the central basic factorization of w is $w = \alpha_1 a_1 \cdots \alpha_{k+1} \cdot a_{k+1}\gamma_{k+1}b_{k+1}\beta_{k+1} \cdots b_1\beta_1$ and γ_{k+1} is called the *remainder of order* $k+1$. We iterate this process while γ_k exists and $c(\gamma_k) = c(w)$. If, for any n, $c(\gamma_n) \neq c(w)$ and γ_{n-1} admits a central basic factorization of the standard form, then $w = \alpha_1 a_1 \cdots \alpha_n a_n \gamma_n b_n \beta_n \cdots b_1 \beta_1$ is called the *iterated central basic factorization of type* 1 of w and is called *standard* and of length n. If $c(\gamma_{n-1}) = c(w)$ and γ_{n-1} has an overlapped central basic factorization, then $w = \alpha_1 a_1 \cdots \alpha_n b_n \gamma_n a_n \beta_n \cdots b_1 \beta_1$ is the *iterated central basic factorization of type* 1 of w and is called *overlapped* and of length n. If $c(\gamma_{n-1}) = c(w)$ and γ_{n-1} has a degenerate central basic factorization, $\gamma_{n-1} = \alpha_n a_n \beta_n$, then $w = \alpha_1 a_1 \cdots \alpha_n a_n \beta_n \cdots b_1 \beta_1$ is the *iterated central basic factorization of type* 1 of w and is called *degenerate* and of length n. We say that, in the first two cases, γ_n is the *remainder* of the central basic factorization of w, while in the degenerate case there is no remainder. Finally, if $c(\gamma_n) = c(w)$, for all n, we say that w admits an *infinite iterated central basic factorization of type* 1 and we write $w = \alpha_1 a_1 \alpha_2 a_2 \cdots \;\cdots b_2 \beta_2 b_1 \beta_1$. We denote the iterated central basic factorization of type 1 by $\mathsf{I_1CBF}(w)$. Note that all the factors involved in this factorization have content strictly contained in $c(w)$.

Now, let $w = \alpha_1 a_1 \gamma_1 b_1 \beta_1$ be the central basic factorization of w. While $\gamma_k \neq 1$ or γ_{k-1} does not admit a degenerate central basic factorization, we consider the central basic factorization of γ_k. The *iterated central basic factorization of type* 2 of w, $\mathsf{I_2CBF}(w)$, is defined by one of the following forms: $w = \alpha_1 a_1 \alpha_2 a_2 \cdots \alpha_n a_n \beta_n \cdots b_2 \cdot \beta_2 b_1 \beta_1$, in case $\mathsf{CBF}(\gamma_{n-1})$ is degenerate, $w = \alpha_1 a_1 \alpha_2 a_2 \cdots \alpha_n a_n b_n \beta_n \cdots b_2 \beta_2 b_1 \beta_1$, in case $\gamma_n = 1$, or $w = \alpha_1 a_1 \alpha_2 a_2 \cdots \;\cdots b_2 \beta_2 \cdot b_1 \beta_1$, if the iteration is infinite. Note that, also in this iterated basic factorization, all the factors involved have content strictly contained in $c(w)$.

Results from Almeida [3, 2] and Almeida and Weil [7] allow us to conclude that we can iterate the central basic factorization of any of these two types and that the infinite product, in fact, converges.

Lemma 3.2 (Almeida and Weil [7]). *Let S be a compact monoid. Then any two accumulation points of every right infinite product in S are $\mathcal{R}$-equivalent.*

Lemma 3.3 (dual of the previous lemma). *Let S be a compact monoid. Then any two accumulation points of every left infinite product in S are $\mathcal{L}$-equivalent.*

Corollary 3.4. *Let S be a pro-DA monoid. Given a right infinite product in S and a left infinite product in S such that their accumulation points are in the same regular $\mathcal{J}$-class, then the product of any right accumulation point by any left accumulation point is independent of the choice of these points.*

Proof. It is enough to observe that the product of any two accumulation points of each of the two infinite products is in the regular $\mathcal{H}$-class $R \cap L$ which is trivial, where R is the regular $\mathcal{R}$-class that contains all the accumulation points of the right infinite product and L is the regular $\mathcal{L}$-class that contains all the accumulation points of the left infinite product. □

We denote by $\overrightarrow{\prod}_{k=1}^{n} u_k$ the product $u_1 u_2 \cdots u_n$ and by $\overleftarrow{\prod}_{k=1}^{n} v_k$ the product $v_n \cdots v_2 v_1$. Given a pro-DA monoid S and sequences $(u_k)_{k\geq 1}$, $(v_k)_{k\geq 1} \in S^{\mathbb{N}}$ in the conditions of the previous corollary, we denote by $\overrightarrow{\prod}_{k\geq 1} u_k \cdot \overleftarrow{\prod}_{k\geq 1} v_k$ the product of an accumulation point of the sequence $(\overrightarrow{\prod}_{k=1}^{n} u_k)_n$ by an accumulation point of the sequence $(\overleftarrow{\prod}_{k=1}^{n} v_k)_n$, when n goes to infinity.

Therefore, given $w \in \overline{\Omega}_A \mathsf{DA} \backslash \{1\}$, the iterated central basic factorization of w is of one of the following forms:

$$w = \overrightarrow{\prod}_{k=1}^{n} (\alpha_k a_k) \cdot \gamma_n \cdot \overleftarrow{\prod}_{k=1}^{n} (b_k \beta_k)$$

if it is of type 1 and it is finite and of the standard form or of the overlapped form, or

$$w = \overrightarrow{\prod}_{k=1}^{n} (\alpha_k a_k) \cdot \overleftarrow{\prod}_{k=1}^{n} (b_k \beta_k)$$

if is of type 2 and it is finite and $\mathsf{CBF}(\gamma_{n-1})$ is of the standard form or of the overlapped form, or

$$w = \overrightarrow{\prod}_{k=1}^{n} (\alpha_k a_k) \cdot \beta_n \cdot \overleftarrow{\prod}_{k=1}^{n-1} (b_k \beta_k)$$

if is of type 1 or 2 and it is finite and degenerate, or

$$w = \overrightarrow{\prod}_{k\geq 1} (\alpha_k a_k) \cdot \overleftarrow{\prod}_{k\geq 1} (b_k \beta_k)$$

if the iteration is infinite (and of any type). In fact, the last equality is valid as we see from Lemma 3.9 or from Lemma 3.11, depending on the type of the iterated factorization. We recall some results that we use to prove these lemmas.

Lemma 3.5 (cf. [2, Lemma 8.1.4]). *Let $S \in \mathsf{DS}$ and let $e \in E(S)$ and $u \in S$ such that $u \geq_{\mathcal{J}} e$. Then $eu\mathcal{R}e\mathcal{L}ue$.*

Corollary 3.6. *Let $S \in \mathsf{DA}$ and let $e, f \in E(S)$ and $u \in S$ be such that $u \geq_{\mathcal{J}} e\mathcal{J}f$. Then $euf = ef$.*

Corollary 3.7 (cf. [2, Theorem 8.1.7]). *Let S be a pro-DA semigroup and let $r, s, t \in S$ be such that $c(s) \subseteq c(r) = c(t)$. Then $r^\omega s t^\omega = r^\omega t^\omega$.*

Proposition 3.8 (cf. [2, Theorem 8.1.10]). *Given $w \in \overline{\Omega}_A\mathsf{DA}$, then w is idempotent if and only if $\left[{w \atop u}\right] \in \{0, \infty\}$, for any $u \in A^+$, where $\left[{w \atop u}\right]$ is the supremum of the integers r such that u^r is a subword of w.*

We are now ready to prove the convergence of the infinite product in the iterated central basic factorization of type 1 of w, as stated in the following lemma:

Lemma 3.9. *Given $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$, if w has an infinite iterated central basic factorization of type 1, then $w = \overrightarrow{\prod}_{k\geq 1}(\alpha_k a_k) \cdot \overleftarrow{\prod}_{k\geq 1}(b_k\beta_k)$.*

Proof. The successive iterations of the central basic factorization of type 1 of w are $w = \overrightarrow{\prod}_{k=1}^{n}(\alpha_k a_k) \cdot \gamma_n \cdot \overleftarrow{\prod}_{k=1}^{n}(b_k\beta_k)$, for all n. By compactness, there exists a subsequence $(\overrightarrow{\prod}_{k=1}^{m}(\alpha_k a_k)\, ,\, \gamma_m\, ,\, \overleftarrow{\prod}_{k=1}^{m}(b_k\beta_k))_{m\in M}$ that converges to some $(\alpha,\, \gamma,\, \beta)$. By Proposition 3.8, α and β are idempotents, since $c(\alpha_k a_k) = c(\alpha)$ and $c(b_k\beta_k) = c(\beta)$, for all k, and also $\alpha \mathcal{J} \beta$, by [2, Theorem 8.1.7], since they have the same content. It follows, by Corollaries 3.6 and 3.4, that $w = \alpha\gamma\beta = \alpha\beta = \overrightarrow{\prod}_{k\geq 1}(\alpha_k a_k)\cdot\overleftarrow{\prod}_{k\geq 1}(b_k\beta_k)$, since $c(\gamma) \subseteq c(\alpha) = c(\beta)$. □

To show that the infinite product in the iterated central basic factorization of type 2 of w converges we need to beware of the fact that the content of the factors $\alpha_k a_k$ and $b_k\beta_k$ could decrease. In fact, the following lemma shows that the sequence of these contents stabilizes.

Lemma 3.10. *Let $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$ be such that the iterated central basic factorization of type 2 of w is infinite, $\mathsf{I}_2\mathsf{CBF}(w) = \alpha_1 a_1 \alpha_2 a_2 \cdots\ \cdots b_2\beta_2 b_1\beta_1$. Then there exists $N \in \mathbb{N}$ such that, if $i \geq N$, $c(\alpha_i a_i)$ and $c(b_i\beta_i)$ are constant and equal, for all i. Moreover, the pseudowords $w_i = \alpha_i a_i \alpha_{i+1} a_{i+1} \cdots\ \cdots b_{i+1}\beta_{i+1} b_i\beta_i$, with $i \geq N$, have a standard central basic factorization.*

Proof. We consider the sequence $(w_n = \gamma_{n-1} = \alpha_n a_n \cdots\ \cdots b_n\beta_n)_{n\geq 1}$ of elements of $\overline{\Omega}_A\mathsf{DA}\backslash\{1\}$. We have $c(w_1) = c(w)$ and $c(w_j) \subseteq c(w_i)$, if $i < j$. Since $\mathsf{I}_2\mathsf{CBF}(w)$ is infinite and A is a finite alphabet, it follows that, from a certain point on, the contents $c(w_i)$ must stabilize. Let N be a integer such that, if $i, j \geq N$, then $c(w_i) = c(w_j)$. We recall that, if the central basic factorization of w_i, $\mathsf{CBF}(w_i) = \alpha_i a_i \gamma_i b_i \beta_i$, is of the overlapped form, then $a_i \notin c(\gamma_i b_i\beta_i)$ and $b_i \notin c(\alpha_i a_i \gamma_i)$ and, therefore, $c(\gamma_i = w_{i+1}) \subsetneq c(w_i)$. On the other hand, if the central basic factorization of w_i is degenerate, then $\mathsf{I}_2\mathsf{CBF}(w_i)$ is finite and, therefore, $\mathsf{I}_2\mathsf{CBF}(w)$ is also finite, which contradicts the hypothesis. So, if $i \geq N$, then $\mathsf{CBF}(w_i)$ is standard. By definition of standard central basic factorization and by the above, it follows that, if $i \geq N$, then $c(\alpha_i a_i) = c(b_i\beta_i) = c(w_N)$.

Finally, we note that, if $i < N$, then the central basic factorization of w_i could be of the standard or of the overlapped form. □

Lemma 3.11. *Given $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$, if w has an infinite iterated central basic factorization of type 2, then $w = \overrightarrow{\prod}_{k\geq 1}(\alpha_k a_k) \cdot \overleftarrow{\prod}_{k\geq 1}(b_k\beta_k)$.*

Proof. Let $\mathsf{I}_2\mathsf{CBF}(w) = \alpha_1 a_1 \alpha_2 a_2 \cdots\ \cdots b_2\beta_2 b_1\beta_1$, let N be an integer satisfying the condition of Lemma 3.10 and let $w_N = \alpha_N a_N \alpha_{N+1} a_{N+1} \cdots\ \cdots b_{N+1}\beta_{N+1} b_N\beta_N$. Note that, by Lemma 3.10, $c(\alpha_k a_k) = c(w_N) = c(b_k\beta_k)$, for all $k \geq N$. So, the iterated central basic factorization of type 2 of w_N coincides with the iterated central basic factorization of type 1 of w_N. Applying Lemma 3.9 to this iterated factorization, it follows that $w_N = \overrightarrow{\prod}_{k\geq N}(\alpha_k a_k) \cdot \overleftarrow{\prod}_{k\geq N}(b_k\beta_k)$ and, therefore, $w = \overrightarrow{\prod}_{k\geq 1}(\alpha_k a_k) \cdot \overleftarrow{\prod}_{k\geq 1}(b_k\beta_k)$. $\square$

Example 3.12. The iterated central basic factorizations of $w = a^\omega b^\omega c^2(ab)^\omega cabc^\omega$ of type 1 and of type 2 are, respectively:

$$\mathsf{I_1CBF}(w) = a^\omega b^\omega \cdot \underline{c} \cdot ca \cdot \underline{b} \cdot (ab)^\omega \cdot \underline{a} \cdot bc \cdot \underline{a} \cdot bc^\omega$$

$$\text{and } \mathsf{I_2CBF}(w) = a^\omega b^\omega \cdot \underline{c} \cdot ca \cdot \underline{b} \cdot a \cdot \underline{b} \cdot a \cdot \underline{b} \cdots \;\; \cdots \underline{a} \cdot b \cdot \underline{a} \cdot b \cdot \underline{a} \cdot bc \cdot \underline{a} \cdot bc^\omega.$$

Note that $\mathsf{I_1CBF}(w)$ is finite, standard and of length 2 while $\mathsf{I_2CBF}(w)$ is infinite.

The next step in our factorizations consists in applying successively the corresponding factorization to each factor α_i, β_i and γ_n, the latter only in the case of a finite iterated central basic factorization of type 1. We observe, by Proposition 3.13, that this process is finite. For that purpose, we define two operations in $\overline{\Omega}_A\mathsf{DA}$, op^1 and op^2, as follows. We start with op^1:

(i) Let $u = a_1 \cdots a_n b_n \cdots b_1 \in A^+$ and $\alpha_1, \ldots, \alpha_n, \gamma_n, \beta_n, \ldots, \beta_1 \in \overline{\Omega}_A\mathsf{DA}$ be such that, for each $i < n$, $a_i \notin c(\alpha_i)$, $b_i \notin c(\beta_i)$, $\alpha_i a_i$ and $b_i\beta_i$ have the same content as the product $\alpha_1 a_1 \cdots \alpha_n a_n \gamma_n b_n \beta_n \cdots b_1\beta_1$ and, for $i = n$, either a_n, b_n, $\alpha_n a_n$, $b_n\beta_n$ satisfy the same conditions, or $a_n \notin c(\gamma_n) \cup \{b_n\} \cup c(\beta_n)$, $b_n \notin c(\alpha_n) \cup \{a_n\} \cup c(\gamma_n)$, and $\alpha_n a_n \gamma_n b_n$ and $a_n\gamma_n b_n\beta_n$ have the same content as $\alpha_1 a_1 \cdots \alpha_n a_n \gamma_n b_n \beta_n \cdots b_1\beta_1$. We define $op^1_u(\alpha_1, \ldots, \alpha_n, \gamma_n, \beta_n, \ldots, \beta_1) = \alpha_1 a_1 \cdots \alpha_n a_n \gamma_n b_n \beta_n \cdots b_1 \beta_1$.

(ii) Let $u = a_1 \cdots a_n b_{n-1} \cdots b_1 \in A^+$ and $\alpha_1, \ldots, \alpha_n, \beta_n, \ldots, \beta_1 \in \overline{\Omega}_A\mathsf{DA}$ be such that, for all i, $a_i \notin c(\alpha_i)$, $b_i \notin c(\beta_i)$, $a_n \notin c(\beta_n)$ and $\alpha_i a_i$, $b_i\beta_i$ and $a_n\beta_n$ have the same content as the product $\alpha_1 a_1 \cdots \alpha_n a_n \beta_n \cdots b_1\beta_1$. We define $op^1_u(\alpha_1, \ldots, \alpha_n, \beta_n, \ldots, \beta_1) = \alpha_1 a_1 \cdots \alpha_n a_n \beta_n \cdots b_1\beta_1$.

(iii) Let $u = \overrightarrow{\prod}_{i\geq 1} a_i \cdot \overleftarrow{\prod}_{i\geq 1} b_i \in A^{\omega+\omega^*}$ and $\alpha_1, \alpha_2, \ldots, \beta_1, \beta_2, \ldots \in \overline{\Omega}_A\mathsf{DA}$ be such that, for all i, $a_i \notin c(\alpha_i)$, $b_i \notin c(\beta_i)$ and $\alpha_i a_i$ and $b_i\beta_i$ have the same content as the product $\overrightarrow{\prod}_{k\geq 1}(\alpha_k a_k) \cdot \overleftarrow{\prod}_{k\geq 1}(b_k\beta_k)$. We define $op^1_u(\alpha_1, \alpha_2, \ldots, \ldots, \beta_2, \beta_1) = \overrightarrow{\prod}_{i\geq 1}(\alpha_i a_i) \cdot \overleftarrow{\prod}_{i\geq 1}(b_i\beta_i)$.

Similarly, we define op^2, but now assuming that the contents could decrease:

(i) Let $u = a_1 \cdots a_n b_n \cdots b_1 \in A^+$ and $\alpha_1, \ldots, \alpha_n, \beta_n, \ldots, \beta_1 \in \overline{\Omega}_A\mathsf{DA}$ be such that, for each $i < n$, $a_i \notin c(\alpha_i)$, $b_i \notin c(\beta_i)$, $\alpha_i a_i$ and $b_i\beta_i$ have the same content as the product $\alpha_i a_i \cdots \alpha_n a_n b_n \beta_n \cdots b_i \beta_i$ and, for $i = n$, either $a_n \notin c(\alpha_n)$, $b_n \notin c(\beta_n)$, and $\alpha_n a_n$ and $b_n\beta_n$ have the same content as $\alpha_n a_n b_n \beta_n$, or $a_n \notin \{b_n\} \cup c(\beta_n)$, $b_n \notin c(\alpha_n) \cup \{a_n\}$ and $\alpha_n a_n b_n$ and $a_n b_n \beta_n$ have the same content as $\alpha_n a_n b_n \beta_n$. We define $op^2_u(\alpha_1, \ldots, \alpha_n, \beta_n, \ldots, \beta_1) = \alpha_1 a_1 \cdots \alpha_n a_n b_n \beta_n \cdots b_1 \beta_1$.

(ii) Let $u = a_1 \cdots a_n b_{n-1} \cdots b_1 \in A^+$ and $\alpha_1, \ldots, \alpha_n, \beta_n, \ldots, \beta_1 \in \overline{\Omega}_A\mathsf{DA}$ be such that, for all i, $a_i \notin c(\alpha_i)$, $b_i \notin c(\beta_i)$, $a_n \notin c(\beta_n)$, and $\alpha_i a_i$ and $b_i\beta_i$ have the same content as the product $\alpha_i a_i \cdots \alpha_n a_n \beta_n \cdots b_i \beta_i$. We define $op^2_u(\alpha_1, \ldots, \alpha_n, \beta_n, \ldots, \beta_1) = \alpha_1 a_1 \cdots \alpha_n a_n \beta_n \cdots b_1 \beta_1$.

(iii) Let $u = \overrightarrow{\prod}_{i\geq 1} a_i \cdot \overleftarrow{\prod}_{i\geq 1} b_i \in A^{\omega+\omega^*}$ and $\alpha_1, \alpha_2, \ldots, \beta_1, \beta_2, \ldots \in \overline{\Omega}_A\mathsf{DA}$ be such that, for all i, $a_i \notin c(\alpha_i)$, $b_i \notin c(\beta_i)$ and $\alpha_i a_i$ and $b_i\beta_i$ have the same content as the product $\overrightarrow{\prod}_{k\geq i}(\alpha_k a_k) \cdot \overleftarrow{\prod}_{k\geq i}(b_k\beta_k)$. We define $op^2_u(\alpha_1, \alpha_2, \ldots, \ldots, \beta_2, \beta_1) = \overrightarrow{\prod}_{i\geq 1}(\alpha_i a_i) \cdot \overleftarrow{\prod}_{i\geq 1}(b_i\beta_i)$.

Proposition 3.13. *Each element of $\overline{\Omega}_A\mathsf{DA}$ can be obtained from 1 by applying successively one of the described operations a number of times that does not exceed $|A|$.*

Proof. Each $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$ has an iterated central basic factorization that involves a finite (or infinite) product of letters a_i and b_i, which depends on the type of the iterated factorization. As the content of each factor is strictly contained in $c(w)$, the result follows by induction on $|c(w)|$. □

We end this section by presenting some characterizations of the implicit operations on DA connected with its iterated central basic factorizations.

Let $\rceil\!\!\lceil w \rceil\!\!\rceil_i$ be the number of iterations until we obtain the iterated central basic factorization of type i of w, with $i \in \{1,2\}$. Note that $\rceil\!\!\lceil w \rceil\!\!\rceil_i \in \mathbb{N}\cup\{\infty\}$. We denote by $\|w\|_i$ the greatest integer n such that $c(\alpha_n a_n) = c(b_n\beta_n) = c(w)$ and $\alpha_n a_n$ and $b_n\beta_n$ are disjoint, in the iterated central basic factorization of type i of w. If this maximum does not exist, we set $\|w\|_i = \infty$. If this condition does not occur for any integer n, we set $\|w\|_i = 0$. Note that $\rceil\!\!\lceil w \rceil\!\!\rceil_1 = \|w\|_1$ for the standard and the overlapped case, and $\rceil\!\!\lceil w \rceil\!\!\rceil_1 - 1 = \|w\|_1$ in the degenerate case. Moreover, we have $\|w\|_1 = \|w\|_2$. The following statements are formulated for the iterated central basic factorization of type 2, although we have similar results for the iterated central basic factorization of type 1. We leave the details to the reader. From hereon, we use the notation $\rceil\!\!\lceil w \rceil\!\!\rceil$ and $\|w\|$ instead of $\rceil\!\!\lceil w \rceil\!\!\rceil_2$ and $\|w\|_2$, respectively.

Lemma 3.14. *Let $u, v \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$. If $c(u) = c(v)$, then $\|uv\| \geq \|u\| + \|v\|$.*

Proof. Let $u = \alpha_{u_1}a_{u_1}\cdots\alpha_{u_m}a_{u_m}\gamma_{u_m}b_{u_m}\beta_{u_m}\cdots b_{u_1}\beta_{u_1}$ and $v = \alpha_{v_1}a_{v_1}\cdots\alpha_{v_n}\cdot a_{v_n}\gamma_{v_n}b_{v_n}\beta_{v_n}\cdots b_{v_1}\beta_{v_1}$ be, respectively, the m-iteration and the n-iteration of the central basic factorization of u and v, with $\|u\| = m$ and $\|v\| = n$, for $m, n \in \mathbb{N}$. Note that $c(\alpha_{u_i}a_{u_i}) = c(b_{u_i}\beta_{u_i}) = c(u) = c(v) = c(\alpha_{v_j}a_{v_j}) = c(b_{v_j}\beta_{v_j})$ for all $i \leq m$ and $j \leq n$. Since $uv = \alpha_{u_1}a_{u_1}\cdots a_{u_m}\gamma_{u_m}b_{u_m}\cdots b_{u_1}\beta_{u_1}\cdot\alpha_{v_1}a_{v_1}\cdots a_{v_n}\gamma_{v_n}b_{v_n}\cdots b_{v_1}\beta_{v_1}$, it is easy to see that $\|uv\| \geq m+n = \|u\| + \|v\|$. The case where $\|u\| = \infty$ or $\|v\| = \infty$ is similar. □

We define the *cumulative content* of $w \in \overline{\Omega}_A\mathsf{DA}$, and we denote it by $\vec{c}(w)$, to be the empty set, if $\mathsf{I_2CBF}(w)$ is finite, or the set of all letters $a \in A$ such that there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, $a \in c(\alpha_n a_n) = c(b_n\beta_n)$, in the case where $\mathsf{I_2CBF}(w)$ is infinite.

Proposition 3.15. *Let $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$. The following conditions are equivalent:*

(i) $\mathsf{DA} \models w^2 = w$;
(ii) $\|w\| = \infty$;
(iii) $c(w) = \vec{c}(w)$.

Proof. $(i) \Rightarrow (ii)$: Suppose that w is idempotent. By Lemma 3.14, we have $\|w\| = \|w^2\| \geq 2\|w\|$ which implies $\|w\| = \infty$ or $\|w\| = 0$. Since $\|w^2\| > 0$, for all $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$, then we have $\|w\| = \infty$.

$(ii) \Rightarrow (i)$: Suppose that $\|w\| = \infty$, i.e., $w = \overrightarrow{\prod}_{k\geq 1}(\alpha_k a_k)\cdot\overleftarrow{\prod}_{k\geq 1}(b_k\beta_k)$ with $c(\alpha_k a_k) = c(b_k\beta_k) = c(w)$, for all k. By Proposition 3.8, it follows that w is an idempotent of $\overline{\Omega}_A\mathsf{DA}$. Let $\varphi : \overline{\Omega}_A\mathsf{DA} \to S$ be a continuous homomorphism in a finite monoid $S \in \mathsf{DA}$. Then $\varphi(w)$ is an idempotent. Thus $\mathsf{DA} \models w^2 = w$.

$(ii) \Leftrightarrow (iii)$: It follows immediately from the definition of cumulative content of w. □

Lemma 3.16. *Let $w = xyz \in \overline{\Omega}_A\mathsf{DA}$ with $c(x)$, $c(z) \subsetneq c(w)$. Then $\|w\| \leq \|y\| + 1$.*

Proof. If $\|y\| = \infty$ or x and z are the empty word, the result is trivial. Otherwise, we proceed by induction on $(|c(y)|, \|y\|)$ where the pairs are in the lexicographic order. Suppose that $|c(y)| = 1$. Since $c(x)$, $c(z) \subsetneq c(w)$, there exists a letter in $c(y) \cup c(z)$ that does not belongs to $c(x)$ and there exists a letter in $c(x) \cup c(y)$ that does not belong to $c(z)$. By definition of central basic factorization, the following cases can occur:

(i) $\mathsf{CBF}(w) = (\alpha, a, \gamma, b, \beta)$ of the standard form: In this case, the letters a and b belong to the content of y (otherwise, $c(x) = c(w)$ or $c(z) = c(w)$). Since $|c(y)| = 1$, we have $a = b$. Moreover, $a \notin c(x) = c(z)$ and $c(\gamma) = \{a\} \subsetneq c(w)$, because γ is a factor of y and $c(y) = \{a\}$. It follows that $\|w\| = 1$.
(ii) $\mathsf{CBF}(w) = (\alpha, a, \gamma, b, \beta)$ of the overlapped form: It follows, immediately, that $\|w\| = 0$.
(iii) $\mathsf{CBF}(w) = (\alpha, a, \beta)$ of the degenerate form: It follows, immediately, that $\|w\| = 0$.

Now, suppose that $|c(y)| > 1$. Let y_1 be a prefix of yz and let y_2 be a suffix of xy such that y_1 is minimal for $c(xy_1) = c(w)$ and y_2 is minimal for $c(y_2z) = c(w)$. Note that xy_1 is one of the products αa or $\alpha a\gamma b$, according to the $\mathsf{CBF}(w)$ is of the standard or of the degenerate form, in the first case, or of the overlapped form, in the second case. We have the dual result for the factor y_2z and, therefore, the existence of these factors is justified. We also note that y_1 and y_2 are non-empty words, by the hypothesis $c(x), c(z) \subsetneq c(w)$. If y_1 is not a prefix of y or y_2 is not a suffix of y, then the central basic factorization of w is overlapped and, therefore, $\|w\| = 0$. Suppose that y_1 is a prefix of y and y_2 is a suffix of y. Two cases can occur: $y = y_1y'y_2$, with y' possibly empty, or y_1 and y_2 are not disjoint factors of y. In the latter case, the central basic factorization of w is of the overlapped form or of the degenerate form and, therefore, $\|w\| = 0$. Suppose that $y = y_1y'y_2$, with y' possibly empty. If $c(y') \subsetneq c(w)$, then $\|w\| = 1$ and the result follows. Suppose that $c(y') = c(w)$ and, therefore, $\|w\| = \|y'\| + 1$. We consider the central basic factorization of y, $\mathsf{CBF}(y)$. The cases where $\mathsf{CBF}(y)$ are of the degenerate or of the overlapped form are trivial, because $\|y'\| = 0$ (note that $c(y') = c(y)$). Suppose that $\mathsf{CBF}(y) = (\alpha, a, \gamma, b, \beta)$ is of the standard form. Then there exist y'_1 and y'_2 such that $\alpha a = y_1y'_1$, $b\beta = y'_2y_2$ and $y'_1\gamma y'_2 = y'$. We have $c(\gamma) \subsetneq c(y)$ or $\|\gamma\| = \|y\| - 1$. If $c(\gamma) \subsetneq c(y)$, then $\|y\| = 1$, $\|y'\| \leq 1$ and $\|w\| \leq 2$. Otherwise, it follows that $c(\gamma) = c(y) = c(y') = c(w)$. If $c(y'_1)$, $c(y'_2) \subsetneq c(\gamma)$, by induction hypothesis, it follows that $\|y'\| \leq \|\gamma\| + 1 = \|y\|$. Thus $\|w\| \leq \|y\| + 1$. Otherwise, $c(y'_1) = c(\gamma)$, $c(y'_2) = c(\gamma)$, or both occur simultaneously. Suppose, by symmetry, that $c(y'_1) = c(\gamma) = c(w)$. Since $\alpha a = y_1y'_1$ is the minimum prefix of y such that $c(\alpha a) = c(w)$, it follows that y'_1 is the minimum prefix of y' such that $c(y'_1) = c(w)$. It follows that $\|w\| = \|y'\| + 1 \leq \|\gamma\| + 2 = \|y\| + 1$. □

Corollary 3.17. *Let $w = x_1 \cdots x_r \in \overline{\Omega}_A\mathsf{DA}$ with $c(x_i) \subsetneq c(w)$, for all i. Then $\|w\| < \frac{r}{2}$.*

Proof. If $r = 2$, then $\|w\| = 0$, since the central basic factorization of w is of the overlapped form. If $r = 3$, then, depending of the type of the central basic factorization of w, we have $\|w\| = 1$ or $\|w\| = 0$. If $r > 3$, by the previous lemma, we have $\|w\| \leq \|x_2 \cdots x_{r-1}\| + 1$. It follows, by induction on r, that

$\|w\| \leq \frac{r-3}{2} + 1 = \frac{r}{2} - \frac{1}{2}$, in the case where r is odd, or $\|w\| \leq \frac{r}{2} - 1$, in the case where r is even. In any case, $\|w\| < \frac{r}{2}$. □

4. Representation of implicit operations on DA

4.1. **Two tree representations.** We present two distinct representations of the implicit operations on DA by trees. The first one comes from ideas used by Almeida and Weil [7] for representing implicit operations over R. In the second one, we extend to DA some techniques developed by Almeida and Zeitoun [8] to solve the word problem for ω-terms over R. We end the subsection by giving a more compact representation of the second tree representation using automata.

4.1.1. *Representation by finite-height trees.* Let A be a finite alphabet. We define the set $T_1(A)$ to be a set of trees of finite height and with a number of vertices that may be infinite. The set $T_1(A)$ consists of all trees with a root and satisfying the following conditions:

(1) The vertices which are direct descendants of a vertex v, and that we call the *progeny of* v, are ordered with order type $\mathbf{n}$, with n finite and odd, or with order type $\omega + \omega^*$.
(2) The direct descendants (or sons) of a vertex are ordered as follows: reading from left to right, and also from right to left, they switch between a vertex which we call *node* and a vertex which we call *leaf* and always starting, whether we read from the left or from the right, by a node.
(3) A node has one and only one of the following properties: either it has descendants, or it is labeled by 1 (in this case we call it a *degenerate node*).
(4) A leaf does not have descendants and it is labeled by a letter in A.
(5) The *content of a leaf* consists of its label; the *content of a node* is the set of labels of its descendants leaves; the content of a degenerate node is $\emptyset$.

For each non-degenerate node v, let v_i and f_i be, respectively, the i-th node and the i-th leaf, when we read from left to right, in the progeny of v and let v'_i and f'_i be, respectively, the i-th node and the i-th leaf, when we read from right to left, in the progeny of v. Three distinct cases can occur, depending if the progeny of v has order type $\omega + \omega^*$, $4\mathbf{m} - 1$ or $4\mathbf{m} + 1$:

(6) The case $\omega + \omega^*$: For all $i < \omega$, we have $c(f_i) \notin c(v_i)$, $c(f'_i) \notin c(v'_i)$ and $c(v_i) \cup c(f_i) = c(f'_i) \cup c(v'_i) = c(v)$.
(7) The case $4\mathbf{m} - 1$: For all $i \leq m$, we have $c(f_i) \notin c(v_i)$, $c(f'_i) \notin c(v'_i)$ and $c(v_i) \cup c(f_i) = c(f'_i) \cup c(v'_i) = c(v)$.
(8) The case $4\mathbf{m} + 1$: For all $i < m$, we have $c(f_i) \notin c(v_i)$, $c(f'_i) \notin c(v'_i)$ and $c(v_i) \cup c(f_i) = c(f'_i) \cup c(v'_i) = c(v)$. For $f = f_m$ or $f = f'_m$, then one and only one of the following cases can occur:
  (8-a) $c(f_m) \notin c(v_m)$, $c(f'_m) \notin c(v'_m)$, $c(v_m) \cup c(f_m) = c(f'_m) \cup c(v'_m) = c(v)$ and $c(v_{m+1}) \neq c(v)$.
  (8-b) $c(f_m) \notin c(v_{m+1}) \cup c(f'_m) \cup c(v'_m)$, $c(f'_m) \notin c(v_m) \cup c(f_m) \cup c(v_{m+1})$ and $c(v_m) \cup c(f_m) \cup c(v_{m+1}) \cup c(f'_m) = c(f_m) \cup c(v_{m+1}) \cup c(f'_m) \cup c(v'_m) = c(v)$.

Note that the contents of the successive descendants nodes of a branch strictly decrease. It follows that the height of $t \in T_1(A)$ is at most $|A|$.

We define the mapping $\varrho : T_1(A) \to \overline{\Omega}_A\mathsf{DA}$ as follows. Let $t \in T_1(A)$. We obtain $\varrho(t)$ as the ordered reading, from left to right, of the labels of the leaves of t. Formally, we define $\varrho(t)$ by induction on the height of t. If t has height 0, then t is only a vertex, which is a degenerate node, and we set $\varrho(t) = 1$. Suppose that t has non-zero height. Let $a_1, \ldots, a_n, b_n, \ldots, b_1$ be the labels of the leaves that are direct descendants of the root (in case the number of leaves is finite and odd, we have $a_1, \ldots, a_n, b_{n-1}, \ldots, b_1$, and, in case that this number is infinite, we have $a_1, a_2, \ldots \ \ldots, b_2, b_1$, respectively) and let $t_1, \ldots, t_n, t_{n+1}, t'_n, \ldots, t'_1$ (respectively, $t_1, \ldots, t_n, t'_n, \ldots, t'_1$ and $t_1, t_2, \ldots \ \ldots, t'_2, t'_1$) be the subtrees attached to each direct descendant node of the root. Note that each subtree has height strictly lower than the height of t. Let

$$\varrho(t) = \overrightarrow{\prod}_{i\geq 1}^{n}(\varrho(t_i)a_i) \cdot \varrho(t_{n+1}) \cdot \overleftarrow{\prod}_{i\geq 1}^{n}(b_i\varrho(t'_i))$$

respectively

$$\varrho(t) = \overrightarrow{\prod}_{i\geq 1}^{n}(\varrho(t_i)a_i) \cdot \varrho(t'_n) \cdot \overleftarrow{\prod}_{i\geq 1}^{n-1}(b_i\varrho(t'_i))$$

or

$$\varrho(t) = \overrightarrow{\prod}_{i\geq 1}(\varrho(t_i)a_i) \cdot \overleftarrow{\prod}_{i\geq 1}(b_i\varrho(t'_i)).$$

Lemma 4.1. *The factorization used in the definition of $\varrho(t)$ is the iterated central basic factorization of type 1 defined in the previous section.*

Proof. It suffices to note that, by properties (6), (7) and (8) from the definition of tree $t \in T_1(A)$, a_1 and b_1 are, respectively, the first occurrence of the last appearing letter when we read from left to right and from right to left. The other labels $a_2, \ldots, b_2$ in the progeny of the root satisfy the same condition in the subtree of t which results from eliminating the first two and the last two branches that leave from the root and so on, until the content of the subtree decreases, as in the iteration of type 1 of the central basic factorization defined previously. Since, by Proposition 3.1 this factorization is unique, it follows that the factorization of $\varrho(t)$ is the iterated central basic factorization of type 1. $\square$

Theorem 4.2. *The mapping $\varrho : T_1(A) \to \overline{\Omega}_A\mathsf{DA}$ is a bijection.*

Proof. Let t and $\bar{t}$ be distinct elements in $T_1(A)$. Let $a_1, a_2, \ldots, b_2, b_1$ be the labels of the leaves which are direct descendants of the root of t and let $t_1, t_2, \ldots, t'_2$, t'_1 be the subtrees attached to each direct descendant node of the root of t (as we have defined previously). Similarly, we define $\bar{a}_1, \bar{a}_2, \ldots, \bar{b}_2, \bar{b}_1$ and $\bar{t}_1, \bar{t}_2, \ldots, \bar{t}'_2, \bar{t}'_1$ with respect to $\bar{t}$. Let k be minimum with $t_k \neq \bar{t}_k$, $a_k \neq \bar{a}_k$, $t'_k \neq \bar{t}'_k$ or $b_k \neq \bar{b}_k$ and consider the k-th iteration of the central basic factorization of $\varrho(t)$ and $\varrho(\bar{t})$:

$$\varrho(t) = \varrho(t_1)a_1 \cdots \varrho(t_k)a_k\gamma_k b_k\varrho(t'_k) \cdots b_1\varrho(t'_1)$$

and

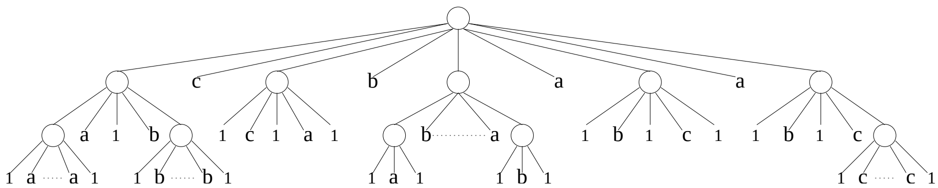
$$\varrho(\bar{t}) = \varrho(\bar{t}_1)\bar{a}_1 \cdots \varrho(\bar{t}_k)\bar{a}_k\bar{\gamma}_k\bar{b}_k\varrho(\bar{t}'_k) \cdots \bar{b}_1\varrho(\bar{t}'_1).$$

Note that, if we have $2n + 1$ nodes in the progeny of the root, with $k = n + 1$, i.e., when t and $\bar{t}$ differ in the subtree attached at the node in the central position, we only iterate $k - 1$ times. Proceeding by induction on the height of t, it follows that $\varrho(t_k) \neq \varrho(\bar{t}_k)$, $a_k \neq \bar{a}_k$, $\varrho(t'_k) \neq \varrho(\bar{t}'_k)$ or $b_k \neq \bar{b}_k$. Since, by Proposition 3.1, the central basic factorization is unique, we deduce that $\varrho(t) \neq \varrho(\bar{t})$.

Let $w \in \overline{\Omega}_A\mathsf{DA}$. To verify that the mapping is onto, we proceed by induction on $|c(w)|$. If $|c(w)| = 0$, then $w = 1$ and $w = \varrho(t)$, where t is a tree with just one vertex. If $|c(w)| \neq 0$, then $w = op^1_u(\alpha_1, \alpha_2, \ldots, \beta_2, \beta_1)$, where $u = a_1a_2 \cdots b_2b_1$ is an element in $A^+ \cup A^{\omega+\omega^*}$ and $\alpha_1, \alpha_2, \ldots, \beta_2, \beta_1$ satisfy the conditions used to define

op^1_u: for all i, $c(w) = c(\alpha_i a_i)$, $c(w) = c(b_i\beta_i)$, $a_i \notin c(\alpha_i)$ and $b_i \notin c(\beta_i)$. Moreover, if $|u|$ is finite and odd, then, since a_n is the letter at the central position, we also have $a_n \notin c(\beta_n)$. If $|u| = n$ is finite and even and the central basic factorization of w is of the overlapped form, we have that, for $k = n$, $a_n \notin c(\gamma_n) \cup \{b_n\} \cup c(\beta_n)$ and $b_n \notin c(\alpha_n) \cup \{a_n\} \cup c(\gamma_n)$. It follows that $c(\alpha_i)$ and $c(\beta_i)$ are strictly contained in $c(w)$, for all i. For all i, let t_i and t'_i be elements in $T_1(A)$ such that $\varrho(t_i) = \alpha_i$ and $\varrho(t'_i) = \beta_i$. Then $w = \varrho(t)$, where t is the subtree of $T_1(A)$ whose leaves of the progeny of the root are labeled with $a_1, a_2, \ldots, b_2, b_1$ and whose nodes of this progeny have the subtrees $t_1, t_2, \ldots, t'_2, t'_1$ attached, respectively. □

Example 4.3. The tree in $T_1(A)$ which represents $w = a^\omega b^\omega c^2(ab)^\omega cabc^\omega \in \overline{\Omega}_A\mathsf{DA}$ is the following:



4.1.2. *Representation by* quasi-ternary *trees.* Let A be a finite alphabet. We define a set $T_2(A)$ of *quasi-ternary* trees with labeled vertices, of which there may be infinitely many, and whose heights of the central branches can be infinite too. The set $T_2(A)$ consists of all trees with a root such that:

(i) The vertices are labeled with a pair of letters in A, or with a letter in A (*degenerate vertex*), or with a letter $\varepsilon \notin A$ (*final vertex*).
(ii) The vertices which are labeled with a pair of letters have three direct descendants. The degenerate vertices have only two direct descendants. The final vertices do not have descendants.

We define recursively the *content* of a vertex v, $c(v)$, to be the union of the contents of their sons with the set of its labels. The *left content* of a vertex v, $c_l(v)$, is the content of its left son. Dually, one may define the *right content* of a vertex v, $c_r(v)$. In the case where the vertex is non-degenerate, we also define the *central content* of a vertex v, $c_c(v)$, to be the content of its central son.

The set $T_2(A)$ also satisfies the following condition:

(iii) For each non-final vertex v one of the following cases occurs:
(iii-a) Non-degenerate case: Let (m_1, m_2) be the label of v. Then only one of the following cases occurs: either
$$m_1 \notin c_l(v),\ m_2 \notin c_r(v) \text{ and } c_l(v) \cup \{m_1\} = \{m_2\} \cup c_r(v) = c(v),$$
or
$$m_1 \notin c_c(v) \cup \{m_2\} \cup c_r(v),\ m_2 \notin c_l(v) \cup \{m_1\} \cup c_c(v) \text{ and}$$
$$c_l(v) \cup \{m_1\} \cup c_c(v) \cup \{m_2\} = \{m_1\} \cup c_c(v) \cup \{m_2\} \cup c_r(v) = c(v).$$
(iii-b) Degenerate case: Let m be the label of v. We have $m \notin c_l(v)$, $m \notin c_r(v)$ and $c_l(v) \cup \{m\} = \{m\} \cup c_r(v) = c(v)$.

Observe that the contents of the vertices of the successive descendants from the right or from the left branches strictly decrease. We define the *depth* of $t \in T_2(A)$, $\mathfrak{d}(t)$, to be the maximum distance to the root of vertices which do not have as an ancestor a vertex from a central branch. It follows that the depth of t is at most

$|A|$, but it can be strictly smaller. As an example, $\mathfrak{d}(a_1 \cdots a_n) = 1$, with $a_i \neq a_j$ if $i \neq j$.

We define the mapping $\rho : T_2(A) \to \overline{\Omega}_A\mathsf{DA}$ recursively as follows. Let $t \in T_2(A)$. If t has non-zero depth, i.e., t is just a vertex with label ε, we set $\rho(t) = 1$. Otherwise, let v_0, v_1 and v_2 (v_0 and v_2 in the degenerate case) be the vertices descending directly from the root v_ε. For each vertex v_i, with $i \in \{0,1,2\}^*$, let v_{i0}, v_{i1} and v_{i2} (v_{i0} and v_{i2} in the degenerate case) be the sons of v_i. Let $l_{(i,0)}$ and $l_{(i,2)}$ be the labels of v_i (in the degenerate case, let $l_{(i,0)}$ be this label). We denote by t_{i0}, t_{i1} and t_{i2} the subtrees that begin at the sons v_{i0}, v_{i1} and v_{i2}, respectively. Then, $\rho(t)$ is described as follows:

$$\rho(t) = \rho(t_0) \cdot l_{(\varepsilon,0)} \cdot \rho(t_{10}) \cdot l_{(1,0)} \cdots \cdots l_{(1,2)} \cdot \rho(t_{12}) \cdot l_{(\varepsilon,2)} \cdot \rho(t_2)$$

if the central body of the tree is infinite. If it is finite, we have one of the following cases:

$$\rho(t) = \rho(t_0) \cdot l_{(\varepsilon,0)} \cdots \rho(t_{1^{n-1}0}) \cdot l_{(1^{n-1},0)} \cdot l_{(1^{n-1},2)} \cdot \rho(t_{1^{n-1}2}) \cdots l_{(\varepsilon,2)} \cdot \rho(t_2)$$

or

$$\rho(t) = \rho(t_0) \cdot l_{(\varepsilon,0)} \cdots \rho(t_{1^{n-1}0}) \cdot l_{(1^{n-1},0)} \cdot \rho(t_{1^{n-1}2}) \cdots l_{(\varepsilon,2)} \cdot \rho(t_2).$$

Lemma 4.4. *The factorization used in the definition of $\rho(t)$ is the iterated central basic factorization of type 2 defined above.*

Proof. By Property (iii) from the tree's definition, we know that $l_{(\varepsilon,0)}$ and $l_{(\varepsilon,2)}$ are, respectively, the first occurrence of the last appearing letter when we read, respectively, from left to right and from right to left. The other labels from this factorization arise from the inductive process of applying the same factorization to the subtree of t that has as a root the central son of the root of t. This process ends when the root of the subtree that we are considering is degenerate or, if it is not degenerate, it has as a central son a final vertex. By Proposition 3.1, this factorization is unique and the iteration ends with the same condition of the iterated central basic factorization of type 2 of $\rho(t)$. Thus the factorizations are equal. $\square$

Theorem 4.5. *The mapping $\rho : T_2(A) \to \overline{\Omega}_A\mathsf{DA}$ is a bijection.*

Proof. Let t and $\bar{t}$ be distinct elements of $T_2(A)$. Let k be minimum for $l_{(1^k,0)} \neq \bar{l}_{(1^k,0)}$, $l_{(1^k,2)} \neq \bar{l}_{(1^k,2)}$, $t_{1^k0} \neq \bar{t}_{1^k0}$ or $t_{1^k2} \neq \bar{t}_{1^k2}$. Consider the $k+1$-iteration of the central basic factorization of $\rho(t)$ and $\rho(\bar{t})$:

$$\rho(t) = \rho(t_0) \cdot l_{(\varepsilon,0)} \cdots \rho(t_{1^k0}) \cdot l_{(1^k,0)} \cdot \rho(t_{1^{k+1}}) \cdot l_{(1^k,2)} \cdot \rho(t_{1^k2}) \cdots l_{(\varepsilon,2)} \cdot \rho(t_2)$$

and

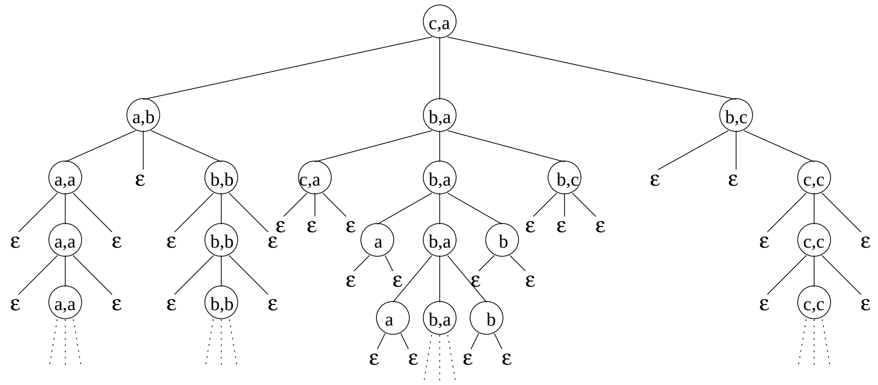
$$\rho(\bar{t}) = \rho(\bar{t}_0) \cdot \bar{l}_{(\varepsilon,0)} \cdots \rho(\bar{t}_{1^k0}) \cdot \bar{l}_{(1^k,0)} \cdot \rho(\bar{t}_{1^{k+1}}) \cdot \bar{l}_{(1^k,2)} \cdot \rho(\bar{t}_{1^k2}) \cdots \bar{l}_{(\varepsilon,2)} \cdot \rho(\bar{t}_2).$$

Proceeding by induction on the depth of t and $\bar{t}$, it follows that $l_{(1^k,0)} \neq \bar{l}_{(1^k,0)}$, $l_{(1^k,2)} \neq \bar{l}_{(1^k,2)}$, $\rho(t_{1^k0}) \neq \rho(\bar{t}_{1^k0})$ or $\rho(t_{1^k2}) \neq \rho(\bar{t}_{1^k2})$. Since, by Proposition 3.1, this factorization is unique, it follows that $\rho(t) \neq \rho(\bar{t})$.

Let $w \in \overline{\Omega}_A\mathsf{DA}$. To verify that ρ is onto, we proceed by induction on $|c(w)|$. If $|c(w)| = 0$, then $w = 1$ and $w = \rho(t)$, where t is the tree with just one vertex labeled by ε. If $|c(w)| \neq 0$, then we consider the iterated central basic factorization of w, $w = \alpha_1 a_1 \alpha_2 a_2 \cdots \cdots b_2 \beta_2 b_1 \beta_1$ (or one of the other previously described cases). Note that $c(\alpha_k)$ and $c(\beta_k)$ are strictly contained in $c(w)$, for all k. For all k, let $t_{1^{k-1}0}$ and $t_{1^{k-1}2}$ be elements in $T_2(A)$ such that $\rho(t_{1^{k-1}0}) = \alpha_k$ and

$\rho(t_{1^{k-1}2}) = \beta_k$. Then $w = \rho(t)$, where t is the tree in $T_2(A)$ whose central vertices are labeled with (a_1, b_1), (a_2, b_2), etc., and with the subtrees $t_{1^{k-1}0}$ and $t_{1^{k-1}2}$, for each k, respectively attached. □

Example 4.6. We have the following representation by a tree in $T_2(A)$ of the pseudoword $w = a^\omega b^\omega c^2(ab)^\omega cabc^\omega \in \overline{\Omega}_A\mathsf{DA}$:



4.1.3. *Representation by automata.* It is sometimes convenient to compress the tree representation described in 4.1.2. We do it by identifying vertices which have the same attached subtrees. We begin with the definition of DA-automaton.

An A-labeled DA-*automaton* is a tuple $\mathcal{A} = (V, \to, q, F, \lambda)$ where $(V, \to, q, F)$ is a non-empty deterministic automaton (which may be infinite) over the alphabet $\mathbb{B} = \{0, 1, 2\}$ and $\lambda : V \to A \times A \cup A \cup \{\varepsilon\}$ is a total function. It also satisfies the following conditions:

(A1) The set of final states is $F = \lambda^{-1}(\varepsilon)$.
(A2) There are no outgoing transitions from F.
(A3) Let $v \in V \backslash F$. Then both $v.0$ and $v.2$ are defined. We also have $v.1$ defined if and only if $\lambda(v) \in A \times A$. Otherwise, $\lambda(v) \in A$.
(A4) Given $v \in V \backslash F$ with $\lambda(v) = (\lambda^1(v), \lambda^2(v)) \in A \times A$ we have one and only one of the following cases:
  (i) $\lambda^1(v.\mathbb{B}^*) \cup \lambda^2(v.\mathbb{B}^*) = [\lambda^1(v.0\mathbb{B}^*) \cup \lambda^2(v.0\mathbb{B}^*)] \mathbin{\mathring{\cup}} \lambda^1(v)$ and $\lambda^1(v.\mathbb{B}^*) \cup \lambda^2(v.\mathbb{B}^*) = [\lambda^1(v.2\mathbb{B}^*) \cup \lambda^2(v.2\mathbb{B}^*)] \mathbin{\mathring{\cup}} \lambda^2(v)$,
  (ii) $\lambda^1(v.\mathbb{B}^*) \cup \lambda^2(v.\mathbb{B}^*) = \lambda^1(v) \mathbin{\mathring{\cup}} [\lambda^1(v.1\mathbb{B}^*) \cup \lambda^2(v.1\mathbb{B}^*) \cup \lambda^2(v) \cup \lambda^1(v.2\mathbb{B}^*) \cup \lambda^2(v.2\mathbb{B}^*)]$ and $\lambda^1(v.\mathbb{B}^*) \cup \lambda^2(v.\mathbb{B}^*) = [\lambda^1(v.0\mathbb{B}^*) \cup \lambda^2(v.0\mathbb{B}^*) \cup \lambda^1(v) \cup \lambda^1(v.1\mathbb{B}^*) \cup \lambda^2(v.1\mathbb{B}^*)] \mathbin{\mathring{\cup}} \lambda^2(v)$,

  where we consider, for each vertex v' such that $\lambda(v') \in A$, $\lambda(v') = \lambda^1(v')$.
(A5) Given $v \in V \backslash F$ with $\lambda(v) = \lambda^1(v) \in A$ we have

$$\lambda^1(v.\mathbb{B}^*) \cup \lambda^2(v.\mathbb{B}^*) = [\lambda^1(v.0\mathbb{B}^*) \cup \lambda^2(v.0\mathbb{B}^*)] \mathbin{\mathring{\cup}} \lambda^1(v)$$

$$\text{and } \lambda^1(v.\mathbb{B}^*) \cup \lambda^2(v.\mathbb{B}^*) = [\lambda^1(v.2\mathbb{B}^*) \cup \lambda^2(v.2\mathbb{B}^*)] \mathbin{\mathring{\cup}} \lambda^1(v)$$

  where, again, we consider, for each vertex v' such that $\lambda(v') \in A$, $\lambda(v') = \lambda^1(v')$.

The trees defined in 4.1.2 are DA-automata such that every state is reached from the initial state by a unique path. We call these trees DA-*trees*.

Given a DA-automaton $\mathcal{A} = (V, \to, q, F, \lambda)$ and $v \in V$, the *subautomaton of $\mathcal{A}$ with initial state v* is the DA-automaton $\mathcal{A}_v = (V \cap v.\mathbb{B}^*, \to, v, F \cap v.\mathbb{B}^*, \lambda_{|V \cap v.\mathbb{B}^*})$.

Observe that, by conditions (A4) and (A5), if $v.\alpha$ is defined, then $|\alpha|_0 + |\alpha|_2 \leq |A|$. Moreover, each time we follow a transition labeled 0 or 2, we end up in a subautomaton labeled by an alphabet strictly contained in the previous one. It follows that, if $p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} p_n = p_0$ is a closed path of $\mathcal{A}$, then $a_i = 1$ for all $i = 1, \ldots, n$.

We say that two DA-automata $\mathcal{A} = (V_i, \to_i, q_i, F_i, \lambda_i)$ $(i = 0, 1)$ are *isomorphic* if there is a bijection $\varphi : V_0 \to V_1$ such that, for all $v \in V$ and for all $a \in \mathbb{B}$, $\varphi(v.a) = \varphi(v).a$ and $\lambda_1(\varphi(v)) = \lambda_0(v)$.

We say that two DA-automata $\mathcal{A} = (V_i, \to_i, q_i, F_i, \lambda_i)$ $(i = 0, 1)$ are *k-equivalent* if

$$\text{for all } \alpha \in \mathbb{B}^*,\ |\alpha| \leq k \Longrightarrow \lambda_0(q_0.\alpha) = \lambda_1(q_1.\alpha).$$

We say that two DA-automata are *equivalent* if they are *k-equivalent* for all $k \geq 0$. We write $\mathcal{A}_0 \sim_k \mathcal{A}_1$ to denote the k-equivalence and we set $\sim = \bigcap \sim_k$. Note that equivalent DA-trees are isomorphic, since, as we said above, each state is completely determined by the unique path starting at the initial state and ending at this state.

Lemma 4.7. *Any DA-automaton has a unique (up to isomorphism of DA-trees) equivalent DA-tree.*

Proof. Let $\mathcal{A} = (V, \to, q, F, \lambda)$ be a DA-automaton. We define the DA-tree $\mathcal{T} = (W, \to, p, G, \nu)$ as follows. Let $W = \{\alpha \in \mathbb{B}^* \mid q.\alpha \text{ is defined}\}$, $p = \varepsilon$, $\nu(\alpha) = \lambda(q.\alpha)$ and $G = \nu^{-1}(\varepsilon)$. If $q.\alpha0$ and $q.\alpha2$ are defined, i.e., if $\lambda(q.\alpha) \neq \varepsilon$, then we define the transitions $\alpha \xrightarrow{0} \alpha0$ and $\alpha \xrightarrow{2} \alpha2$. Moreover, if $q.\alpha1$ is also defined, which corresponds to $\lambda(q.\alpha) \in A \times A$, then we have also the transition $\alpha \xrightarrow{1} \alpha1$ of $\mathcal{T}$. It is easy to see that the properties (A1)-(A5) remain valid and the uniqueness of this construction results from the fact that equivalent DA-trees are isomorphic. □

We call the *unfolding* of $\mathcal{A}$ the unique (up to isomorphism) DA-tree $\vec{\mathcal{A}}$ equivalent to the DA-automaton $\mathcal{A}$.

Corollary 4.8. *Let $\mathcal{A}$ and $\mathcal{A}'$ be DA-automata. Then $\mathcal{A} \sim \mathcal{A}'$ if and only if $\vec{\mathcal{A}} = \vec{\mathcal{A}}'$.*

We define the *value* $\pi(\mathcal{A}) \in \overline{\Omega}_A\mathsf{DA}$ of a DA-automaton $\mathcal{A}$ by $\pi(\mathcal{A}) = \rho(\vec{\mathcal{A}})$. Given a DA-automaton $\mathcal{A} = (V, \to, q, F, \lambda)$ and $v \in V$, let $[v] = \pi(\mathcal{A}_v)$.

Lemma 4.9. *Let $\mathcal{A} = (V, \to, q, F, \lambda)$ be a DA-automaton and $v \in V \backslash F$. Then, the central basic factorization of $[v]$ is $[v.0] \cdot \lambda^1(v) \cdot [v.1] \cdot \lambda^2(v) \cdot [v.2]$, if $v.1$ is defined, or $[v.0] \cdot \lambda^1(v) \cdot [v.2]$, otherwise. Thus, by uniqueness of the central basic factorization, we have, for $u, v \in V \backslash F$,*

$$[u] = [v] \Longrightarrow \begin{cases} \lambda^1(u) = \lambda^1(v) \\ \lambda^2(u) = \lambda^2(v) \\ [u.0] = [v.0] \\ [u.1] = [v.1] \\ [u.2] = [v.2] \end{cases}$$

respectively,

$$[u] = [v] \Longrightarrow \begin{cases} \lambda^1(u) = \lambda^1(v) \\ [u.0] = [v.0] \\ [u.2] = [v.2]. \end{cases}$$

Proof. It suffices to proceed by induction on $c([v])$ taking into account the definition of central basic factorization. □

We call the *wrapping* of a DA-automaton $\mathcal{A} = (V, \to, q, F, \lambda)$ the DA-automaton $[\mathcal{A}] = ([V], \to, [q], [F], \nu)$ where:

(i) $[V] = \{[v] \mid v \in V\} \subseteq \overline{\Omega}_A\mathsf{DA}$;
(ii) $[v].0 = [v.0]$, $[v].1 = [v.1]$ and $[v].2 = [v.2]$;
(iii) $\nu([v]) = \lambda(v)$.

This automaton is obtained from $\mathcal{A}$ by identifying states representing the same pseudoword. For $w \in \overline{\Omega}_A\mathsf{DA}$ we define the *wrapped* DA-*automaton* of w to be $\mathcal{A}(w) = [\rho^{-1}(w)]$.

The *value* of a path $p_0 \xrightarrow{\delta_0} p_1 \xrightarrow{\delta_1} \cdots \xrightarrow{\delta_n} p_{n+1}$ in a DA-automaton is the product $\prod_{i=0}^{n}(\delta_i, \lambda(p_i)) \in (\mathbb{B} \times (A \times A \cup A))^*$. The *language* $\mathcal{L}(v) \subseteq (\mathbb{B} \times (A \times A \cup A))^*$ *associated to the state* $v \in \mathcal{A}$ is the set of values of all the paths starting at v and ending at a final state. The *language* $\mathcal{L}(\mathcal{A})$ *associated to* $\mathcal{A}$ is the language associated to its initial state. Finally, the *language* $\mathcal{L}(w)$ *associated to* w is $\mathcal{L}(w) = \mathcal{L}(\mathcal{A}(w))$.

Note that, if we consider the automaton obtained from $\mathcal{A}(w)$ by replacing the label of each edge in $\mathcal{A}(w)$ by the pair whose first component is the label that this edge has in $\mathcal{A}(w)$ and the second component is the label of the initial vertex of the edge in $\mathcal{A}(w)$, then it is an automaton that recognizes $\mathcal{L}(w)$. We see this replacement in Example 4.12.

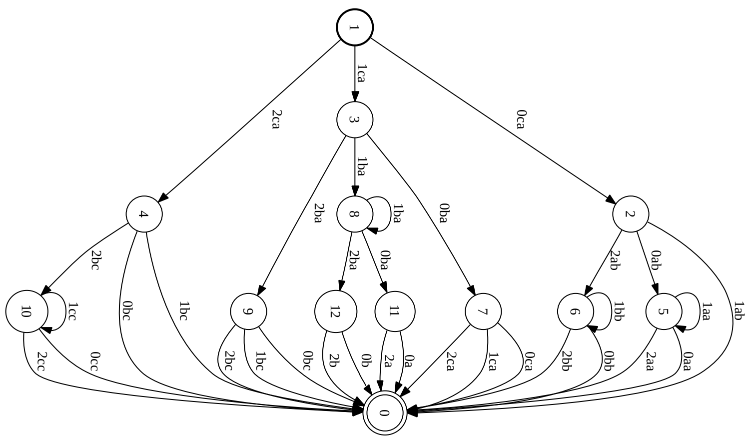
Lemma 4.10. *Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two* DA*-automata. We have $\mathcal{L}(\mathcal{A}_1) = \mathcal{L}(\mathcal{A}_2)$ if and only if $\vec{\mathcal{A}}_1 = \vec{\mathcal{A}}_2$.*

Proof. It suffices to note that $\mathcal{L}(\mathcal{A})$ uniquely determines the maximal paths in $\mathcal{A}$, which in turn determines $\vec{\mathcal{A}}$. □

Proposition 4.11. *Let $v, w \in \overline{\Omega}_A\mathsf{DA}$. Then $\mathsf{DA} \models v = w$ if and only if $\mathcal{L}(v) = \mathcal{L}(w)$.*

Proof. By Theorem 4.5, we have $t(v) = t(w)$ if and only if $\mathsf{DA} \models v = w$. The result now follows from the previous lemma. □

Example 4.12. The wrapped DA-automaton of $w = a^\omega b^\omega c^2 (ab)^\omega cabc^\omega \in \overline{\Omega}_A\mathsf{DA}$ is presented in the following figure. Note that it is a finite automaton.



4.2. **Representation by labeled orderings.** We consider, in the set of linear orderings, the subalgebra generated by the ordinal 1 and by the following operations:

$$+ : (o_1, o_2) \mapsto o_1 + o_2,$$
$$\theta : (o_1, o_2, \dots, \dots, o_2', o_1') \mapsto o_1 + o_2 + \cdots \; \cdots + o_2' + o_1'.$$

We call an element of this subalgebra a *$*$-linear ordering*. Note that a $*$-linear ordering has the following properties: it is countable since it is a countable sum of countable orderings, it has a minimum element and a maximum element, any element except the maximum has a successor and any element except the minimum has a predecessor.

Given a $*$-linear ordering o, we consider a representation of o in the free algebra generated by $\{1\}$ and by the operations $+$ and θ, $\langle\{1\}; +, \theta\rangle$. We define the *rank of this representation* to be the maximum number of nested operations θ. The *rank of o*, $\mathfrak{r}(o)$, is the minimum of the ranks of the representations of o.

Lemma 4.13. *Given a $*$-linear ordering o, any closed interval in o is also a $*$-linear ordering.*

Proof. We proceed by induction on the *rank* of o. The case $\mathfrak{r}(o) = 0$ is trivial, since it corresponds to the finite linear orderings, and the case $\mathfrak{r}(o) = 1$ is easy to show: if o is a $*$-linear ordering such that $\mathfrak{r}(o) = 1$, then $o = (\omega + \omega^*)m$, with m finite, and the closed intervals in o are of the form m, with m finite, or $(\omega + \omega^*)m'$, with $m' \leq m$. In any case, they also are $*$-linear orderings. Suppose that the result is verified for $\mathfrak{r}(o) < n$. Let o be a $*$-linear ordering such that $\mathfrak{r}(o) = n > 1$. Then $o = \theta(o_1, o_2, \dots, o_2', o_1')m$ for some representation of o, where o_i and o_i' are $*$-linear orderings such that $\mathfrak{r}(o_i) < n$ and $\mathfrak{r}(o_i') < n$, for all i, and m is finite. Let o' be a closed interval in o. Then one of the following cases can occur, for some integers j and k:

(1) $o' \subseteq o_j$;
(2) $o' \subseteq o_k'$;
(3) $o' \subseteq o_j + o_{j+1} + \cdots + o_{k-1} + o_k$ and $o_{j+1} + \cdots + o_{k-1} \subset o'$;
(4) $o' \subseteq o_j' + o_{j-1}' + \cdots + o_{k+1}' + o_k'$ and $o_{j-1}' + \cdots + o_{k+1}' \subset o'$;
(5) $o' \subseteq o_j + o_{j+1} + \cdots \; \cdots + o_{k+1}' + o_k'$ and $o_{j+1} + \cdots \; \cdots + o_{k+1}' \subset o'$;
(6) $o' \subseteq \bar{o} + \theta(o_1, o_2, \dots, \dots, o_2', o_1')m' + \bar{o}'$ and $\theta(o_1, o_2, \dots, \dots, o_2', o_1')m' \subset o$, with $0 < m' < m$, $\bar{o}$ is of the form (2), (4) or (5) with $k = 1$, and $\bar{o}'$ is of the form (1), (3) or (5) with $j = 1$.

We treat the case (5) in detail. The other cases are similar. In this case, we have that $o_j \cap o'$ and $o_k' \cap o'$ are closed intervals in o_j and o_k', respectively, and $\mathfrak{r}(o_j) < n$ and $\mathfrak{r}(o_k') < n$. By induction hypothesis, we have that $o_j \cap o'$ and $o_k' \cap o'$ are $*$-linear orderings. It follows that $o' = \theta(o_j \cap o', o_{j+1}, \dots, \dots, o_{k+1}', o_k' \cap o')$ is a $*$-linear ordering. □

Given a finite alphabet A, let $\mathbf{LO}^*(A)$ be the set of *$*$-labeled linear orderings in A*, where an element $(o, l) \in \mathbf{LO}^*(A)$ is such that o is a linear ordering in $\langle\{1\}; +, \theta\rangle$ and l is a labeling $l : o \to A$.

Let $\mathbf{o} = (o, l)$ be a $*$-labeled linear ordering. A partition of $\mathbf{o}$ in two non-empty intervals $(\mathbf{o_1}, \mathbf{o_2})$, where each element of $\mathbf{o_2}$ is greater than all elements of $\mathbf{o_1}$, is called a *Dedekind cut in* $\mathbf{o}$. We say that a Dedekind cut $(\mathbf{o_1}, \mathbf{o_2})$ is a *gap in* $\mathbf{o}$ if the first interval does not have a maximum and the second interval does not have a minimum. An ordering is *complete* if it does not have any gap. Given an incomplete ordering $\mathbf{o} = (o, l) \in \mathbf{LO}^*(A)$, its completion is isomorphic to the set of Dedekind cuts in $\mathbf{o}$ ordered by the relation $(\mathbf{o_1}, \mathbf{o_2}) \leq (\mathbf{o_1'}, \mathbf{o_2'})$ if $\mathbf{o_1} \subseteq \mathbf{o_1'}$. For details, see Rosenstein [12].

Given a Dedekind cut $(\mathbf{o_1}, \mathbf{o_2})$ in $\mathbf{o}$, we define the *set of right cofinal letters of* $\mathbf{o_1}$, $c_r(\mathbf{o_1})$ and the *set of left cofinal letters of* $\mathbf{o_2}$, $c_l(\mathbf{o_2})$ to be the following:

$$c_r(\mathbf{o_1}) = \{a \in A \mid \forall p \in o_1 \ \exists q \in o_1 : p < q \wedge l(q) = a\},$$
$$c_l(\mathbf{o_2}) = \{a \in A \mid \forall p \in o_2 \ \exists q \in o_2 : q < p \wedge l(q) = a\}.$$

We say that a $*$-linear ordering $\mathbf{o}$ satisfies the *cofinal property* if, for every Dedekind cut $(\mathbf{o_1}, \mathbf{o_2})$ in $\mathbf{o}$, $c_r(\mathbf{o_1}) = c_l(\mathbf{o_2})$. We call this set the *cofinal set* of $(\mathbf{o_1}, \mathbf{o_2})$ and we write $c((\mathbf{o_1}, \mathbf{o_2})) = c_r(\mathbf{o_1}) = c_l(\mathbf{o_2})$.

Let $\mathbf{o} \in \mathbf{LO}^*(A)$ be a labeled linear ordering satisfying the cofinal property. We define the following labeling function of the set $\mathcal{D}_o$ of Dedekind cuts of $\mathbf{o}$:

$$\begin{array}{rcl} l: \ \mathcal{D}_o & \rightarrow & \mathcal{P}(A) \\ (\mathbf{o_1}, \mathbf{o_2}) & \mapsto & \begin{cases} \{a\} & \text{if } (\mathbf{o_1}, \mathbf{o_2}) \text{ is not a gap and } l(\max \mathbf{o_1}) = a \\ c((\mathbf{o_1}, \mathbf{o_2})) & \text{if } (\mathbf{o_1}, \mathbf{o_2}) \text{ is a gap.} \end{cases} \end{array}$$

We say that $\mathbf{o} = (o, l) \in \mathbf{LO}^*(A)$ is a *reduced A-labeled $*$-linear ordering* if it satisfies the following conditions:

(i) $\mathbf{o}$ satisfies the cofinal property;
(ii) Given two distinct gaps in $\mathbf{o}$ with the same cofinal set, there exists a Dedekind cut between them whose label is not contained in this cofinal set.

Let $\mathbf{rLO}^*(A)$ be the set of all reduced A-labeled $*$-linear orderings. To each tree $t \in T_1(A)$ we associate a $*$-labeled linear ordering $\mu(t)$ by ordering the set of the leaves of t from left to right. Formally, the set of the leaves of t, $F(t)$, is ordered as follows: given two elements f and f', let $v_{ff'}$ be the deepest node which is a common ancestor of f and f', and v_f and $v_{f'}$ the sons of $v_{ff'}$ which are ancestors of f and f', respectively (note that we can have $v_f = f$ or $v_{f'} = f'$). If $v_f < v_{f'}$ in the progeny of $v_{ff'}$, then we say that $f < f'$. It is easy to verify that this defines a linear ordering in $F(t)$. Let o be the corresponding linear ordering and let $l : o \rightarrow A$ be the labeling function which maps each leaf to its label. We put $\mu(t) = (o, l)$.

Proposition 4.14. *For each $t \in T_1(A)$, $\mu(t) \in \mathbf{rLO}^*(A)$.*

Proof. We proceed by induction on the height of t, $\mathfrak{h}(t)$, where the case $\mathfrak{h}(t) = 0$ is trivial. Recall that the order type of each progeny is $\mathbf{m}$, with m finite, or $\omega + \omega^*$. Let t be a tree with non-zero height and let $t_1, t_2, \ldots, \ldots, t_2', t_1'$ be the subtrees attached to the nodes which are sons of the root of t. By induction hypothesis, $o(t_i), o(t_i') \in \langle \{1\}; +, \theta \rangle$ for all i. It follows that $o(t) = o(t_1) + 1 + o(t_2) + 1 + \cdots \cdots + 1 + o(t_2') + 1 + o(t_1')$ is an element of the subalgebra, since it is a finite or infinite sum of elements of the subalgebra (in the first case, we apply a sufficient number of times the operator $+$ and, in the second case, we apply the operator θ). Since the height of $t \in T_1(A)$ is at most $|A|$, it follows that $\mu(t) \in \mathbf{LO}^*(A)$.

Let $(\mathbf{o_1}, \mathbf{o_2})$ be a Dedekind cut in $\mu(t)$. If this cut is not a gap in $\mu(t)$, then $c_r(\mathbf{o_1}) = c_l(\mathbf{o_2}) = \emptyset$. Suppose that it is a gap in $\mu(t)$. This corresponds to dividing the tree in the middle of the progeny of a node v with order $\omega + \omega^*$. By Property (4) from the definition of tree in $T_1(A)$, we have that $c(v_i) \cup c(f_i) = c(v_i') \cup c(f_i') = c(v)$, for all i, where v_i and f_i are, respectively, the i-th node and the i-th leaf of the progeny of v, when we count from left to right, and v_i' and f_i' are, respectively, the i-th node and the i-th leaf of the progeny of v, when we count from right to left. Recall that the content of a node is the set of the labels of the descendants leaves. It follows that $c_r(\mathbf{o_1}) = c_l(\mathbf{o_2}) = c(v)$ and, therefore, $\mu(t)$ satisfies the cofinal property.

Now, we show that condition (ii) holds. Let $(\mathbf{o_1}, \mathbf{o_2}) < (\mathbf{o_1'}, \mathbf{o_2'})$ be two distinct gaps in $\mathbf{o}$ with the same cofinal set and let v and v' be the nodes whose progenies are split in two by the intervals of the gaps $(\mathbf{o_1}, \mathbf{o_2})$ and $(\mathbf{o_1'}, \mathbf{o_2'})$, respectively. Let $\bar{v}$ and $\bar{v}'$ be the the deepest ancestor of v and v', respectively, such that $\bar{v}$ and $\bar{v}'$ are in the same progeny. Note that $\bar{v} < \bar{v}'$ and that $c((\mathbf{o_1}, \mathbf{o_2})) = c((\mathbf{o_1'}, \mathbf{o_2'})) \subseteq c(\bar{v})$ and $c((\mathbf{o_1}, \mathbf{o_2})) = c((\mathbf{o_1'}, \mathbf{o_2'})) \subseteq c(\bar{v}')$. Suppose that this progeny has order $\omega + \omega^*$. If $\bar{v}$ is in a position of ω, let f be the leaf that succeeds $\bar{v}$ in the progeny and, in case $\bar{v}$ is in a position of ω^*, and so is $\bar{v}'$, let f be the leaf that precedes $\bar{v}'$ in the progeny. Then, by Property (6) from the definition of tree, we have $c(f) \notin c(\bar{v})$ or $c(f) \notin c(\bar{v}')$. We consider the Dedekind cut in $\mathbf{o}$ and between the two gaps, $(\mathbf{o_1''}, \mathbf{o_2''})$, such that $\max o_1'' = f$. It follows that $c((\mathbf{o_1''}, \mathbf{o_2''})) \not\subseteq c((\mathbf{o_1}, \mathbf{o_2})) = c((\mathbf{o_1'}, \mathbf{o_2'}))$. The other cases hold similarly. Thus $\mu(t)$ is reduced. $\square$

We denote by $\overrightarrow{\sum}_{k=1}^{n} o_k$ the sum $o_1 + o_2 + \cdots + o_n$ and by $\overleftarrow{\sum}_{k=1}^{n} o_k$ the sum $o_n + \cdots + o_2 + o_1$.

Lemma 4.15. *Let $t \in T_1(A)$ be a tree with non-zero height, let $a_1, a_2, \ldots, \ldots, b_2, b_1$ be the labels of the leaves that are sons of the root and let $t_1, t_2, \ldots, \ldots, t_2', t_1'$ be the subtrees of t attached to each son of the root. If $\mu(t_m) = (o_m, l_m)$ and $\mu(t_m') = (o_m', l_m')$, for all m, then $\mu(t) = (o, l)$ where o is a $*$-linear ordering of one of the following forms:*

(1) $\overrightarrow{\sum}_{n\geq 1}(o_n + 1) + \overleftarrow{\sum}_{n\geq 1}(1 + o_n')$, *if the leaf sons have order $\omega + \omega^*$;*
(2) $\overrightarrow{\sum}_{i=1}^{n}(o_i + 1) + o_{n+1} + \overleftarrow{\sum}_{i=1}^{n}(1 + o_i')$, *if the number of leaf sons is even;*
(3) $\overrightarrow{\sum}_{i=1}^{n}(o_i + 1) + o_n' + \overleftarrow{\sum}_{i=1}^{n-1}(1 + o_i')$, *if the number of leaf sons is odd;*

and l is the labeling $l : o \to A$ satisfying the following conditions:

(i) $l\left(\overrightarrow{\sum}_{i=1}^{m-1}(o_i + 1) + \gamma\right) = l_m(\gamma)$, *if γ is an initial segment of o_m,*
(ii) $l\left(\overrightarrow{\sum}_{i\geq 1}(o_i + 1) + \overleftarrow{\sum}_{i>m}(1 + o_i') + 1 + \gamma\right) = l_m'(\gamma)$, *if γ is an initial segment of o_m',*
(iii) $l\left(\overrightarrow{\sum}_{i=1}^{m}(o_i + 1)\right) = a_m$,
(iv) $l\left(\overrightarrow{\sum}_{i\geq 1}(o_i + 1) + \overleftarrow{\sum}_{i>m}(1 + o_i') + 1\right) = b_m$.

Proof. The verification follows directly by induction on the height of t. $\square$

To establish that μ is a bijection, we construct and iterate a central basic partition of a non-zero reduced $*$-labeled linear ordering. We start with the following lemma.

Lemma 4.16. *Let $\mathbf{o} = (o, l) \in \mathbf{rLO}^*(A)$. Then, for each $a \in c(\mathbf{o})$, there exist the smallest position of $\mathbf{o}$ labeled a and the largest position of $\mathbf{o}$ labeled a.*

Proof. For each $a \in c(\mathbf{o})$, let $D_a = \{(\mathbf{o_1}, \mathbf{o_2}) \in \mathcal{D}_o \mid a \in l((\mathbf{o_1}, \mathbf{o_2}))\}$. This subset is non-empty and bounded below, so, since $\mathcal{D}_o$ is a complete ordering, it has an infimum, $(\mathbf{o_1}, \mathbf{o_2})$. Suppose that it is a gap in $\mathbf{o}$. By definition of infimum, for every Dedekind cut $(\mathbf{o_1'}, \mathbf{o_2'}) > (\mathbf{o_1}, \mathbf{o_2})$, there exists a Dedekind cut $(\mathbf{o_1''}, \mathbf{o_2''})$ such that $(\mathbf{o_1}, \mathbf{o_2}) \leq (\mathbf{o_1''}, \mathbf{o_2''}) < (\mathbf{o_1'}, \mathbf{o_2'})$ and $a \in l((\mathbf{o_1''}, \mathbf{o_2''}))$. But this is equivalent to $a \in c((\mathbf{o_1}, \mathbf{o_2}))$ and, therefore, there exists a sequence of Dedekind cuts less than $(\mathbf{o_1}, \mathbf{o_2})$ whose label contains a, which is contrary to the hypothesis of $(\mathbf{o_1}, \mathbf{o_2})$ be an infimum. Thus $(\mathbf{o_1}, \mathbf{o_2})$ is not a gap and, therefore, it is minimum for, otherwise, its successor will be also a lower bound of D_a. Since $(\mathbf{o_1}, \mathbf{o_2})$ is not a gap, $\mathbf{o_1}$ has a maximum and $\mathbf{o_2}$ has a minimum. The maximum of $\mathbf{o_1}$ is the smallest position of $\mathbf{o}$ labeled a. Similarly, we prove the existence of the largest position of $\mathbf{o}$ labeled a. □

Let $\mathbf{o} = (o, l) \in \mathbf{rLO}^*(A)$. For each letter $a \in A$, let p_a^o be the smallest position of $\mathbf{o}$ such that $l(p) = a$, let $\bar{p}_a^o$ be the largest position of $\mathbf{o}$ such that $l(p) = a$ and let $p^o = \max\{p_a^o \mid a \in A\}$ and $\bar{p}^o = \min\{\bar{p}_a^o \mid a \in A\}$ (if there is no position labeled by a, we set $p_a^o = \min o$ and $\bar{p}_a^o = \max o$). Note that, since A is finite, there exist p^o and $\bar{p}^o$. Three cases can occur: $p^o < \bar{p}^o$, $p^o > \bar{p}^o$ or $p^o = \bar{p}^o$. We begin with the case $p^o < \bar{p}^o$. Let $\alpha_1 = [\min o, p^o[$, $\gamma_1 =]p^o, \bar{p}^o[$ and $\beta_1 =]\bar{p}^o, \max o]$, which are also $*$-linear orderings by Lemma 4.13, since they are closed intervals in o. We have $o = \alpha_1 + 1 + \gamma_1 + 1 + \beta_1$. We call this equality the *central basic partition of* (o, l). Let $\gamma_0 = o$, $p_1 = p^o$ and $\bar{p}_1 = \bar{p}^o$. While $p_i < \bar{p}_i$ and $c(\gamma_i) = c(o)$, let $p_{i+1} = \max\{p_a^o \in \gamma_i \mid a \in A\}$, $\bar{p}_{i+1} = \min\{\bar{p}_a^o \in \gamma_i \mid a \in A\}$ and $\alpha_{i+1} = [\min \gamma_i, p_{i+1}[$, $\gamma_{i+1} =]p_{i+1}, \bar{p}_{i+1}[$ and $\beta_{i+1} =]\bar{p}_{i+1}, \max \gamma_i]$ be the $*$-linear orderings such that $\gamma_i = \alpha_{i+1} + 1 + \gamma_{i+1} + 1 + \beta_{i+1}$. If, for any k, $c(\gamma_{k+1}) \neq c(o)$, or $p_k > \bar{p}_k$, or $p_k = \bar{p}_k$, then we put, respectively:

(i) $\gamma_k = \alpha_{k+1} + 1 + \gamma_{k+1} + 1 + \beta_{k+1}$, where $\alpha_{k+1} = [\min \gamma_k, p_{k+1}[$, $\gamma_{k+1} =]p_{k+1}, \bar{p}_{k+1}[$ and $\beta_{k+1} =]\bar{p}_{k+1}, \max \gamma_k]$,

(ii) $\gamma_k = \alpha_{k+1} + 1 + \gamma_{k+1} + 1 + \beta_{k+1}$, where $\alpha_{k+1} = [\min \gamma_k, \bar{p}_{k+1}[$, $\gamma_{k+1} =]\bar{p}_{k+1}, p_{k+1}[$, and $\beta_{k+1} =]p_{k+1}, \max \gamma_k]$,

(iii) $\gamma_k = \alpha_{k+1} + 1 + \beta_{k+1}$ where $\alpha_{k+1} = [\min \gamma_k, p_k[= [\min \gamma_k, \bar{p}_k[$ and $\beta_{k+1} =]p_k, \max \gamma_k] =]\bar{p}_k, \max \gamma_k]$.

In these cases, we stop the iteration and we obtain one of the following equalities:

$$o = \overrightarrow{\sum}_{i=1}^{k+1} (\alpha_i + 1) + \gamma_{k+1} + \overleftarrow{\sum}_{i=1}^{k+1} (1 + \beta_i),$$

if $c(\gamma_{k+1}) \neq c(o)$ or $p_k > \bar{p}_k$, or

$$o = \overrightarrow{\sum}_{i=1}^{k+1} (\alpha_i + 1) + \beta_{k+1} + \overleftarrow{\sum}_{i=1}^{k} (1 + \beta_i),$$

if $p_k = \bar{p}_k$. Let $l_0 = l$. For each i, let l_i' be the restriction of l_{i-1} to the initial segment α_i of γ_{i-1}, l_i'' be the labeling of β_i defined by $l_i''(\delta) = l_{i-1}(\alpha_i + 1 + \gamma_i + 1 + \delta)$, where δ is an initial segment of β_i, and l_i be the labeling of γ_i defined by $l_i(\delta) = l_{i-1}(\alpha_i + 1 + \delta)$, where δ is an initial segment of γ_i. For each $m \geq 1$, we obtain

(i) $l_m'(\delta) = l\left(\overrightarrow{\sum}_{i=1}^{m-1} (\alpha_i + 1) + \delta\right)$, if δ is an initial segment of α_m,

(ii) $l_m(\delta) = l\left(\overrightarrow{\sum}_{i=1}^{m} (\alpha_i + 1) + \delta\right)$, if δ is an initial segment of γ_m,

(iii) $l_m''(\delta) = l\left(\overrightarrow{\sum}_{i=1}^{m} (\alpha_i + 1) + \gamma_m + 1 + \delta\right)$, if δ is an initial segment of β_m.

This defines the *iterated central basic partition of* $\mathbf{o}$.

In case $p_i^o < \bar{p}_i^o$ and $c(\gamma_i) = c(o)$, for all i, we iterate indefinitely the partition defined above and we obtain $o = \overrightarrow{\sum}_{i \geq 1} (\alpha_i + 1) + \overleftarrow{\sum}_{i \geq 1} (1 + \beta_i)$ as justified by the following lemma.

Lemma 4.17. *Let* $\mathbf{o} = (o, l) \in \mathbf{rLO}^*(A)$*. If* $\mathbf{o}$ *has an infinite iterated central basic partition, then* $o = \overrightarrow{\sum}_{i\geq 1}(\alpha_i + 1) + \overleftarrow{\sum}_{i\geq 1}(1 + \beta_i)$.

Proof. Let $P = \{p_i^o : i \geq 1\}$ and $\bar{P} = \{\bar{p}_i^o : i \geq 1\}$. These sets are infinite, P is bounded above and it does not have a maximum element and $\bar{P}$ is bounded below and it does not have a minimum element. We consider the subsets of $\mathcal{D}_o$ that follow:

$$D = \{(\mathbf{o_1}, \mathbf{o_2}) \in \mathcal{D}_o \mid \exists p_i^o \in P, p_i^o = \max o_1\},$$

$$\bar{D} = \{(\mathbf{o_1}, \mathbf{o_2}) \in \mathcal{D}_o \mid \exists \bar{p}_i^o \in \bar{P}, \bar{p}_i^o = \max o_1\}.$$

Let $(\mathbf{o_1}, \mathbf{o_2}) = \sup D$ and $(\mathbf{o_1'}, \mathbf{o_2'}) = \inf \bar{D}$, which exist, since $\mathcal{D}_o$ is a complete ordering, D is bounded above and $\bar{D}$ is bounded below. Note that these Dedekind cuts are gaps in $\mathbf{o}$, since D does not have a maximum and $\bar{D}$ does not have a minimum. Moreover, by definition of iterated central basic partition of an ordering $\mathbf{o}$ and by definition of cofinal set of a gap in $\mathbf{o}$, we have $c((\mathbf{o_1}, \mathbf{o_2})) = c((\mathbf{o_1'}, \mathbf{o_2'})) = c(\mathbf{o})$. If these gaps are distinct, then, by definition of reduced $*$-labeled linear ordering, there exists a Dedekind cut between them whose label does not belong to $c(\mathbf{o})$, which is a contradiction. It follows that $(\mathbf{o_1}, \mathbf{o_2}) = (\mathbf{o_1'}, \mathbf{o_2'})$ and, therefore, $o = \overrightarrow{\sum}_{i\geq 1}(\alpha_i + 1) + \overleftarrow{\sum}_{i\geq 1}(1 + \beta_i)$. □

Let $\nu : \mathbf{rLO}^*(A) \to T_1(A)$ be the mapping defined as follows. If $o = 0$, then $\nu(o, l)$ is the tree which consists of a unique degenerate node. If $o \neq 0$, then we consider the iterated central basic partition of (o, l), $o = \overrightarrow{\sum}_{i\geq 1}(\alpha_i+1)+\overleftarrow{\sum}_{i\geq 1}(1+\beta_i)$. Note that $c(\alpha_i, l_i')$ and $c(\beta_i, l_i'')$ are strictly contained in $c(o, l)$, for all i. We defined the tree $\nu(o, l)$ by induction on the content of (o, l): to each element of the sum $o = \overrightarrow{\sum}_{i\geq 1}(\alpha_i + 1) + \overleftarrow{\sum}_{i\geq 1}(1 + \beta_i)$ corresponds a son of the root. The vertex which corresponds to an element α_i or β_i, for all i (and also to γ_k, if it exists) is a node whose content is the image of the labelings l_i' and l_i'', respectively (if γ_k exists, then the content of the corresponding vertex is the image of l_k). We note that, if one of those orderings is 0, the corresponding vertex is a degenerate node. To the rest of the elements of the sum we associate a leaf labeled, respectively, $l(\alpha_1 + 1)$, $l(\alpha_1 + 1 + \alpha_2 + 1)$, $\ldots$, $l(\alpha_1 + 1 + \alpha_2 + 1 + \cdots \cdots + 1 + \beta_3 + 1)$, $l(\alpha_1 + 1 + \alpha_2 + 1 + \cdots \cdots + 1 + \beta_3 + 1 + \beta_2 + 1)$. We attach to the vertices which do not correspond to a degenerate node the subtrees $\nu(\alpha_1, l_1')$, $\nu(\alpha_2, l_2')$, $\ldots$, $\ldots$, $\nu(\beta_2, l_2'')$, $\nu(\beta_1, l_1'')$. We notice that the trees constructed in this manner are, effectively, in $T_1(A)$. In fact, we can verified immediately the properties (1)-(5) from the definition of tree in $T_1(A)$ in the construction made. The properties (6)-(8) follow by the definition of iterated central basic factorization of (o, l) and the fact that the orderings involved are reduced.

Theorem 4.18. *The mapping* $\mu : T_1(A) \to \mathbf{rLO}^*(A)$ *is a bijection.*

Proof. Let $t \in T(A)$. If $\mathfrak{h}(t) = 0$, then $\nu(\mu(t)) = t$ by definition. Suppose that t has non-zero height and let $\mu(t) = (o, l)$. By Lemma 4.15 and by definition of iterated central basic partition of a reduced $*$-labeled linear ordering, we have that $o = \overrightarrow{\sum}_{n\geq 1}(o_n + 1) + \overleftarrow{\sum}_{n\geq 1}(1 + o_n')$ is exactly the iterated central basic partition of (o, l) and, therefore, $\nu(\mu(t)) = t$. Now, let $(o, l) \in \mathbf{rLO}^*(A)$. The case $o = 0$ follows by definition. Suppose that $o \neq 0$ and let $o = \overrightarrow{\sum}_{i\geq 1}(\alpha_i + 1) + \overleftarrow{\sum}_{i\geq 1}(1 + \beta_i)$ be the iterated central basic partition of $o = (o, l)$. For each i, let $l_{\alpha_i} : \alpha_i \to A$ be defined by $l_{\alpha_i}(\delta) = l_i'(\delta)$, where δ is an initial segment of α_i, and $l_{\beta_i} : \beta_i \to A$

is defined by $l_{\beta_i}(\delta) = l''_i(\delta)$, where δ is an initial segment of β_i. Consider the tree whose leaf sons of the root are labeled $l(\alpha_1 + 1)$, $l(\alpha_1 + 1 + \alpha_2 + 1)$, $\dots$, $l(\alpha_1+1+\alpha_2+1+\cdots\cdots+1+\beta_3+1)$, $l(\alpha_1+1+\alpha_2+1+\cdots\cdots+1+\beta_3+1+\beta_2+1)$ and whose subtrees attached to the nodes of the progeny of the root are $\nu(\alpha_1, l_{\alpha_1})$, $\nu(\alpha_2, l_{\alpha_2})$, $\dots$, $\nu(\beta_2, l_{\beta_2})$, $\nu(\beta_1, l_{\beta_1})$. It follows, by Lemma 4.15, that $\mu(\nu(o, l)) = (o, l)$. □

We finish by showing the relation between the representations of implicit operations by finite-height trees and by labeled orderings. It will be useful to relate also these representations with the representation by *quasi-ternary* trees. In fact, we can construct a bijection $\xi : T_2(A) \to T_1(A)$ recursively as follows. Let $t \in T_2(A)$ be such that $\rho(t) = \rho(t_0) \cdot l_{(\varepsilon,0)} \cdot \rho(t_{10}) \cdot l_{(1,0)} \cdots \cdots l_{(1,2)} \cdot \rho(t_{12}) \cdot l_{(\varepsilon,2)} \cdot \rho(t_2)$, where $\rho : T_2(A) \to \overline{\Omega}_A\mathsf{DA}$ is the bijection defined in 4.1.2. Let i be maximum for $c(v_{1^i}) = c(\rho(t))$, where v_{1^i} is the root of the subtree t_{1^i}. Then the progeny of the root of $\xi(t)$ consists of the leaves labeled by $l_{(\epsilon,0)}$, $l_{(1,0)}$, $\dots$, $l_{(1^i,0)}$, $l_{(1^i,2)}$, $\dots$, $l_{(1,2)}$, $l_{(\epsilon,0)}$ and the trees $\rho(t_0)$, $\rho(t_{10})$, $\dots$, $\rho(t_{1^i0})$, $\rho(t_{1^{i+1}})$, $\rho(t_{1^i2})$, $\dots$, $\rho(t_{12})$, $\rho(t_2)$ attached to the nodes. Proceeding recursively, and since A is finite, we obtain the tree $\xi(t) \in T_1(A)$. We leave the details to the reader.

Acknowledgments

This work is part of the author's doctoral thesis, written under the supervision of Prof. Jorge Almeida, from whose advice the author has greatly benefited. This work was supported by the *Fundação para a Ciência e a Tecnologia (FCT)* through the *PhD grant* SFRH/BD/19720/2004, through the *Centro de Matemática da Universidade do Porto (CMUP)* and also through the project PTDC/MAT/65481/2006, which is partly funded by the European Community Fund FEDER.

References

[1] J. Almeida, *Implicit operations on finite* J*-trivial semigroups and a conjecture of I. Simon*, J. Pure and Applied Algebra **69** (1990) 205–218.
[2] ———, *Finite Semigroups and Universal Algebra*, World Scientific, Singapore, 1994. English translation.
[3] ———, *A syntactical proof of locality of* DA, Int. J. Algebra and Comp. **6** (1996) 165–177.
[4] ———, *Finite semigroups: an introduction to a unified theory of pseudovarieties*, in Semigroups, Algorithms, Automata and Languages, J.-E. P. G. M. S. Gomes and P. V. Silva, eds., Singapore, 2002, World Scientific, 3–64.
[5] J. Almeida, J. C. Costa, and Zeitoun, *Iterated periodicity over finite aperiodic semigroups*, Tech. Rep. CMUP 2009-03, Univ. of Porto, 2009.
[6] J. Almeida, J. C. Costa, and M. Zeitoun, *Some structural properties of the free profinite aperiodic semigroup*, in Proc. of the conference "AutoMathA: from Mathematics to Applications", Univ. Liège, June 2009, 2009.
[7] J. Almeida and P. Weil, *Free profinite* R*-trivial monoids*, Int. J. Algebra and Comput. **7** (1997) 625–671.
[8] J. Almeida and M. Zeitoun, *An automata-theoretic approach to the word problem for ω-terms over* R, Theor. Comp. Sci. **370** (2007) 131–169.
[9] A. Azevedo, *Operations preserving homomorphisms on the class of finite semigroups* DS, in Actas II Encontro de Algebristas Portugueses, J. Almeida, ed., Porto, 1987, Universidade do Porto, 33–43.
[10] S. Eilenberg, *Automata, Languages and Machines*, vol. B, Academic Press, New York, 1976.
[11] J. Reiterman, *The Birkhoff theorem for finite algebras*, Algebra Universalis **14** (1982) 1–10.
[12] J. G. Rosenstein, *Linear Orderings*, Academic Press, New York, 1982.

[13] M. P. Schützenberger, *Sur le produit de concaténation non ambigu*, Semigroup Forum **13** (1976) 47–75.
[14] P. Tesson and D. Thérien, *Diamonds are forever: the variety* DA, in Semigroups, Algorithms, Automata and Languages, World Scientific, 2002, 475–499.
[15] P. Trotter and P. Weil, *The lattice of pseudovarieties of idempotent semigroups and a non-regular analogue*, Algebra Univers. **37** (1997) 491–526.

Instituto Superior de Engenharia do Porto/LEMA and Centro de Matemática da Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal
E-mail address: aim@isep.ipp.pt/amoura@fc.up.pt

Paper 2

THE WORD PROBLEM FOR ω-TERMS OVER DA

A. MOURA

ABSTRACT. In this paper, we solve the word problem for ω-terms over DA. We extend to DA the ideas used by Almeida and Zeitoun to solve the analogous problem for the pseudovariety R applying also a representation by automata of implicit operations on DA, which was recently obtained by the author. Considering certain types of factors of an implicit operation on DA, we can prove that a pseudoword on DA is an ω-term if and only if the associated minimal DA-automaton is finite. Finally, we complete the result by effectively computing in polynomial time the minimal DA-automaton associated to an ω-term.

1. INTRODUCTION

The pseudovariety DA, the class of finite monoids whose regular $\mathcal{D}$-classes are aperiodic semigroups, has been the subject of recent studies due to its various applications. It is known that languages whose syntactic monoids lie in DA have important algebraic, combinatorial, automata-theoretical and logical characterizations that enable us to solve problems in computational and complexity theory (see Tesson and Thérien [13]).

On the other hand, word problems have long played an important role in various branches of Mathematics. In this paper, we solve the word problem for ω-terms over DA, which consists of deciding if two ω-terms are equal over all elements of this pseudovariety. Almeida and Zeitoun [5, 4] solved the analogous problem for the pseudovariety R. Based on this work, we characterize ω-terms over DA by the finiteness of certain types of sets of factors and by the finiteness of the associated minimal DA-automaton. We also construct in polynomial time this minimal DA-automaton.

In [11], we exhibited three representations of implicit operations over DA: by means of labeled trees of finite height, by means of *quasi-ternary* labeled trees, and by means of labeled linear orderings. The paper has also an improvement of the representation by *quasi-ternary* labeled trees, which may be infinite, consisting of *wrapping* the DA-tree of an implicit operation. We obtain a representation by means of DA-automata and we prove here that an ω-term has a finite representation by the minimal DA-automaton. Since this paper depends on several definitions and results from [11], the reader should refer to that paper as needed.

The paper is organized as follows. In Section 2, we introduce some notions and notation about implicit signatures, concluding the corresponding section from [11]. We also recall the notion of central basic factorization of an implicit operation on DA and the representation of implicit operations on DA by automata. Based on Almeida and Zeitoun [5], we construct in Section 3 certain types of factors of

2000 *Mathematics Subject Classification.* 20M05; 20M07; 20M35.

Key words and phrases. Finite monoid; pseudovariety; word problem; pseudoword; omega-term; aperiodic; regular D-class.

an implicit operation. This allows us to characterize an ω-term on DA by the finiteness of these sets of factors and by the finiteness of the associated minimal DA-automaton, which is done in Section 4. Finally, in Section 5 we exhibit an algorithm to compute a finite DA-automaton associated to an ω-term and we prove that the minimal DA-automaton associated to an ω-term can be constructed in polynomial time.

2. Preliminaries

We complete the introduction of notions and notation given in the corresponding section of [11]. For further information on the basic background see, for instance, [1, 3].

In this paper, $\overline{\Omega}_A\mathsf{V}$ denotes the *free pro-*V *monoid on* A. The *natural interpretation* of $u \in \overline{\Omega}_A\mathsf{V}$ in a pro-V monoid S is the mapping $u_S : S^A \to S$ which associates to each function $\mu : A \to S$ the element $\hat{\mu}(u) \in S$. For $u \in \overline{\Omega}_A\mathsf{V}$, the sequence $(u^{n!-1})_n$ converges and we denote the limit by $u^{\omega-1}$. Similarly, u^ω denotes the limit of the sequence $(u^{n!})_n$, which is the unique idempotent in the closed subsemigroup generated by u. The elements of $\overline{\Omega}_A\mathsf{V}$ are called *implicit operations over* V or *pseudowords over* V. Usually, the first name is used when these elements are viewed, via their natural interpretation, as operations on finite semigroups, and the second name is used when the elements are viewed as combinatorial entities generalizing finite words. Recall that, if $\iota : A \to \overline{\Omega}_A\mathsf{V}$ is the natural generating function, then the submonoid generated by $\iota(A)$ is a dense submonoid of $\overline{\Omega}_A\mathsf{V}$.

An *implicit signature* is a set of implicit operations containing the monoid multiplication, $\_\cdot\_$. The *canonical signature* $\kappa = \{\_\cdot\_, \_^{\omega-1}\}$ consists of the monoid multiplication and the unary $(\omega-1)$-power. A κ*-term on the set* A is an element of the unary semigroup T_A^κ freely generated by A , and $\Omega_A^\kappa\mathsf{V}$ is the κ-submonoid of the pro-V monoid freely generated by A, whose elements are called κ*-words* or κ*-terms over* V. A κ*-identity over* V is an equality $u = v$, with u and v κ-words over V. The κ*-word problem for* V consists in deciding if two κ-terms of T_A^κ have the same image under the natural homomorphism into the free pro-V monoid, $\iota : T_A^\kappa \to \overline{\Omega}_A\mathsf{V}$. The signature $\omega = \{\_\cdot\_, \_^\omega\}$ is also of interest. Since, in an aperiodic monoid, any κ-term coincides with the ω-term obtained by replacing all $(\omega-1)$-powers by ω-powers, we can work and formulate the results in terms of the signature ω, which we do from hereon.

In Section 5, we adopt the simplified notation of McCammond [10] using the curved parentheses to represent the ω-power, and so, the ω-terms are seen as words on the extended alphabet $A \cup \{(,)\}$.

Given $w \in \overline{\Omega}_A\mathsf{DA}\backslash\{1\}$, we consider the central basic factorization of w, under the conditions described by Almeida [2], as the tuple $(\alpha, a, \gamma, b, \beta) \in \overline{\Omega}_A\mathsf{DA} \times A \times \overline{\Omega}_A\mathsf{DA} \times A \times \overline{\Omega}_A\mathsf{DA}$ or as the triple $(\alpha, a, \beta) \in \overline{\Omega}_A\mathsf{DA} \times A \times \overline{\Omega}_A\mathsf{DA}$ satisfying one of the following conditions:

(i) **standard form:** $w = \alpha a \gamma b \beta$ with $a, b \in A$, $\alpha, \beta, \gamma \in \overline{\Omega}_A\mathsf{DA}$, $a \notin c(\alpha)$, $b \notin c(\beta)$ and $c(\alpha a) = c(b\beta) = c(w)$;

(ii) **overlapped form:** $w = \alpha b \gamma a \beta$ with $a, b \in A$, $\alpha, \beta, \gamma \in \overline{\Omega}_A\mathsf{DA}$, $a \notin c(\alpha b \gamma)$, $b \notin c(\gamma a \beta)$ and $c(\alpha b \gamma a) = c(b \gamma a \beta) = c(w)$;

(iii) **degenerate form:** $w = \alpha a \beta$ with $a \in A$, $\alpha, \beta \in \overline{\Omega}_A\mathsf{DA}$, $a \notin c(\alpha)$, $a \notin c(\beta)$ and $c(\alpha a) = c(a\beta) = c(w)$.

Almeida proved that this factorization exists and is unique and we denote it by $\mathsf{CBF}(w)$. We iterate this factorization by applying it to the central factor γ until it becomes 1 or the central basic factorization is of the degenerate form. We denote this iterated central basic factorization (called of type 2 in [11]) by $\mathsf{I_2CBF}(w)$ and it has one of the following forms: $\mathsf{I_2CBF}(w) = \alpha_1 a_1 \cdots \alpha_n a_n b_n \beta_n \cdots b_1 \beta_1$, $\mathsf{I_2CBF}(w) = \alpha_1 a_1 \cdots \alpha_n a_n \beta_n \cdots b_1 \beta_1$ or $\mathsf{I_2CBF}(w) = \alpha_1 a_1 \cdots \ \cdots b_1 \beta_1$.

To solve the word problem over DA, we use a result from [11] which states that two pseudowords have the same DA-*quasi-ternary* tree if and only if they are equal over DA. As these DA-trees may be infinite, and, as such, they may not be calculated in full form, we use the improvement of this representation which consists of representing the implicit operations on DA by means of DA-automata.

Briefly, the tree $t(w) \in T_2(A)$ which represents the pseudoword $w \in \overline{\Omega}_A\mathsf{DA}$ is constructed recursively as follows: it has a root corresponding to the pseudoword w and, assuming that $\mathsf{CBF}(w) = \alpha a \gamma b \beta$, the root is labeled by the pair (a, b) and it has three sons, corresponding to the pseudowords α, γ and β, with edges labeled 0, 1 and 2, respectively. If $\mathsf{CBF}(v) = \alpha a \beta$ is degenerate, for some vertex corresponding to a pseudoword v, then this vertex is labeled a and it has only two sons with edges labeled 0 and 2, respectively. Any tree of $T_2(A)$ is a DA-automaton, $t(w) = (V, \rightarrow, q, F, \lambda)$, where q is the root, F is the set of vertices corresponding to the empty word and λ is the state labeling function. The *wrapped* DA-*automaton* of w, $\mathcal{A}(w)$, is the automaton obtained from $t(w)$ by identifying states corresponding to the same pseudoword. Moving the label of each state and adding it to the labels of the edges starting in such state, we obtain an automaton that recognizes the language associated to w, $\mathcal{L}(w)$ (see [11]), and which is minimal for the condition $\mathcal{L}(\mathcal{A}) = \mathcal{L}(w)$. We end this section with the following powerful result from [11]:

Proposition 2.1. *Let* $v, w \in \overline{\Omega}_A\mathsf{DA}$. *Then* $\mathsf{DA} \models v = w$ *if and only if* $\mathcal{L}(v) = \mathcal{L}(w)$.

3. Factors of a pseudoword over DA

In this paper, we prove that the word problem can be effectively solved when we work with ω-terms over DA. We start by considering certain types of factors of a pseudoword $w \in \overline{\Omega}_A\mathsf{DA}$.

Let $w \in \overline{\Omega}_A\mathsf{DA}$. We define certain sets of factors of w: $\mathcal{F}(w)$, which consists of the so called DA-*factors* of w; $\mathcal{R}(w)$, consisting of the *relative remainders* of w; and $\mathcal{S}(w)$, the set of the *absolute remainders* of w.

We define $f_\delta(w)$, $l_{(\delta,0)}(w)$ and $l_{(\delta,2)}(w)$ by induction on the length of $\delta \in \{0, 1, 2\}^*$ as follows:

$$\begin{array}{rcl} f_\varepsilon(w) & = & w \\ (f_{\delta 0}(w), l_{(\delta,0)}(w), f_{\delta 1}(w), l_{(\delta,2)}(w), f_{\delta 2}(w)) & \stackrel{def}{=} & \mathsf{CBF}(f_\delta(w)) \\ \text{or } (f_{\delta 0}(w), l_{(\delta,0)}(w), f_{\delta 2}(w)) & \stackrel{def}{=} & \mathsf{CBF}(f_\delta(w)) \end{array}$$

depending on whether the central basic factorization of $f_\delta(w)$ is of the standard or of the overlapped form, or if it is of the degenerate form. The set of DA-factors of w is

$$\mathcal{F}(w) = \{f_\delta(w) \mid \delta \in \{0, 1, 2\}^* \text{ and } f_\delta(w) \text{ is defined}\} \subseteq \overline{\Omega}_A\mathsf{DA}.$$

It consists of the set of images under $\rho : T_2(A) \mapsto \overline{\Omega}_A\mathsf{DA}$ of the subtrees $t(w)$, which correspond to some factor of the form α_δ, β_δ or γ_δ of the iterated factorization of some factor of the iterated central basic factorization of type 2 of w.

The set of relative remainders of w is the set $\mathcal{R}(w)$ of elements of $\mathcal{F}(w)$, which consists of the images under ρ of the subtrees attached to a vertex which is a son from a central branch of a given vertex. These subtrees are the trees corresponding to the factors γ_δ of some iterated factorization of a factor of the iterated central basic factorization of type 2 of w. Formally, we write

$$\mathcal{R}(w) = \{f_\delta(w) \mid \delta \in \{0,1,2\}^*1 \text{ and } f_\delta(w) \text{ is defined}\} = f_1(\mathcal{F}(w)).$$

Let $u, v \in \overline{\Omega}_A\mathsf{DA}$ be such that u is a prefix of v. We use the notation $u^{-1}v$ to represent any suffix of v such that $v = u \cdot u^{-1}v$ in $\overline{\Omega}_A\mathsf{DA}$. Similarly, if u is a suffix of v, we use vu^{-1} to denote any prefix of v such that $v = vu^{-1} \cdot u$ in $\overline{\Omega}_A\mathsf{DA}$. We define the set of absolute remainders of w, $\mathcal{S}(w)$, to be the smallest subset containing w and satisfying to the following conditions:

(i) $u \in \mathcal{S}(w) \Rightarrow f_0(u) \in \mathcal{S}(w)$;
(ii) $u \in \mathcal{S}(w) \Rightarrow f_2(u) \in \mathcal{S}(w)$;
(iii) $u, v \in \mathcal{S}(w)$, $a \in A$ and $\mathbf{o}(ua)$ is an initial segment of $\mathbf{o}(v)$ implies that $(ua)^{-1}v \subseteq \mathcal{S}(w)$;
(iv) $u, v \in \mathcal{S}(w)$, $a \in A$ and $\mathbf{o}(au)$ is a final segment of $\mathbf{o}(v)$ implies that $v(au)^{-1} \subseteq \mathcal{S}(w)$.

Recall that $\mathbf{o}(u)$ is the reduced $*$-labeled linear ordering representing $u \in \overline{\Omega}_A\mathsf{DA}$, notation introduced in [11]. See Rosenstein [12] for the basics on linear orderings. Let us see the relation that exists between the elements of $\mathcal{S}(w)$ and the closed intervals of $\mathbf{o}(w)$, starting by observing some auxiliary results.

Lemma 3.1. *Given $w \in \overline{\Omega}_A\mathsf{DA}$, we have $\mathcal{F}(w) \subseteq \mathcal{S}(w)$.*

Proof. We obviously have $w \in \mathcal{S}(w)$. By conditions (i) and (ii) we have, respectively, the elements $f_0(w)$ and $f_2(w)$ in $S(w)$. By [11, Lemma 4.16], it follows that, for each $a \in A$, there exist the smallest position of $\mathbf{o}(w)$ labeled a, $p_a^{o(w)}$, and the largest position of $\mathbf{o}(w)$ labeled a, $\bar{p}_a^{o(w)}$. Therefore, and since A is finite, there exist $p^{o(w)} = \max\{p_a^{o(w)} | a \in A\}$ and $\bar{p}^{o(w)} = \min\{\bar{p}_a^{o(w)} | a \in A\}$. By definition of $f_0(w)$ and $f_2(w)$, it follows that $\mathbf{o}(f_0(w)l(p^{o(w)}))$ is an initial segment of $\mathbf{o}(w)$ and $\mathbf{o}(l(\bar{p}^{o(w)})f_2(w))$ is a final segment of $\mathbf{o}(w)$. By conditions (iii) and (iv), $f_1(w) \in S(w)$. Proceeding inductively on the factors $f_0(w)$, $f_1(w)$ and $f_2(w)$, we deduce that all the elements of $\mathcal{F}(w)$ are in $\mathcal{S}(w)$. $\square$

Lemma 3.2. *Let $w \in \overline{\Omega}_A\mathsf{DA}$. For each position p in $\mathbf{o}(w)$, there exists a closed interval $\mathbf{o}' \subseteq \mathbf{o}(w)$ such that:*

(i) *$\mathbf{o}' \simeq \mathbf{o}(f_\delta(w))$ with $f_\delta(w) \in \mathcal{F}(w)$;*
(ii) *$p = p^{o'}$ or $p = \bar{p}^{o'}$.*

Proof. We proceed by induction on the content of w, $c(w)$. If $|c(w)| = 1$, suppose that $c(w) = \{a\}$, then $w = a^n$, with n finite, or $w = a^\omega$. If $w = a^n$, with n finite, then $\mathbf{o}(w) = \mathbf{n}$. In this case, $\mathbf{o}' = \mathbf{o}(f_{1^{p-1}}(w))$, if $p \leq \lceil n/2 \rceil$, or $\mathbf{o}' = \mathbf{o}(f_{1^{n-p}}(w))$, if $p > \lceil n/2 \rceil$ satisfies the desired conditions (it is enough to observe that in the iterated central basic factorization of w the factors α_i and β_i are all empty and each letter a at a given position is a distinguished label in a position $p^{o'}$ or $\bar{p}^{o'}$ of some iteration). In the case where $w = a^\omega$, we have $\mathbf{o}(w) = \omega + \omega^*$. If p is a position in ω, then we set $\mathbf{o}' = \mathbf{o}(f_{1^{p-1}}(w))$. Otherwise, we set $\mathbf{o}' = \mathbf{o}(f_{1^{q-1}}(w))$, where q is the positive integer corresponding to the position p in $\omega + \omega^*$ when we count from right to left. In any case, the chosen orderings satisfy the conditions (i) and (ii).

Now, suppose that $|c(w)| > 1$. We consider the iterated central basic factorization of type 2 of w, $\mathsf{I_2CBF}(w)$. Then, by [11, Theorem 4.5] and by the analogue version of [11, Theorem 4.18] for $T_2(A)$, we have $\mathbf{o}(w) = \mathbf{o_1} + \mathbf{1} + \mathbf{o_2} + \mathbf{1} + \cdots \cdots + \mathbf{1} + \bar{\mathbf{o}}_\mathbf{2} + \mathbf{1} + \bar{\mathbf{o}}_\mathbf{1}$, for some orderings $\mathbf{o_i}$ and $\bar{\mathbf{o}}_\mathbf{i}$. If $p \in \mathbf{o}(w)$ corresponds to any position labeled a_i or b_i of $\mathsf{I_2CBF}(w)$, then $\mathbf{o}' = \mathbf{o}(f_{1^{i-1}}(w))$ satisfies the desired conditions. Otherwise, p is a position in $\mathbf{o_i}$ or $\bar{\mathbf{o}}_\mathbf{i}$, for some i. Since the content of the pseudoword represented by this ordering is strictly contained in $c(w)$, the result follows by induction. $\square$

Given $f_\delta(w) \in \mathcal{F}(w)$, with $\delta \in \{0,1,2\}^*$, we define the *depth* of $f_\delta(w)$, $\mathfrak{d}(f_\delta(w))$, as the length of the word $\delta \in \{0,1,2\}^*$.

Lemma 3.3. *Let $w \in \overline{\Omega}_A\mathsf{DA}$. Given $f_\delta(w) \in \mathcal{F}(w)$, with $\delta \in \{0,1,2\}^*$, there exist $k \geq 0$, $f_{\delta_1}, f_{\delta_2}, \ldots, f_{\delta_k} \in \mathcal{F}(w)$ and $a_{\delta_1}, a_{\delta_2}, \ldots, a_{\delta_k} \in A$ such that $\mathbf{o}(f_{\delta_1}a_{\delta_1}f_{\delta_2}a_{\delta_2}\cdots f_{\delta_k}a_{\delta_k}f_\delta(w))$ is an initial segment of $\mathbf{o}(w)$.*

Proof. We proceed by induction on $\mathfrak{d}(f_\delta(w))$. The case where $\mathfrak{d}(f_\delta(w)) = 0$, i.e., $f_\delta(w) = f_\varepsilon(w)$, is trivial since $f_\varepsilon(w) = w$. Let $f_\delta(w) \in \mathcal{F}(w)$, with $\delta \in \{0,1,2\}^*$, be such that $\mathfrak{d}(f_\delta(w)) = 1$. Three cases can occur: $f_\delta(w) = f_0(w)$, $f_\delta(w) = f_1(w)$ or $f_\delta(w) = f_2(w)$. It follows, respectively, that $\mathbf{o}(f_0(w))$, $\mathbf{o}(f_0 l(p^{o(w)})f_1(w))$ and $\mathbf{o}(f_0 l(p^{o(w)})f_1 l(\bar{p}^{o(w)})f_2(w))$ are initial segments of $\mathbf{o}(w)$. Now, suppose that $\mathfrak{d}(f_\delta(w)) = n > 1$. Let η be the prefix of δ with length $|\delta| - 1$. By induction hypothesis, there exist $f_{\eta_1}, f_{\eta_2}, \ldots, f_{\eta_m} \in \mathcal{F}(w)$ and $a_{\eta_1}, a_{\eta_2}, \ldots, a_{\eta_m} \in A$ such that $\mathbf{o}(f_{\eta_1}a_{\eta_1}f_{\eta_2}a_{\eta_2}\cdots f_{\eta_m}a_{\eta_m}f_\eta(w))$ is an initial segment of $\mathbf{o}(w)$. By definition of $f_\delta(w) \in \mathcal{F}(w)$, it follows that $f_\delta(w)$ is a factor of $f_\eta(w)$ if and only if η is a prefix of δ. Consider the factors $f_{\eta 0}, f_{\eta 1}, f_{\eta 2} \in \mathcal{F}(w)$. Note that one of them is the factor $f_\delta(w)$. We have, respectively, $\mathbf{o}(f_{\eta_1}a_{\eta_1}f_{\eta_2}a_{\eta_2}\cdots f_{\eta_m}a_{\eta_m}f_{\eta 0}(w))$, $\mathbf{o}(f_{\eta_1}a_{\eta_1}f_{\eta_2}a_{\eta_2}\cdots f_{\eta_m}a_{\eta_m}f_{\eta 0}l(p^{o(\eta)})f_{\eta 1}(w))$ and $\mathbf{o}(f_{\eta_1}a_{\eta_1}f_{\eta_2}a_{\eta_2}\cdots f_{\eta_m}a_{\eta_m}f_{\eta 0} \cdot l(p^{o(\eta)})f_{\eta 1}l(\bar{p}^{o(\eta)})f_{\eta 2}(w))$ as initial segments of $\mathbf{o}(w)$, in the cases where $f_\delta(w) = f_{\eta 0}(w)$, $f_\delta(w) = f_{\eta 1}(w)$ and $f_\delta(w) = f_{\eta 2}(w)$, respectively. $\square$

Lemma 3.4. *Let $w \in \overline{\Omega}_A\mathsf{DA}$. We have:*

(1) *$u \in \mathcal{S}(w) \Rightarrow \exists\, p, q \in \mathbf{o}(w) : \mathbf{o}(u) \simeq [p, q]$;*
(2) *$p, q \in \mathbf{o}(w), p < q \Rightarrow \exists\, u \in \mathcal{S}(w) : \mathbf{o}(u) \simeq [p, q]$.*

Proof. 1. By definition of $f_0(w)$, $f_1(w)$ and $f_2(w)$ and also by definition of $p^{o(w)}$ and $\bar{p}^{o(w)}$, it follows that $f_0(w) \simeq [\min o(w), p^{o(w)}[$, $f_1(w) \simeq]p^{o(w)}, \bar{p}^{o(w)}[$ and $f_2(w) \simeq]\bar{p}^{o(w)}, \max o(w)]$. Note that the predecessors and the successors of $p^{o(w)}$ and $\bar{p}^{o(w)}$ exist in any $*$-labeled linear ordering. Applying [11, Lemma 4.16] to each interval isomorphic to the elements $f_0(w)$, $f_1(w)$ and $f_2(w)$, respectively, and proceeding inductively, we deduce that all elements of $\mathcal{F}(w)$ are isomorphic to closed intervals of $\mathbf{o}(w)$. Let $u \in \mathcal{S}(w)$ and $a \in A$ be such that $\mathbf{o}(ua)$ is an initial segment of $\mathbf{o}(w)$. Then $\mathbf{o}((ua)^{-1}w)$ is a reduced $*$-labeled linear ordering, by [11, Lemma 4.13], because it is a closed interval on $\mathbf{o}(w)$. Hence there exist $p, q \in \mathbf{o}(w)$ such that $\mathbf{o}((ua)^{-1}w) \simeq [p, q]$ (in this case $q = \max o(w)$). We obtain a similar result using the condition (iv). Proceeding inductively, we conclude that all elements of $\mathcal{S}(w)$ are isomorphic to some closed interval of $\mathbf{o}(w)$.

2. Let $p, q \in \mathbf{o}(w)$ be such that $p < q$ and consider the closed interval $[p, q]$. By [11, Lemma 4.13], $[p, q]$ is a reduced $*$-linear ordering. We want to prove

that it is isomorphic to the $*$-linear ordering corresponding to an element of $\mathcal{S}(w)$. Let $p' = \text{predecessor}(p)$ and $q' = \text{successor}(q)$. Consider the interval $[p', q']$. By Lemma 3.2, there exists $f_\delta(w) \in \mathcal{F}(w)$ such that $p' = p^{o(f_\delta(w))}$ or $p' = \bar{p}^{o(f_\delta(w))}$. If $p' = p^{o(f_\delta(w))}$, we choose the factor $f_{\delta 0}$, and if $p' = \bar{p}^{o(f_\delta(w))}$, we choose the factor $f_{\delta 0} l(p^{o(f_\delta(w))}) f_{\delta 1}$. Let $f_{\delta_1}, \ldots, f_{\delta_k} \in \mathcal{F}(w)$ and $a_{\delta_1}, \ldots, a_{\delta_k} \in A$ be such that $\mathbf{o}(f_{\delta_1} a_{\delta_1} \cdots f_{\delta_k} a_{\delta_k} f_\delta(w))$ is an initial segment of $\mathbf{o}(w)$, as we had shown in Lemma 3.3. Then either $\mathbf{o}(f_{\delta_1} a_{\delta_1} \cdots f_{\delta_k} a_{\delta_k} f_{\delta 0} l(p^{o(f_\delta(w))})(w)) \simeq [\min o(w), p']$ or $\mathbf{o}(f_{\delta_1} a_{\delta_1} \cdots f_{\delta_k} a_{\delta_k} f_{\delta 0} l(p^{o(f_\delta(w))}) f_{\delta 1} l(\bar{p}^{o(f_\delta(w))})(w)) \simeq [\min o(w), p']$, depending on the case. By condition (iii) applied either $k+1$ or $k+2$ times and, depending on the case, using the factors f_{δ_i} of this initial segment, $f_{\delta 0}$ and $f_{\delta 1}$, the letters a_{δ_i}, $l(p^{o(f_\delta(w))})$ and $l(\bar{p}^{o(f_\delta(w))})$ and the pseudoword w, we obtain a pseudoword $v \in \mathcal{S}(w)$ such that $\mathbf{o}(v) \simeq]p', \max o(w)] = [p, \max o(w)]$. We proceed similarly with q' and using condition (iv) and the pseudoword v. It follows that there exists $u \in \mathcal{S}(w)$ is such that $\mathbf{o}(u) \simeq [p, q]$. $\square$

We conclude, by Lemma 3.4, that the elements of $\mathcal{S}(w)$ correspond to the closed intervals of $\mathbf{o}(w)$. Let $u \in \mathcal{S}(w)$ and let $p, q \in \mathbf{o}(w)$ be such that $\mathbf{o}(u) \simeq [p, q]$ as we have seen in the previous lemma. Let $f_\delta(w), f_\eta(w) \in \mathcal{F}(w)$, with $\delta, \eta \in \{0,1,2\}^*$, satisfy the conditions of Lemma 3.2, respectively, to p and q. We call p and q the *borders* of u and $|\delta|$ and $|\eta|$ are, respectively, the *depth* of each border.

4. Characterizations of ω-terms over DA

We solve the word problem for ω-terms over DA. For this purpose, we present, in this section, several characterizations of an ω-term over DA. We start by observing that the factors involved in the central basic factorization of an ω-term over DA are also ω-terms over DA. As a tool to be used in inductive processes that follow, we define, inductively, the length of an ω-term by $|a| = 1$, with $a \in A$, $|uv| = |u| + |v|$ and $|u^\omega| = |u| + 1$.

Lemma 4.1. *Let $w \in \Omega^\omega_A \mathsf{DA} \backslash \{1\}$ and let $(\alpha, a, \gamma, b, \beta)$ (respectively, (α, a, β)) be the central basic factorization of w. Then α, γ and β (respectively, α and β) are also ω-terms over DA.*

Proof. We proceed by induction on $(c(w), |w|)$, where the pairs are ordered lexicographically. The case $w \in A$ is trivial. Suppose that $w = x^\omega$ with $x \in \Omega^\omega_A \mathsf{DA} \backslash \{1\}$ and that the factors involved in the central basic factorization of x, $\mathsf{CBF}(x) = (\alpha, a, \gamma, b, \beta)$ (respectively, $\mathsf{CBF}(x) = (\alpha, a, \beta)$ in the degenerate case), are ω-terms over DA. Then the central basic factorization of w is of one of the following forms: $(\alpha, a, \gamma b \beta w^2 \alpha a \gamma, b, \beta)$, in the standard case (note that $w = x x^{\omega-2} x = x(x^{\omega-1})^2 x = x(x^\omega)^2 x = x w^2 x$), $(\alpha a \gamma, b, \beta w^2 \alpha, a, \gamma b \beta)$, in the overlapped case, and $(\alpha, a, \beta w^2 \alpha, a, \beta)$, in the degenerate case. In any case, the factors involved are also ω-terms over DA.

Now, suppose that $w = xy$, where the factors involved in the central basic factorization of x and y are ω-terms over DA. Let $\mathsf{CBF}(x) = (\alpha_x, a_x, \gamma_x, b_x, \beta_x)$ or $\mathsf{CBF}(x) = (\alpha_x, a_x, \beta_x)$, and $\mathsf{CBF}(y) = (\alpha_y, a_y, \gamma_y, b_y, \beta_y)$ or $\mathsf{CBF}(y) = (\alpha_y, a_y, \beta_y)$, be the central basic factorizations of x and y, respectively, depending on the type of factorization. Several cases can occur:

(i) Suppose that $c(x) = c(y) = c(w)$. Then the central basic factorization of w is $(\alpha_x, a_x, \gamma_x b_x \beta_x \alpha_y a_y \gamma_y, b_y, \beta_y)$, or $(\alpha_x a_x \gamma_x, b_x, \beta_x \alpha_y a_y \gamma_y, b_y, \beta_y)$, or $(\alpha_x, a_x, \gamma_x b_x \beta_x \alpha_y,$

$a_y, \gamma_y b_y \beta_y)$, or $(\alpha_x a_x \gamma_x, b_x, \beta_x \alpha_y, a_y, \gamma_y b_y \beta_y)$, depending on whether the central basic factorizations of x and y are, respectively, both of the standard form, $\mathsf{CBF}(x)$ is of the standard form and $\mathsf{CBF}(y)$ is of the overlapped form, $\mathsf{CBF}(x)$ is of the overlapped form and the $\mathsf{CBF}(y)$ is of the standard form, or both of the factorizations are of the overlapped form. In the cases where at least one of the central basic factorizations of x and y is degenerate, we also have analogous central basic factorizations of w. In fact, in the case where $\mathsf{CBF}(x) = (\alpha_x, a_x, \beta_x)$, we have $\mathsf{CBF}(w) = (\alpha_x, a_x, \beta_x\alpha_y a_y \gamma_y, b_y, \beta_y)$, $\mathsf{CBF}(w) = (\alpha_x, a_x, \beta_x\alpha_y, a_y, \gamma_y b_y \beta_y)$ or $\mathsf{CBF}(w) = (\alpha_x, a_x, \beta_x\alpha_y, a_y, \beta_y)$, depending on whether the central basic factorization of y is standard, overlapped or degenerate. In any case, the factors involved are finite products of ω-terms and, therefore, they are ω-terms.

(ii) Now, we suppose that $c(x) \neq c(w)$ and $c(y) = c(w)$. We also suppose that the central basic factorization of y is of the standard form, $\mathsf{CBF}(y) = (\delta_{y_k}, a_{y_0}, \gamma_y, b_y, \beta_y)$, where $k = |c(y)| - 1$. Let $(\delta_{y_{(k-1)}}, a_{y_1}, \alpha_{y_1})$ be the left basic factorization of δ_{y_k}, as defined in [5]. Since $c(\delta_{y_{(k-1)}}) \subsetneq c(\delta_{y_k}) \subsetneq c(y)$, we repeat the process a finite number of times until we obtain the factorization $y = \delta_{y_0} a_{y_k} \cdots a_{y_1} \alpha_{y_1} a_{y_0} \gamma_y b_y \beta_y$. Remember that the factors involved in this factorization are also ω-terms, by [5, Lemma 2.2] and by induction hypothesis. Let i be maximum such that $c(w) = c(x \cdot \delta_{y_0} a_{y_k} \cdots a_{y_i} \alpha_{y_i} a_{y_{(i-1)}})$. Then we have $\mathsf{CBF}(w) = (x \cdot \delta_{y_0} a_{y_k} \cdots a_{y_i} \alpha_{y_i}, a_{y_{(i-1)}}, \alpha_{y_{(i-1)}} \cdots a_{y_0} \gamma_y, b_y, \beta_y)$, where all the factors involved are ω-terms. In the case where the central basic factorization of y is degenerate, we use the same argument and we obtain $\mathsf{CBF}(w) = (x \cdot \delta_{y_0} a_{y_k} \cdots a_{y_i} \alpha_{y_i}, a_{y_{(i-1)}}, \alpha_{y_{(i-1)}} \cdots a_{y_1} \alpha_{y_1}, a_{y_0}, \beta_y)$ or $\mathsf{CBF}(y) = (x\delta_{y_k}, a_{y_0}, \beta_y)$. Let us see the case where the central basic factorization of y is of the overlapped form, $\mathsf{CBF}(y) = (\alpha_y, a_y, \gamma_y, b_y, \beta_y)$. If $c(x\alpha_y) = c(w)$ then, by a similar argument to the one used in the previous case, we obtain $\mathsf{CBF}(w) = (x \cdot \delta_{y_0} a_{y_k} \cdots a_{y_i} \alpha_{y_i}, a_{y_{(i-1)}},$
$\alpha_{y_{(i-1)}} \cdots \alpha_{y_1}, a_y, \gamma_y b_y \beta_y)$. If $c(x\alpha_y) \neq c(w)$ and $c(x\alpha_y a_y) = c(w)$, then $\mathsf{CBF}(w) = (x\alpha_y, a_y, \gamma_y b_y \beta_y)$. In the case where $c(x\alpha_y a_y) \neq c(w)$, we use a similar argument for γ_y and we obtain $\mathsf{CBF}(w) = (x\alpha_y, a_y, \delta_{y_0} b_{y_k} \cdots b_{y_i} \gamma_{y_i}, b_{y_{(i-1)}}, \gamma_{y_{(i-1)}} \cdots \gamma_{y_1} b_y \beta_y)$. We obtain the dual result for the case where $c(y) \neq c(w)$ and $c(x) = c(w)$.

(iii) Finally, we can verify the case where $c(x) \neq c(w)$ and $c(y) \neq c(w)$ using, again, an argument similar to that given for (ii). $\square$

We say that an ω-term is *reduced* if it has no subterm of the form $r^\omega s t^\omega$, with $c(s) \subseteq c(r) = c(t)$, and no subterm of the form $(rs^\omega z^\omega t)^\omega$, with r and t pseudowords which may be empty and with $c(t) \cup c(r) \subseteq c(s) = c(z)$. Recall that, in a pro-DA monoid, $r^\omega s t^\omega = r^\omega t^\omega$, if $c(s) \subseteq c(r) = c(t)$ (see [1, Lemma 8.1.4 and Theorem 8.1.7]).

Lemma 4.2. *Let w be an ω-term which defines an idempotent in $\overline{\Omega}_A\mathsf{DA}$. Then we have one of the following conditions:*

(i) *There exist ω-terms x, y, z, t such that $\mathsf{DA} \models w = xy^\omega z^\omega t$, $c(y) = c(z) = c(w)$, $|x| + |y| + |z| + |t| < |w|$ and x and t satisfy one of the following conditions: they do not define idempotents over DA or $c(s) \subsetneq c(w)$ for both $s = x$ and $s = t$;*

(ii) *There exist ω-terms x, y, z such that $\mathsf{DA} \models w = xy^\omega z$, $c(y) = c(w)$, $|x| + |y| + |z| < |w|$ and x and z satisfy one of the following conditions: they do not define idempotents over DA or $c(s) \subsetneq c(w)$ for both $s = x$ and $s = z$.*

We also have that $xy^\omega z^\omega t$ (respectively, $xy^\omega z$) is reduced.

Proof. We begin by noting that the substitutions $r^\omega st^\omega \to r^\omega t^\omega$, if $c(s) \subseteq c(r) = c(t)$, and $(rs^\omega yz^\omega t)^\omega \to rs^\omega z^\omega t$, if $c(ryt) \subseteq c(s) = c(z)$, do not change the value of an ω-term over DA. Moreover, the length of the terms decrease when we apply these substitutions. Let v be a reduced ω-term obtained from w by applying these substitutions. Since w is idempotent over DA, v is also idempotent. Moreover, $|v| \leq |w|$. We write $v = x_1 \cdots x_r$, where each x_i is a letter or a term of the form y_i^ω. By [11, Corollary 3.17], there exists x_i such that $c(x_i) = c(v)$ and $x_i = y_i^\omega$. Suppose that there exists another factor x_j with $c(x_j) = c(v)$ and $x_j = y_j^\omega$, for some y_j. Considering the fact that v is reduced, the factors x_i and x_j must be consecutive and, therefore, $v = xy^\omega z^\omega t$, with x and t not satisfying one of the conditions $c(s) = c(v) = c(w)$ or $s = y_i^\omega$, with $s = x$ or $s = t$. Thus, either x is not an idempotent, or x is an idempotent and $c(x) \subsetneq c(v)$, and similarly for t. Now, suppose that no other x_j is such that $c(x_j) = c(v)$ and $x_j = y_j^\omega$. Then $v = x_1 \cdots x_{i-1}x_ix_{i+1} \cdots x_r = xy^\omega z$ for some x, y, z, where x and z are not idempotents or, if any of them is idempotent, then it has strictly smaller content than v. We also have $|x|+|y|+|z|+|t| < |w|$ in the first case, and $|x|+|y|+|z| < |w|$ in the second case. □

We are now ready to present some characterizations of the ω-terms over DA. The following is a sort of periodicity theorem for DA.

Theorem 4.3. *Let $w \in \overline{\Omega}_A\mathsf{DA}$. The following conditions are equivalent:*

(a) *$\mathcal{L}(w)$ is rational.*
(b) *$\mathcal{A}(w)$ is finite.*
(c) *The set $\{\rho(t(w)_v) \mid v \in V\}$ is finite, where $t(w) = \langle V, \to, q, F, \lambda\rangle$.*
(d) *$\mathcal{F}(w)$ is finite.*
(e) *$\mathcal{R}(w)$ is finite.*
(f) *$\mathcal{S}(w)$ is finite.*
(g) *$w \in \Omega_A^\omega\mathsf{DA}$.*

Proof. $(a) \Leftrightarrow (b)$: Given $\mathcal{A}(w)$, which is finite, we construct a finite automaton recognizing $\mathcal{L}(w)$, by replacing the label of each edge in $\mathcal{A}(w)$ by the pair whose first component is the label that the edge has in $\mathcal{A}(w)$ and the second component is the label of the initial vertex of the edge in $\mathcal{A}(w)$. For the direct implication we do the converse: given the minimal automaton that recognizes the language $\mathcal{L}(w)$ (and it is unique by [11, Lemma 4.10]), we construct the automaton $\mathcal{A}(w)$ whose states are labeled with the second component of the label of the edges that starts from that state.

$(b) \Leftrightarrow (c)$: Note that, by definition, there exists a bijection between the set of states in $\mathcal{A}(w)$ and the pseudowords $\rho(t(w)_v)$, with $v \in V$. Hence, the result follows.

$(c) \Leftrightarrow (d)$: Applying [11, Lemma 4.9] to $t(w)$, we have that the set of vertices, $\{\rho(t(w)_v) \mid v \in V\}$, is in bijection with $\mathcal{F}(w)$.

$(d) \Rightarrow (e)$: It is obvious, because $\mathcal{R}(w) \subseteq \mathcal{F}(w)$.

$(e) \Rightarrow (f)$: Suppose that $\mathcal{R}(w)$ is finite. To show that $\mathcal{S}(w)$ is also finite, we proceed by induction on $|A|$, where the case $|A| = 0$ is trivial. Now, suppose that $|A| \geq 1$. Let $\mathcal{S}_n(w) = \{u \in \mathcal{S}(w) \mid \text{the borders of } u \text{ have depth not exceeding } n\}$.

Then, we have

$$\begin{array}{rcl}\mathcal{S}_{n+1}(\mathcal{R}(w)) & \subseteq & \mathcal{S}_n[f_0(\mathcal{R}(w))]\cdot A\cdot f_1(\mathcal{R}(w))\cdot A\cdot \mathcal{S}_n[f_2(\mathcal{R}(w))]\\ && \cup\,\mathcal{S}_n[f_0(\mathcal{R}(w))]\cdot A\cdot \mathcal{S}_n[f_1(\mathcal{R}(w))]\\ && \cup\,\mathcal{S}_n[f_1(\mathcal{R}(w))]\cdot A\cdot \mathcal{S}_n[f_2(\mathcal{R}(w))]\\ && \cup\,\mathcal{S}_n[f_0(\mathcal{R}(w))]\cup \mathcal{S}_n[f_1(\mathcal{R}(w))]\cup \mathcal{S}_n[f_2(\mathcal{R}(w))]\\ & \subseteq & \mathcal{S}[f_0(\mathcal{R}(w))]\cdot A\cdot \mathcal{R}(w)\cdot A\cdot \mathcal{S}[f_2(\mathcal{R}(w))]\\ && \cup\,\mathcal{S}[f_0(\mathcal{R}(w))]\cdot A\cdot \mathcal{S}_n(\mathcal{R}(w))\\ && \cup\,\mathcal{S}_n(\mathcal{R}(w))\cdot A\cdot \mathcal{S}[f_2(\mathcal{R}(w))]\\ && \cup\,\mathcal{S}[f_0(\mathcal{R}(w))]\cup \mathcal{S}_n(\mathcal{R}(w))\cup \mathcal{S}[f_2(\mathcal{R}(w))].\end{array}$$

By induction on n and by definition of $\mathcal{S}(w)$, we obtain

$$\begin{array}{rcl}\mathcal{S}_{n+1}(\mathcal{R}(w)) & \subseteq & \bigcup_{i=0}^{n}(\mathcal{S}[f_0(\mathcal{R}(w))]\cdot A)^i\cdot\\ && (\mathcal{S}[f_0(\mathcal{R}(w))]\cdot A\cdot \mathcal{R}(w)\cdot A\cdot \mathcal{S}[f_2(\mathcal{R}(w))]\cup \mathcal{S}[f_0(\mathcal{R}(w))]\cdot A\cup \mathcal{R}(w))\\ && \bigcup\,(\mathcal{S}[f_0(\mathcal{R}(w))]\cdot A\cdot \mathcal{R}(w)\cdot A\cdot \mathcal{S}[f_2(\mathcal{R}(w))]\cup A\cdot \mathcal{S}[f_2(\mathcal{R}(w))]\cup \mathcal{R}(w))\cdot\\ && \bigcup_{i=0}^{n}(A\cdot \mathcal{S}[f_2(\mathcal{R}(w))])^i\\ && \bigcup\,\mathcal{S}[f_0(\mathcal{R}(w))]\cup \mathcal{R}(w)\cup \mathcal{S}[f_2(\mathcal{R}(w))]\end{array}$$

for all n and, therefore, $\mathcal{S}(\mathcal{R}(w))$ is contained in the union of these sets. We also have

$$\begin{array}{rcl}\mathcal{S}(w) & \subseteq & \{w\}\cup \mathcal{S}(f_0(w))\cdot A\cdot \mathcal{R}(w)\cdot A\cdot \mathcal{S}(f_2(w))\\ && \cup\,\mathcal{S}(f_0(w))\cdot A\cdot \mathcal{S}(\mathcal{R}(w))\\ && \cup\,\mathcal{S}(\mathcal{R}(w))\cdot A\cdot \mathcal{S}(f_2(w))\\ && \cup\,\mathcal{S}(f_0(w))\cup \mathcal{S}(\mathcal{R}(w))\cup \mathcal{S}(f_2(w)).\end{array}$$

Considering the last two inclusions, it is enough to show that the following sets are finite: $\mathcal{S}(f_0(w))$, $\mathcal{S}(f_2(w))$, $\mathcal{S}[f_0(\mathcal{R}(w))]$ and $\mathcal{S}[f_2(\mathcal{R}(w))]$. Let $u \in \{f_0(w),\ f_2(w)\}\cup f_0(\mathcal{R}(w))\cup f_2(\mathcal{R}(w))$. Since $c(f_0(v)), c(f_2(v)) \subsetneq c(v)$, for all $v \neq 1$, it follows that $c(u) \subsetneq c(w)$. Moreover, since

$$\mathcal{R}(\mathcal{F}(w)) = f_1[\mathcal{F}(\mathcal{F}(w))] = f_1(\mathcal{F}(w)) = \mathcal{R}(w),$$

we have, in particular, $\mathcal{R}(u) \subseteq \mathcal{R}(w)$ and, therefore, $\mathcal{R}(u)$ is finite. Applying the induction hypothesis to u, which has a smaller content, we conclude that $\mathcal{S}(u)$ is finite. Hence $\mathcal{S}(w)$ is finite.

$(f) \Rightarrow (g)$: Let $\mathcal{S}(w)$ be finite. We proceed by induction on $|c(w)|$ to show that w is an ω-term. If $c(w) = \{a\}$, then $w = a^n$, with n finite, or $w = a^\omega$ and, therefore, it is an ω-term. Now, suppose that $|c(w)| \geq 1$. Let $w = \overrightarrow{\prod}_{i=0}^{\llbracket w\rrbracket-1}(\alpha_i a_i)\cdot \overleftarrow{\prod}_{i=0}^{\llbracket w\rrbracket-1}(b_i\beta_i)$ be the iterated central basic factorization of w. Suppose that $\llbracket w\rrbracket$ is finite. Recall that $\llbracket w\rrbracket$ is the number of iterations until we obtain the iterated central basic factorization of w. Note that $\mathcal{S}(\alpha_i)$, $\mathcal{S}(\beta_i) \subseteq \mathcal{S}(w)$, for all i, because $\alpha_i, \beta_i \subseteq \mathcal{S}(w)$. Since, by induction hypothesis, $\mathcal{S}(w)$ is finite, then $\mathcal{S}(\alpha_i)$ and $\mathcal{S}(\beta_i)$ are also finite, for all i. Moreover, $c(\alpha_i)$, $c(\beta_i) \subsetneq c(w)$, for all i. It follows, by induction hypothesis, that α_i and β_i are ω-terms, for all i. Hence, w is an ω-term.

Now, suppose that $\llbracket w\rrbracket$ is infinite. Let $u_{l,k} = \overrightarrow{\prod}_{i=l}^{l+k-1}(\alpha_i a_i)$, $v_{l,k} = \overleftarrow{\prod}_{i=l}^{l+k-1}(b_i\beta_i)$ and $w_{l,k} = \overrightarrow{\prod}_{i\geq l}(\alpha_i a_i)\cdot \overleftarrow{\prod}_{i\geq k}(b_i\beta_i)$, with $k, l \geq 0$. We have $w = u_{0,i}\cdot \alpha_i a_i \cdot w_{i+1,i+1}\cdot b_i\beta_i\cdot v_{0,i}$. By definition, $w_{i,i} = f_{1^i}(w) \in \mathcal{S}(w)$. Let N be an integer satisfying the condition of [11, Lemma 3.10], i.e., if $i, j, k \geq N$, then $c(w_{i,i}) = c(w_{j,j})$ and $c(\alpha_k a_k) = c(b_k\beta_k)$. Since $\mathcal{S}(w)$ is finite, there exist $l \geq N$ and $k > 0$ such that $w_{l+k,l+k} = w_{l,l} = u_{l,l+k}\cdot w_{l+k,l+k}\cdot v_{l,l+k}$ and, therefore, $w_{l,l} = u_{l,l+k}^{\omega}\cdot w_{l,l}\cdot v_{l,l+k}^{\omega}$. Since $c(w_{l,l}) \subseteq c(u_{l,l+k}) = c(v_{l,l+k})$, we have, by [11, Corollary 3.7], $w_{l,l} = u_{l,l+k}^{\omega}\cdot v_{l,l+k}^{\omega}$. Hence $w = u_{0,l}w_{l,l}v_{0,l} = u_{0,l}u_{l,l+k}^{\omega}v_{l,l+k}^{\omega}v_{0,l}$ which is an ω-term.

$(g) \Rightarrow (d)$: Let $w \in \Omega_A^\omega \mathsf{DA}$. We proceed by induction on $(|c(w)|, |w|)$ where the pairs are ordered lexicographically.

If $c(w) = a$, then or $w = a^n$ is a word and $\mathcal{F}(w) = \{1, a^2, a^4, \ldots, a^n\}$ or $\mathcal{F}(w) = \{1, a, a^3, \ldots, a^n\}$, depending on whether n is even or odd, or $w = a^\omega$ and we have $\mathcal{F}(w) = \{1, a^\omega\}$. In any case, $\mathcal{F}(w)$ is finite.

If $|c(w)| > 1$, we start by showing that the set $f_{1^*}(w)$ is finite. Let $w = \overrightarrow{\prod}_{i=0}^{[\![w]\!]-1}(\alpha_i a_i) \cdot \overleftarrow{\prod}_{i=0}^{[\![w]\!]-1}(b_i \beta_i)$ be the iterated central basic factorization of w. If $\|w\|$ is finite, where $\|w\|$ denotes the largest integer n such that $c(\alpha_n a_n) = c(b_n \beta_n) = c(w)$ with $\alpha_n a_n$ and $b_n \beta_n$ disjoint (notation introduced in [11]), then we can write $w = \alpha_0 a_0 \cdots \alpha_k a_k \gamma_k b_k \beta_k \cdots b_0 \beta_0$, with $a_i, b_i \in A$, $c(\alpha_i), c(\beta_i) \subsetneq c(w)$, for all i, and $c(\gamma_k) \subsetneq c(w)$. By Lemma 4.1, these factors are also ω-terms. Since $c(\gamma_k) \subsetneq c(w)$, it follows, by induction on $|c(w)|$, that $f_{1^*}(\gamma_k)$ is finite. Since $f_{1^*}(w) = f_{1^*}(\gamma_k) \cup \{\alpha_i a_i \cdots \gamma_k \cdots b_i \beta_i \mid i \leq k\}$, it follows that $f_{1^*}(w)$ is finite.

If $\|w\|$ is infinite, then, by [11, Proposition 3.15], w is idempotent. By Lemma 4.2, we can write w in one of the following forms: $w = xy^\omega z$, with $|x| + |y| + |z| < |w|$, or $w = xy^\omega z^\omega t$, with $|x| + |y| + |z| + |t| < |w|$. Suppose that we have the first case. Since $c(y) \subseteq c(w)$ and $|y| < |w|$, by induction hypothesis applied to y, $\mathcal{F}(y)$ is finite. Since $(d) \Rightarrow (f)$, it follows that $\mathcal{S}(y)$ is also finite. Similarly, the sets $\mathcal{S}(x)$ and $\mathcal{S}(z)$ are finite. Hence we have

$$f_{1^*}(w) = f_{1^*}(xy^\omega z) \subseteq \mathcal{S}(x)y^\omega \mathcal{S}(z) \cup \mathcal{S}(x)y^\omega \mathcal{S}(y) \cup \mathcal{S}(y)y^\omega \mathcal{S}(z) \cup \mathcal{S}(y)y^\omega \mathcal{S}(y).$$

It follows that $f_{1^*}(w)$ is finite. The second case is similar.

Let $l \geq N$ and $k > 0$ be such that $f_{1^{l+k}}(w) = f_{1^l}(w)$, where N satisfies the condition of [11, Lemma 3.10]. Then the following equalities are satisfied by DA:

$$\begin{aligned}
f_{1^l}(w) &= \alpha_l a_l \cdots \alpha_{l+k-1} a_{l+k-1} f_{1^{l+k}}(w) b_{l+k-1} \beta_{l+k-1} \cdots b_l \beta_l \\
&= \alpha_l a_l \cdots \alpha_{l+k-1} a_{l+k-1} f_{1^l}(w) b_{l+k-1} \beta_{l+k-1} \cdots b_l \beta_l \\
&= (\alpha_l a_l \cdots \alpha_{l+k-1} a_{l+k-1})^\omega f_{1^l}(w) (b_{l+k-1} \beta_{l+k-1} \cdots b_l \beta_l)^\omega \\
&= (\alpha_l a_l \cdots \alpha_{l+k-1} a_{l+k-1})^\omega (b_{l+k-1} \beta_{l+k-1} \cdots b_l \beta_l)^\omega
\end{aligned}$$

where the last equality follows from [11, Corollary 3.7]. It follows that

$$w = \alpha_0 a_0 \cdots \alpha_{l-1} a_{l-1} (\alpha_l a_l \cdots \alpha_{l+k-1} a_{l+k-1})^\omega (b_{l+k-1} \beta_{l+k-1} \cdots b_l \beta_l)^\omega b_{l-1} \beta_{l-1} \cdots b_0 \beta_0.$$

Note that $f_{1^*0}(w) \subseteq W_0 = \{\alpha_0, \ldots, \alpha_{l+k-1}\}$ and $f_{1^*2}(w) \subseteq W_2 = \{\beta_0, \ldots, \beta_{l+k-1}\}$. Hence we have $\mathcal{F}(w) = f_{1^*}(w) \cup \mathcal{F}(f_{1^*0}(w)) \cup \mathcal{F}(f_{1^*2}(w)) \subseteq f_{1^*}(w) \cup \mathcal{F}(W_0) \cup \mathcal{F}(W_2)$. Since W_0 and W_2 are finite sets of ω-terms on a smaller alphabet than $c(w)$, we have, by induction hypothesis, that $\mathcal{F}(W_0)$ and $\mathcal{F}(W_2)$ are finite. The implication $(g) \Rightarrow (d)$ follows and this completes the proof of the theorem. □

5. An algorithm to compute the minimal DA-automaton

Given two ω-terms on an alphabet A, we wish to show that it is possible to decide if they coincide over all elements of DA. By Theorem 4.3, we know that, if w is an ω-term over DA, then the wrapped DA-automaton (which is minimal) that represents w is finite. Moreover, by Proposition 4.11, Lemma 4.10, and Corollary 4.8 from [11], and by definition of minimal DA-automaton, two ω-terms coincide over DA if and only if their wrapped DA-automata are isomorphic.

In this section, the aim is to construct the minimal DA-automaton of an ω-term. For that purpose, we present an algorithm which constructs a finite DA-automaton of an ω-term and, using existing tools, this automaton may be efficiently minimized.

5.1. **The main function.** Let w be an ω-term and let $\bar{w} = \mathbf{word}(w)$ be a well-parenthesized word on the alphabet $A \cup \{(,)\}$, which results from replacing the ω-powers of w by a pair of parentheses. In the automaton that we want to construct, each state represents a word $\bar{u} = \mathbf{word}(u)$ that corresponds to an ω-term u which defines a DA-factor of the ω-word defined by the initial ω-term w. The automaton has as initial state the vertex corresponding to the ω-term w. For each state $\bar{u}$, the *sons* of $\bar{u}$, which are the terminal states from an edge whose initial state corresponds to $\bar{u}$, represent the words which define the factors of the central basic factorization of u.

For better understanding the algorithm, we present the programming of some routines. The complete programming of the algorithm in Python may be found in http://cmup.fc.up.pt/cmup/amoura/DAautomaton_complete.py.

The main routine, called DAautomaton and described in Algorithm 1, constructs the automaton $\mathcal{A} = (V, E, \iota, e, \nu)$ by a recursive process. Initially, the automaton is presented as follows: the set of states, V, consists of the initial state ι, which corresponds to the word $\bar{w}$, and of the final state e, which corresponds to the word ε, the set of transitions is empty and the labeling relation has only the pair (e, ε).

```
def DAautomaton(input):
    e = ''
    iota = input
    V = [e,input]
    E = []
    nu = [[e,'epsilon']]
    V0 = []
    V1 = [input]
    while V1 != []:
        for w in V1:
            ll = LeftLabel(w)
            rl = RightLabel(w)
            F = Factorization(w,ll,rl)
            nu = nu+[F[0]]
            desc = F[1]
            if len(desc) == 3:
                for j in range(3):
                    E += [[w,j,desc[j]]]
                    if desc[j] not in V:
                        V += [desc[j]]
                        V0 += [desc[j]]
            else:
                E += [[w,0,desc[0]]]
                E += [[w,2,desc[1]]]
                for i in range(2):
                    if desc[i] not in V:
                        V += [desc[i]]
                        V0 += [desc[i]]
        V1 = V0
        V0 = []
    A = [V,E,iota,e,nu]
    return A
```

Algorithm 1

Let V_0 and V_1 be, respectively, the set of states which were not yet processed and the set of states which will be processed in the following step (which corresponds to run the **while** cycle once). Initially, V_0 is the empty set and V_1 consists of the initial state. The algorithm stops when these sets are both empty.

The process consists in the computation that we proceed to describe. Given a state of V_1, which corresponds to an ω-term u, we calculate the positions ll and rl in $\bar{u}$ of the labels of the central basic factorization of u. For that, we use two functions called LeftLabel and RightLabel, respectively.

We apply to this ω-term u and its label-positions the function Factorization, that it is described in detail in 5.2. The function computes the label of the state and keeps it in the list ν. It also produces the sons of this state. Then the main routine tests if each one of the sons is already in V. If it is not, it is added to V and to V_0 to be processed later. A transition is created that goes from the state that we are processing to the state corresponding to each son and labeled by the order of such son (i.e., 0, 1 or 2).

When $V_0 = \emptyset = V_1$, the routine stops. This means that all the elements have already been processed and all the states corresponding to DA-factors of the initial ω-term are already in the set of states of the automaton. Hence the DA-automaton, that we denote by $\mathcal{G}(w)$, is constructed.

5.2. **The factorization of an ω-term.** It is the function Factorization, described in Algorithm 2, that analyzes a state corresponding to a DA-factor of the initial ω-term. It takes as input the word that corresponds to the state that we are processing, $\bar{u}$, and the positions ll and rl, corresponding to the left label and to the right label of the central basic factorization of this DA-factor. It uses the function Parenthesis to compute the image of the partial function which associates to each position in $\bar{u}$ whose letter is a parenthesis the position corresponding to its matching pair. So that this information will be easily found, the function Parenthesis creates a list of length equal to $|\bar{u}|$ and puts the value -1 on the entries corresponding to the positions of $\bar{u}$ whose letter belongs to A.

```
def Factorization(w,ll,rl):
    m = -1
    P = Parenthesis(w)
    for i in range(len(P)):
        if i < ll < P[i] and i < rl < P[i]:
            m = i
            break
    if ll < rl or m != -1:
        nu = [w,w[ll]+w[rl]]
        desc = [S0forget(w,P,ll),S1remind(w,P,ll,rl,m),S2forget(w,P,rl)]
    elif ll > rl:
        nu = [w,w[rl]+w[ll]]
        desc = [S0remind(w,P,rl),S1forget(w,P,rl,ll),S2remind(w,P,ll)]
    else:
        nu = [w,w[ll]]
        desc = [S0forget(w,P,ll),S2forget(w,P,rl)]
    return [nu,desc]
```

ALGORITHM 2

The function Factorization verifies if the labels LeftLabel and RightLabel are inside a same ω-power and keeps the information, in a variable m, of the position where the largest ω-power that contains these labels begins. Then, it compares the values ll and rl, corresponding to the positions of the labels in the word $\bar{u}$. With this data, it determines the type of the central basic factorization. We have the following cases: if $ll < rl$ or $m \neq -1$, then the central basic factorization is of the standard form; if $ll > rl$ and $m = -1$, then the central basic factorization is of the overlapped form; if $ll = rl$ and $m = -1$, then the central basic factorization is degenerate. In the first case, we use the functions S0forget, S1remind and S2forget to construct the sons, while in the second and third cases we use, respectively, the functions S0remind, S1forget and S2remind, and the functions S0forget and S2forget. These functions are presented in the next subsection.

5.3. **The computation of the sons of an ω-term.** We present the functions that compute the sons of any state of the automaton. The functions consist on the construction of words from the word corresponding to the state that is being processed.

The functions whose name includes the word *forget* consider the subword of the initial word ending at $ll-1$, between $ll+1$ and $rl-1$, or starting at $rl+1$, depending on whether we are computing the son of the transition 0, 1 or 2, respectively, and consisting of all letters in A and all the matching parentheses in the considered interval. We show, for example, the function S0forget in Algorithm 3, which constructs the son of w from the transition labeled by 0.

```
def S0forget(w,P,ll):
    w0 = ''
    for i in range(ll):
        if w[i] != '(' or P[i] < ll:
            w0 += w[i]
    return w0
```

Algorithm 3

The functions whose name includes the word *remind* construct a word from the initial word considering all the ω-powers where the labels are inserted. We describe in detail the most intricate one, the function S1remind, presented in Algorithm 4.

```
def S1remind(w,P,ll,rl,m):
    w1 = ''
    if m == -1:
        for i in range(ll+1,rl):
            if (w[i] != '(' and w[i] != ')') or \
               (w[i] == ')' and P[i] > ll) or \
               (w[i] == '(' and P[i] < rl):
                w1 += w[i]
            elif w[i] == ')' and P[i] < ll:
                for l in range(P[i],i+1):
                    w1 += w[l]
            else:
                for l in range (i,P[i]+1):
                    w1 += w[l]
    else:
        M = P[m]
        for i in range(ll+1,M):
            if w[i] != ')' or P[i] > ll:
                w1 += w[i]
            else:
                for l in range(P[i],i+1):
```

```
                w1 += w[l]
        for i in range(m,M+1):
            w1 += w[i]
        for i in range(m+1,rl):
            if w[i] != '(' or P[i] < rl:
                w1 += w[i]
            else:
                for l in range(i,P[i]+1):
                    w1 += w[l]
    return w1
```

Algorithm 4

Firstly, the routine verifies the value of the parameter m. If it is different from -1, it means that the labels are in a same ω-power and the value of m is the position where the largest ω-power containing both labels begins. The son consists of the concatenation of the suffix of this ω-term beginning in the left label, with the respective ω-term and the prefix of it ending in the right label. Moreover, all the ω-powers containing one of the labels are concatenated as they are read. If the parameter m is equal to -1, meaning that the labels are not in a same ω-power, the routine constructs the son just reading the word from left to right and concatenating all the ω-powers containing one of the labels.

We finish with an example of a DA-automaton $\mathcal{G}(w)$ constructed by the described algorithm:

Example 5.1. Let $w = (ab^\omega caa^\omega)^\omega$ and $\bar{w} = (a(b)ca(a))$. We have LeftLabel $= c$ in the position $ll = 5$ and RightLabel $= b$ in the position $rl = 3$. As these labels are in the same ω-power, corresponding to the interval $[0, 10]$, it follows that the first occurrence of c appears before the last occurrence of b, when we read from left to right. Thus the central basic factorization of w is standard. The sons are calculated with the functions S0forget, S1remind and S2forget, respectively, and correspond to the following words: $\bar{w}_0 = a(b)$, $\bar{w}_1 = a(a)(a(b)ca(a))a(b)$ and $\bar{w}_2 = ca(a)$. The other states are constructed recursively. The DA-automaton $\mathcal{G}(w)$ associated to the ω-term $w = (ab^\omega caa^\omega)^\omega$ is described in Figure 1.

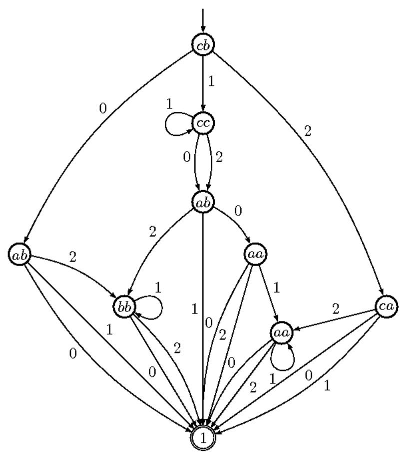


Figure 1. The DA-automaton $\mathcal{G}(w)$ associated to the ω-term $w = (ab^\omega caa^\omega)^\omega$.

5.4. **The complexity of the algorithm.** As explained in the previous subsections, the algorithm constructs, in each step, the factors of the central basic factorization of the ω-term that we are considering. However, nothing so far guarantees that the algorithm stops and, consequently, that the automaton $\mathcal{G}(w)$ is finite. This is what we propose to prove in this subsection together with the study of the complexity of the algorithm.

Let $\bar{w} = \mathbf{word}(w)$ be the input and let $|\bar{w}| = n$. For $l < n$, k_l is the number of pairs of parentheses containing the position l, $K = \max_{l<n} k_l$, $\mathfrak{l}_{(i,j)}$ is the length of the subword bounded by the pair of parentheses (i, j), with $0 \leq i, j < n$, i.e., the length of the subword corresponding to the ω-power (i, j), Φ_l is the sum of the lengths of the subwords corresponding to the ω-powers containing the position l and $\Phi = \max_{l<n} \Phi_l$.

Lemma 5.2. *Let $w \in T_A^\omega$. The length of an ω-term representing a* DA*-factor of the ω-word $\iota(w)$ is bounded above by $n + 2\Phi$.*

Proof. We give an upper bound for the length of the words corresponding to each vertex of $\mathcal{G}(w)$, using the parameters defined above.

We start by observing that the functions whose name includes *forget* create a word with length strictly smaller than the length of the input given to that function. So, it is enough to verify what happens when we apply to a word a function whose name includes *remind*. Consider the function S0remind and suppose that $\bar{w}_0 = \text{S0remind}(\bar{w})$. Note that, in this case, the central basic factorization of w is overlapped. Let rl be the position of the right label. Then we have $|\bar{w}_0| = rl + \sum_{i<rl<j}(\mathfrak{l}_{(i,j)} - 1) = rl + \Phi_{rl} - k_{rl} < n + \Phi$, because we insert in the prefix of the word $\bar{w}$ ending in rl the subwords corresponding to the ω-powers containing rl. Similarly, for $\bar{w}_2 = \text{S2remind}(\bar{w})$, we have $|\bar{w}_2| = ll + \Phi_{ll} - k_{ll} < n + \Phi$. If the central basic factorization of w is standard, we have $\bar{w}_1 = \text{S1remind}(\bar{w})$. It follows that $|\bar{w}_1| \leq (rl - ll - 1) + \sum_{i<ll<j}(\mathfrak{l}_{(i,j)} - 1) + \sum_{i<rl<j}(\mathfrak{l}_{(i,j)} - 1) = (rl - ll - 1) + (\Phi_{ll} - k_{ll}) + (\Phi_{rl} - k_{rl}) < n + 2\Phi$.

In the following iterations, we have the same procedure. When we cut the word to create the three sons, the functions *remind* add the subwords corresponding to the ω-powers containing the position where we cut. Note that, when this cut is done in a factor which had been added previously to the subword that issued from $\bar{w}$, the number of pairs of parentheses containing this position decreases and we have just those corresponding to the ω-powers which had not been added (when we read from the center to the borders). It follows that, in any depth that we are working, $|\bar{u}| < n + 2\Phi$, where $\bar{u}$ is a word corresponding to a state of the automaton. $\square$

We note that, in the above proof, we could use the number $(2K + 1)|\bar{w}|$ as an upper bound of $|\bar{u}|$. However, the upper bound that we have considered is smaller and easily computable. As the length of a word corresponding to a state of the automaton is bounded above and A is a finite alphabet, it follows that V, the set of states of the automaton, is finite. Hence $\mathcal{G}(w)$ is finite.

Corollary 5.3. *The automaton $\mathcal{G}(w)$ produced by the algorithm is finite.*

Although the previous lemma tells us that the number of states of $\mathcal{G}(w)$ is finite, we need to find a smaller upper bound for this number so we can show that the complexity of this construction is polynomial.

We consider the following sets:

$$Q(\bar{w}) = \{(i, j, p_i, p_j) \mid -1 \le i, j \le |\bar{w}|,\ \lambda(i), \lambda(j) \notin \{(,)\},\ 0 \le p_i \le k_i,\ 0 \le p_j \le k_j\}$$

and

$$T(\bar{w}) = \{\bar{w}_{(i,j,p_i,p_j)} \mid (i, j, p_i, p_j) \in Q(\bar{w})\}$$

where k_i and k_j are the numbers of pairs of parentheses containing the positions i and j, respectively, and $\bar{w}_{(i,j,p_i,p_j)}$ is the word obtained from $\bar{w}$ beginning at the position $i+1$, ending at the position $j-1$, reading, from left to right, the first p_i ω-powers containing i and reading, from right to left, the first p_j ω-powers containing the position j. We also use these parentheses as *bridges* to go from a higher position to a lower position (or the dual, when we read from right to left) and this is done at the largest ω-power containing both positions and that is read in any of the ways. If there is no ω-power to be read, then we go from a higher to a lower position by the smaller ω-power containing both positions i and j. The following example should help to understand this definition.

Example 5.4. Let $\bar{w} = a(b(cb)ab)a$. We have, for example, the following elements of $T(\bar{w})$:

$$\begin{aligned}
\bar{w}_{(-1,7,0,0)} &= ab(cb)\\
\bar{w}_{(-1,7,0,1)} &= a(b(cb)ab)b(cb)\\
\bar{w}_{(-1,5,0,1)} &= ab(cb)c\\
\bar{w}_{(-1,5,0,2)} &= a(b(cb)ab)b(cb)c\\
\bar{w}_{(5,4,0,0)} &= \varepsilon\\
\bar{w}_{(5,4,1,0)} &= (cb)\\
\bar{w}_{(5,4,1,2)} &= (cb)ab(b(cb)ab)b(cb)\\
\bar{w}_{(5,4,2,2)} &= (cb)ab(b(cb)ab)(b(cb)ab)b(cb).
\end{aligned}$$

Let $\Lambda_{\bar{w}} : Q(\bar{w}) \to T(\bar{w})$ be the function that maps each tuple $(i, j, p_i, p_j) \in Q(\bar{w})$ to the word $\bar{w}_{(i,j,p_i,p_j)} \in T(\bar{w})$.

Proposition 5.5. *The function $\Lambda_{\bar{w}} : Q(\bar{w}) \to T(\bar{w})$ has in its image all words corresponding to the states of $\mathcal{G}(w)$.*

Proof. Let $\bar{u}$ be a word corresponding to a state of $\mathcal{G}(w)$. Then $\bar{u}$ is a son of a word $\bar{v}$ and, therefore, $\bar{u}$ begins and ends, respectively, at positions i and j corresponding to the left and the right labels of the central basic factorization of $\bar{v}$ ($\bar{u} = \bar{v}_1$), or $i+1$ is the initial position of $\bar{v}$ and j is the position corresponding to one of the labels ($\bar{u} = \bar{v}_0$), or the dual ($\bar{u} = \bar{v}_2$). The numbers p_i and p_j correspond to the ω-powers containing i and j, respectively, that are considered when we read from i to j and from j to i, respectively. Note that the order in which these ω-powers appear, when we read from the borders to the center, is from that of the smallest length to that of the largest length. It follows that $\bar{u} = \bar{w}_{(i,j,p_i,p_j)}$ for the values i, j, p_i and p_j chosen above. $\square$

We note that $\Lambda_{\bar{w}}$ is not an injective function. For example, the empty word is the image of all pairs of the form $(i, i+1, 0, 0)$, $-1 \le i < |\bar{w}|$. Moreover, the elements $\bar{w}_{(i,j,p_i,p_j)}$ and $\bar{w}_{(i,j,p_i-1,p_j+1)}$ may have the same image under $\Lambda_{\bar{w}}$. This follows from the fact that the p_i-th ω-power when we read from the left coincides with the (p_j+1)-th ω-power when we read from the right. By Proposition 5.5, we have the following:

Corollary 5.6. *The number of states of $\mathcal{G}(w)$ is at most $(|\bar{w}|+2)^2(K+1)^2$.*

Now, we are ready to determine the complexity of our algorithm. The main function that constructs the automaton consists of a routine that processes each state of the automaton once. For each element of V, it tests if this state has already been processed, involving $\mathcal{O}(|V|) \leq \mathcal{O}(|\bar{w}|^2K^2)$ steps. Then, it computes the left and the right labels with the respective functions. These functions read each letter of the word and, whenever a new letter is found, it is kept in the variable ll (respectively, rl). The complexity of these functions is $\mathcal{O}(|A|(|\bar{w}| + \Phi))$. Afterwards, the algorithm constructs the sons of the state that is being processed using the function Factorization. This function uses the function Parenthesis and the functions to compute the sons. The function Parenthesis reads the word and computes a list with the positions of the pairs of matching parentheses, with complexity $\mathcal{O}(|\bar{w}|+\Phi)$. The functions which construct the sons read the word corresponding to the state and the ω-powers that will be considered in the new word. So, the complexity of that is $\mathcal{O}(|\bar{w}| + \Phi)$. It follows that the complexity of the function Factorization is $\mathcal{O}(|\bar{w}| + \Phi)$. Hence, the complexity of the algorithm is:

$$|V| \cdot \mathcal{O}(|V| + 2|A|(|\bar{w}| + \Phi) + 3(|\bar{w}| + \Phi)) \leq \mathcal{O}(|\bar{w}|^4K^4). \tag{1}$$

We have proved the following theorem:

Theorem 5.7. *The algorithm that constructs the automaton $\mathcal{G}(w)$, described in the previous subsections, has complexity not exceeding $\mathcal{O}(|\bar{w}|^4K^4)$.*

We have already observed, after Lemma 5.2, that $(2K + 1)|\bar{w}|$ is a higher upper bound for the length of a word corresponding to a state than the upper bound established in the proof of the lemma. However, we use this number make it easier to prove the inequality 1.

Probably, an improvement of the programming and the discovery of a smaller upper bound for the number of states of the automaton allow us to compute a smaller upper bound to the complexity of the computation of $\mathcal{G}(w)$. However, this upper bound can not be smaller than $\mathcal{O}(|\bar{w}|^2)$, as we can see by the following example:

Example 5.8. We consider the sequence of words $(\bar{w}_n)_{n\in\mathbb{N}}$ where $\bar{w}_n = (a_n(a_{n-1}(\cdots(a_1))))$, with $a_i \neq a_j$, if $i \neq j$. We have $|\bar{w}_n| = 3n$ and $|A_n| = n$, where A_n is the alphabet involved in $\bar{w}_n$. We compute the number of states of $\mathcal{G}(w_n)$, $|V_n|$, by recurrence.

For $n = 1$, $\bar{w}_1 = (a_1)$, and for $n = 2$, $\bar{w}_2 = (a_2(a_1))$, the words corresponding to the DA-factors are, respectively, (a_1) and ε, and $(a_2(a_1))$, a_2, $(a_1)(a_2(a_1))$, (a_1) and ε. Hence $\mathcal{G}(w_1)$ and $\mathcal{G}(w_2)$ have, respectively, 2 and 5 states.

Let $\bar{w}_n = (a_n(a_{n-1}(\cdots(a_1))))$, with $n \geq 3$. The central basic factorization of w_n produces the following sons: $a_na_{n-1}\cdots a_2$, $(a_1)(a_2(a_1))\cdots(a_n(a_{n-1}(\cdots(a_1)))) = \bar{w}_1\bar{w}_2\cdots\bar{w}_n$ and $(a_{n-1}(\cdots(a_1))) = \bar{w}_{n-1}$. Thus, the number of states of $\mathcal{G}(w_n)$ is the sum of the number of states of $\mathcal{G}(w_{n-1})$ with the other states corresponding to the DA-factors of w_n and that are not DA-factors of w_{n-1}. Let $\bar{w}_{n(0)} = a_na_{n-1}\cdots a_2$ and $\bar{w}_{n(1)} = (a_1)(a_2(a_1))\cdots(a_n(a_{n-1}(\cdots(a_1)))) = \bar{w}_1\bar{w}_2\cdots\bar{w}_n$ be, respectively, the sons of $\bar{w}_n$ by the edges labeled 0 and 1. The successive iterations of the central basic factorization of $\bar{w}_{n(0)}$ produce the factors $a_{n-1}\cdots a_3$, $a_{n-2}\cdots a_4$, $\ldots$, and $a_{\frac{n+3}{2}}a_{\frac{n+1}{2}}$ (respectively, $a_{\frac{n}{2}}$, if n is even). Note that these factors are not states of $\mathcal{G}(w_{n-1})$. Hence $\bar{w}_n$ has $\frac{n-1}{2}$ factors (respectively, $\frac{n}{2}$ factors, if n is even) which are descendants from the left edge of the state $\bar{w}_n$. On the other hand, the central basic factorization of $\bar{w}_{n(1)}$ produces the factors $\bar{w}_1\bar{w}_2\cdots\bar{w}_{n-1}$, $\bar{w}_{n-1}\bar{w}_n$ and $\bar{w}_{n-1}$. Note that $\bar{w}_{n-1}\bar{w}_n$ is the only factor which is not a state of $\mathcal{G}(w_{n-1})$, since it has in its content the letter a_n. Moreover, the central basic factorization of this factor produces the factors $\bar{w}_{n-1}$ and $\bar{w}_{n-1}\bar{w}_n$, which were already counted. Thus, we count two new factors which are descendants from the central branch. We have the following recurrence formula for the number of states of $\mathcal{G}(w_n)$, with $n \geq 3$:

$$|V_n| = |V_{n-1}| + 3 + \left\lfloor \frac{n}{2} \right\rfloor$$

and, therefore, using basic calculus, we have, for $m \geq 1$,

$$|V_{2m+1}| = 9 + (m+8)(m-1)$$

and

$$|V_{2m}| = 5 + (m+7)(m-1).$$

Hence, the number of states of $\mathcal{G}(w_n)$ is $\Omega(|\bar{w}|^2)$.

Given an automaton $\mathcal{G}(w)$, we construct the finite automaton that recognizes $\mathcal{L}(w)$ by replacing the label of each edge of $\mathcal{G}(w)$ by the ordered pair whose first component is the label of the edge in $\mathcal{G}(w)$ and the second component is the label of the initial state of the edge in $\mathcal{G}(w)$. After that, we minimize the automaton. Brzozowski's Algorithm [7] and Hopcroft's Algorithm [8] to minimize a finite deterministic automaton are well known and they have exponential and $\mathcal{O}(lm\log m)$ complexity, respectively, where l is the cardinality of the alphabet and m is the number of states of the automaton. However, Almeida and Zeitoun [6] described an algorithm to minimize a finite deterministic automaton whose strongly connected non-trivial components are cycles, in time $\mathcal{O}(l+d)$, where d is the number of transitions of the automaton. Note that $\mathcal{G}(w)$ satisfies this condition, since in any cycle of $\mathcal{G}(w)$ the edges are labeled by $(1,x)$, with $x \in A \times A \cup A$ and there is only one edge going from each state with first component labeled 1. As the number of states of the automaton is bounded above by $(|\bar{w}|+2)^2(K+1)^2$, the number $3(|\bar{w}|+2)^2(K+1)^2$ is an upper bound for the number of transitions of the automaton. Furthermore, in 1971, Hopcroft and Karp [9] presented a linear algorithm for testing the equivalence of two finite deterministic automata without requiring previous minimization. So, we have established the following result:

Theorem 5.9. *The word problem for ω-terms over* DA *has a solution in polynomial time, not exceeding $\mathcal{O}((nK)^4)$, where n is the length of the word corresponding to the ω-term and K is the maximum depth of ω-powers.*

Example 5.10. The minimal DA-automaton of the ω-term $w = (ab^\omega caa^\omega)^\omega$ is represented in Figure 2. It follows from identifying states v_{120} and v_{1201} of the automaton $\mathcal{G}(w)$ presented in the Example 5.1. Note that the state v_{120} corresponds to the ω-term a^ω and the state v_{1201} corresponds to the ω-term aa^ω, which are equal over DA.

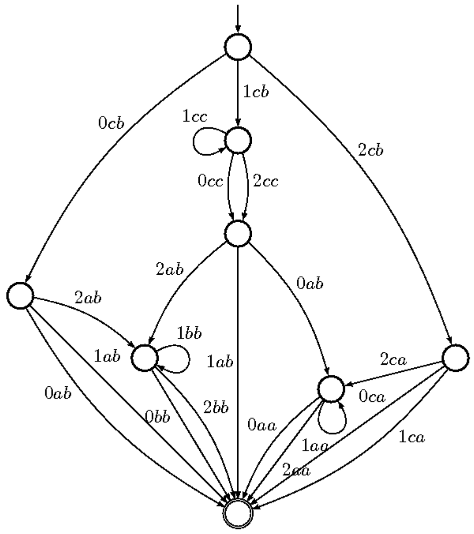


FIGURE 2. The minimal DA-automaton associated to $w = (ab^\omega caa^\omega)^\omega$.

ACKNOWLEDGMENTS

This work is part of the author's doctoral thesis, written under the supervision of Prof. Jorge Almeida, from whose advice the author has greatly benefited. This work was supported by the *Fundação para a Ciência e a Tecnologia (FCT)* through the *PhD grant* SFRH/BD/19720/2004, through the *Centro de Matemática da Universidade do Porto (CMUP)* and also through the project PTDC/MAT/65481/2006, which is partly funded by the European Community Fund FEDER.

REFERENCES

[1] J. Almeida, *Finite Semigroups and Universal Algebra*, World Scientific, Singapore, 1994. English translation.

[2] ———, *A syntactical proof of locality of* DA, Int. J. Algebra and Comp. **6** (1996) 165–177.

[3] ———, *Finite semigroups: an introduction to a unified theory of pseudovarieties*, in Semigroups, Algorithms, Automata and Languages, J.-E. P. G. M. S. Gomes and P. V. Silva, eds., Singapore, 2002, World Scientific, 3–64.

[4] J. Almeida and M. Zeitoun, *The equational theory of ω-terms for finite* R*-trivial semigroups*, in Proc. of the Workshop on Semigroups and Languages, Lisbon 2002, V. H. F. I. M. Araújo M. J. J. Branco and G. M. S. Gomes, eds., World Scientific, 2004, 1–22.

[5] ———, *An automata-theoretic approach to the word problem for ω-terms over* R, Theor. Comp. Sci. **370** (2007) 131–169.

[6] ———, *Description and analysis of a bottom-up DFA minimization algorithm*, Inform. Process. Lett. **107** (2008) 52–59.

[7] J. Brzozowski, *Canonical regular expressions and minimal state graphs for definitive events*, in Proc. Symp. Math. Theor. Automata, New York, 1963, 529–561.

[8] J. E. Hopcroft, *An $n \log n$ algorithm for minimizing states in a finite automaton*, in Theory of machines and computations (Proc. Internat. Sympos., Technion, Haifa, 1971), Z. Kohavi, ed., Academic Press, 1971, 189–196.

[9] J. E. Hopcroft and R. M. Karp, *A linear algorithm for testing equivalence of finite automata*, tech. rep., Cornell University, 1971.

[10] J. McCammond, *Normal forms for free aperiodic semigroups*, Int. J. Algebra and Comp. **11** (2001) 581–625.

[11] A. Moura, *Representations of the free profinite object over* DA, Tech. Rep. CMUP 2009-26, Univ. of Porto, 2009. `http://cmup.fc.up.pt/cmup/v2/include/filedb.php?id=275&table=publicacoes&field=file`.

[12] J. G. Rosenstein, *Linear Orderings*, Academic Press, New York, 1982.

[13] P. Tesson and D. Thérien, *Diamonds are forever: the variety* DA, in Semigroups, Algorithms, Automata and Languages, World Scientific, 2002, 475–499.

Instituto Superior de Engenharia do Porto/LEMA and Centro de Matemática da Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

E-mail address: `aim@isep.ipp.pt/amoura@fc.up.pt`

Paper 3

IDEMPOTENT-GENERATED SEMIGROUPS AND PSEUDOVARIETIES

J. ALMEIDA AND A. MOURA

ABSTRACT. The operator which constructs the pseudovariety generated by the idempotent-generated semigroups of a given pseudovariety is investigated. Several relevant examples of pseudovarieties generated by their idempotent-generated elements are given as well as some properties of this operator. Particular attention is paid to the pseudovarieties in $\{\mathsf{J}, \mathsf{R}, \mathsf{L}, \mathsf{DA}\}$ concerning this operator and their generator ranks and idempotent-generator ranks.

1. INTRODUCTION

Due to its applications in Computer Science, the theory of finite semigroups saw significant advancements in the 1960's driven by developments in the theory of finite automata. This connection with finite semigroups was firstly explored to obtain computational results. In parallel, combinatorial and algebraic properties of finite semigroups were studied. Eilenberg [4] established a correspondence between certain families of rational languages and certain classes of finite semigroups, called *pseudovarieties*, which provided a suitable framework for the bridges between the two theories.

There are many important pseudovarieties, often constructed from other ones by applying suitable operators. Some natural operators have been extensively studied. In this paper, we introduce a new one which constructs the subpseudovariety generated by the idempotent-generated semigroups of a given pseudovariety.

Several works have been dedicated to idempotent-generated semigroups. It is well-known that any finite semigroup embeds into a finite regular idempotent-generated semigroup, which was proved by Howie [6] using full transformations semigroups. Howie [7] also proved that the full transformations subsemigroup consisting of all order-preserving and contractive full transformations is idempotent-generated. Laradji and Umar [8] improved this result and showed that, in fact, every order-preserving and contractive full transformation is expressible as a product of idempotents of the same type and with the same range. The analogous result for the subsemigroup of contractive full transformations also holds [8].

On the other hand, Petrich and Reilly [14] proved that every completely simple semigroup embeds into an idempotent-generated one. Furthermore, Petrich [13] presents a concrete model of the embedding due to Pastijn and Yan [12] of a semigroup into an idempotent-generated Rees matrix semigroup that fixes some properties.

2000 *Mathematics Subject Classification.* 20M07.

Key words and phrases. Finite semigroup; pseudovariety; operator; idempotent-generated semigroup; generator rank; idempotent generator rank.

In this paper, while we do not obtain a complete characterization of the pseudovarieties which are generated by their idempotent-generated semigroups, we prove that many familiar pseudovarieties have this property. The techniques used for this purpose include the representations of free profinite semigroups over R, J and DA due to Almeida and Weil [3], Almeida [1], and Moura [9], respectively. In the cases of R and J, we also observe an alternative approach based on some results concerning transformations of a finite chain due to Pin [15] and Straubing [18] and the results concerning idempotent-generated subsemigroups of full transformations from Howie, and Laradji and Umar. On the other hand, the work of Petrich [13] allows us to show that the pseudovarieties $\bar{\mathsf{H}}$, where H is a pseudovariety of groups, CS and CR also have this property.

The new approach in the case of the pseudovarieties R and J is justified as it gives a significant improvement in terms of the *generator rank* and *idempotent generator rank*. More generally, we show that both ranks are infinite for every pseudovariety in the interval $[\mathsf{J}, \mathsf{DS}]$. We also prove that every semigroup in the subpseudovariety generated by all n-generated members of any of the pseudovarieties $\mathsf{J}, \mathsf{R}, \mathsf{L}, \mathsf{DA}$ divides a semigroup in the same pseudovariety generated by at most $n + 1$ idempotents. We compare these results with the works of Umar [19], and Laradji and Umar [8] concerning the ranks and idempotent-ranks of the subsemigroups of all contractive full transformations, and contractive and order-preserving full transformations, respectively. We observe that, in fact, we decrease the number of idempotent generators of the idempotent-generated semigroups when we use the embeddings of the semigroups of R and J presented in this paper.

The paper is organized as follows. In Section 2, we recall some basics of the theory of pseudovarieties of semigroups and profinite semigroups and we introduce some notation concerning operators on pseudovarieties. We also present a list of the pseudovarieties and bases of pseudoidentities defining them that will be used in our study. In Section 3, we observe some properties of the operator $\_{\mathsf{E}}$, we determine some pseudovarieties of the form V_{E} and we make a short introduction of the main question addressed in the paper: what are the pseudovarieties that are generated by their idempotent-generated elements? We present in the following sections some pseudovarieties having this property: in Section 4 using the embedding in a Rees matrix semigroup constructed by Petrich, and in Section 5 using representations of the free profinite semigroup. Finally, in Section 6 we determine the generator rank and idempotent generator rank of every pseudovariety in the interval $[\mathsf{J}, \mathsf{DS}]$ and we also determine a lower bound for the idempotent generator rank of the subpseudovarieties generated by all n-generated members of any pseudovariety in the interval $[\mathsf{J}, \mathsf{DA}]$. Combining with the results of Section 5, we improve the last result showing that the lower bound is the exact value in the case of the pseudovarieties $\mathsf{J}, \mathsf{R}, \mathsf{L}, \mathsf{DA}$. To introduce some relevant results in our study, we develop some existing techniques that need to be recalled. Rather than including them in Section 2, we briefly introduce them along as need.

2. Preliminaries

We assume acquaintance with notions concerning pseudovarieties of semigroups and profinite semigroups. We briefly recall some basics and we refer the reader to [1, 2, 15] for detailed information.

For a semigroup S, let S^1 be the monoid obtained by adjoining a neutral element 1 to S in case S does not already possess one, and $S^1 = S$ otherwise. We denote by $E(S)$ the set of idempotents of S and by $\langle E(S)\rangle$ the subsemigroup of S generated by $E(S)$. For $s \in S$, s^ω denotes the unique idempotent in the subsemigroup generated by s. We say that a semigroup S *divides* a semigroup T, and we write $S \prec T$, if there exists a surjective homomorphism of a subsemigroup of T onto S.

A *pseudovariety* of semigroups is a class of finite semigroups that is closed under taking subsemigroups, homomorphic images and finite direct products. Equivalently, a pseudovariety of semigroups is a class of finite semigroups closed under taking division and finite direct products. For example, S is the pseudovariety of all finite semigroups.

There are many ways to construct new pseudovarieties from known ones, that is by applying operators to pseudovarieties. For example, given a pseudovariety V, the following classes of finite semigroups are pseudovarieties:

- EV consists of all $S \in \mathsf{S}$ such that $\langle E(S)\rangle \in \mathsf{V}$;
- DV consists of all $S \in \mathsf{S}$ such that, for every regular $\mathcal{D}$-class D of S, $D \in \mathsf{V}$;
- for a pseudovariety H of groups, $\bar{\mathsf{H}}$ consists of all $S \in \mathsf{S}$ such that every subgroup G of S belongs to H.

We also have other types of operators that construct new pseudovarieties by describing the generators. The new pseudovariety is then the smallest pseudovariety containing such semigroups. In this way, we introduce the operator V_E, which is the topic of this paper. Given a pseudovariety V, we define V_E as the pseudovariety generated by the subsemigroups generated by the idempotents of each member of V, i.e.,

$$\mathsf{V}_\mathsf{E} = \langle \langle E(S)\rangle \mid S \in \mathsf{V}\rangle.$$

Note that $\mathsf{V}_\mathsf{E} \subseteq \mathsf{V}$ as the indicated generators of V_E belong to V.

Because it will be useful in our study, we present an obvious observation about the subsemigroup generated by a subset of idempotents of a given semigroup:

Lemma 2.1. *Let $S \in \mathsf{S}$ and $X \subseteq E(S)$. Then $\langle E\langle X\rangle\rangle = \langle X\rangle$. In particular, we have $\langle E\langle E(S)\rangle\rangle = \langle E(S)\rangle$ and $\langle E\langle E(D)\rangle\rangle = \langle E(D)\rangle$, for every regular $\mathcal{D}$-class D of S.*

A semigroup equipped with a topology for which the multiplication is a continuous function is called a *topological semigroup*. Finite semigroups are endowed with the discrete topology. A topological semigroup S is a *profinite semigroup* (respectively, a *pro-*V *semigroup*) if it is a compact semigroup which is *residually finite* (respectively, *residually in* V), which means that, for any two distinct elements of S, there exists a continuous homomorphism into a finite semigroup (respectively, into a member of V) that separates them. Equivalently, profinite semigroups are compact 0-*dimensional*, which means that the topology has an open basis consisting of clopen sets. The elements of a pseudovariety V are pro-V semigroups.

We denote by $\overline{\Omega}_A\mathsf{V}$ the *free pro-*V *semigroup on* A, which is the unique (up to isomorphism of topological semigroups) pro-V semigroup such that every mapping $\mu : A \to S$ into a pro-V semigroup S can be extended to a unique continuous homomorphism $\hat{\mu} : \overline{\Omega}_A\mathsf{V} \to S$ such that $\hat{\mu} \circ \iota = \mu$, where $\iota : A \to \overline{\Omega}_A\mathsf{V}$ is the natural generating function (i.e., its image generates a dense subsemigroup of $\overline{\Omega}_A\mathsf{V}$). The elements of $\overline{\Omega}_A\mathsf{V}$ are called *implicit operations over* V. For $u \in \overline{\Omega}_A\mathsf{V}$ the sequence

$(u^{n!})_n$ converges and we denote the limit by u^ω, which is the unique idempotent in the closed subsemigroup generated by u.

An equality of the form $u = v$, with $u, v \in \overline{\Omega}_A\mathsf{S}$, is called a *pseudoidentity* and $|A|$ is its *arity*. The pseudoidentity is *valid* in a profinite semigroup S, and we write $S \models u = v$, if, for every continuous homomorphism $\varphi : \overline{\Omega}_A\mathsf{S} \to S$, we have $\varphi(u) = \varphi(v)$. It is easy to see that the validity of a pseudoidentity in a finite semigroup is preserved under taking homomorphic images, subsemigroups and finite direct products. Conversely, Reiterman's Theorem [16] says that every pseudovariety is defined by some set of finitary pseudoidentities. We end this section with a list of pseudovarieties that will be used in this paper and some corresponding bases of pseudoidentities defining them:

$\mathsf{J} = [\![(xy)^\omega x = (xy)^\omega = y(xy)^\omega]\!]$	$\mathcal{J}$-trivial semigroups;
$\mathsf{R} = [\![(xy)^\omega x = (xy)^\omega]\!]$	$\mathcal{R}$-trivial semigroups;
$\mathsf{L} = [\![y(xy)^\omega = (xy)^\omega]\!]$	$\mathcal{L}$-trivial semigroups;
$\mathsf{A} = [\![x^{\omega+1} = x^\omega]\!]$	aperiodic (or $\mathcal{H}$-trivial) semigroups;
$\mathsf{G} = [\![x^\omega = 1]\!]$	groups;
$\mathsf{LG} = [\![(x^\omega y)^\omega x^\omega = x^\omega]\!]$	local groups;
$\mathsf{CR} = [\![x^{\omega+1} = x]\!]$	completely regular semigroups;
$\mathsf{CS} = [\![x^{\omega+1} = x, (xyx)^\omega = x^\omega]\!]$	completely simple semigroups;
$\mathsf{RB} = [\![x^2 = x, xyx = x]\!]$	rectangular bands;
$\mathsf{DA} = [\![((xy)^\omega x)^2 = (xy)^\omega x]\!]$	regular $\mathcal{D}$-classes are aperiodic semigroups;
$\mathsf{DG} = [\![(xy)^\omega = (yx)^\omega]\!]$	regular $\mathcal{D}$-classes are groups;
$\mathsf{DO} = [\![(xy)^\omega (yx)^\omega (xy)^\omega = (xy)^\omega]\!]$	regular $\mathcal{D}$-classes are orthodox semigroups;
$\mathsf{DS} = [\![((xy)^\omega x)^{\omega+1} = (xy)^\omega x]\!]$	regular $\mathcal{D}$-classes are semigroups.

3. Properties of the operator $\_{\mathsf{E}}$

We establish some basic properties of the operator $\_{\mathsf{E}}$. We start by observing that the definition given for $\mathsf{V_E}$, where V is a pseudovariety, is equivalent to $\mathsf{V_E}$ being generated by the idempotent-generated semigroups of V.

Lemma 3.1. *The operator* $\_{\mathsf{E}}$ *has the following properties, where* V *and* W *are arbitrary pseudovarieties:*

(1) $\mathsf{V_E} = \langle T \in \mathsf{V} \mid T = \langle E(T)\rangle\rangle$*;*
(2) $\mathsf{V} \subseteq \mathsf{W}$ *implies* $\mathsf{V_E} \subseteq \mathsf{W_E}$*;*
(3) $(\mathsf{V} \cap \mathsf{W})_\mathsf{E} \subseteq \mathsf{V_E} \cap \mathsf{W_E}$*;*
(4) $(\mathsf{V_E})_\mathsf{E} = \mathsf{V_E}$*;*
(5) $(\mathsf{EV})_\mathsf{E} = \mathsf{V_E}$*;*
(6) $\mathsf{E}(\mathsf{V_E}) = \mathsf{EV}$*.*

Proof. (1) Let $T = \langle E(S)\rangle$ with $S \in \mathsf{V}$. Since $\langle E(S)\rangle$ is a subsemigroup of S, it follows that $T \in \mathsf{V}$. By Lemma 2.1, we have that $\langle E(T)\rangle = \langle E\langle E(S)\rangle\rangle = \langle E(S)\rangle = T$. Hence the generators of the two pseudovarieties are the same.

(2) is immediate from the definition of $\mathsf{V_E}$ and (3) follows from (2).

(4) The direct inclusion follows from $\mathsf{V_E} \subseteq \mathsf{V}$ and (2). Conversely, since the generators of $\mathsf{V_E}$ are the semigroups $\langle E(S)\rangle$, with $S \in \mathsf{V}$, it suffices to show that

$\langle E(S)\rangle \in (\mathsf{V_E})_\mathsf{E}$, for all $S \in \mathsf{V}$. Indeed, since $\langle E(S)\rangle \in \mathsf{V_E}$, by Lemma 2.1 and definition of $\_\mathsf{E}$ we have $\langle E(S)\rangle = \langle E\langle E(S)\rangle\rangle \in (\mathsf{V_E})_\mathsf{E}$.

(5) We have $(\mathsf{EV})_\mathsf{E} = \langle\langle E(S)\rangle \mid S \in \mathsf{EV}\rangle = \langle\langle E(S)\rangle \mid \langle E(S)\rangle \in \mathsf{V}\rangle$. Let us see that the generators of $(\mathsf{EV})_\mathsf{E}$ are in $\mathsf{V_E}$. In fact, as $\langle E(S)\rangle \in \mathsf{V}$, Lemma 2.1 yields $\langle E(S)\rangle = \langle E\langle E(S)\rangle\rangle \in \mathsf{V_E}$. The reverse inclusion follows from $\mathsf{V} \subseteq \mathsf{EV}$ and (2).

(6) Since $\mathsf{V_E} \subseteq \mathsf{V}$, applying the increasing operator $\mathsf{E}\_$, we obtain $\mathsf{E}(\mathsf{V_E}) \subseteq \mathsf{EV}$. If $S \in \mathsf{EV}$, i.e., $\langle E(S)\rangle \in \mathsf{V}$, then Lemma 2.1 gives that $\langle E(S)\rangle = \langle E\langle E(S)\rangle\rangle$ is one of the generators of $\mathsf{V_E}$ so that, in particular, $S \in \mathsf{E}(\mathsf{V_E})$. □

A natural question, for which we have no answer, is whether we always have equality in part (3) of Lemma 3.1.

Corollary 3.2. *Let* V *and* W *be pseudovarieties such that* $\mathsf{V_E} = \mathsf{V}$ *and* $\mathsf{EV} = \mathsf{EW}$. *Then* $\mathsf{V} \subseteq \mathsf{W}$.

Proof. Applying the operator $\_\mathsf{E}$ to $\mathsf{EV} = \mathsf{EW}$ and using property (5) of Lemma 3.1, it follows that $\mathsf{V} = \mathsf{V_E} = (\mathsf{EV})_\mathsf{E} = (\mathsf{EW})_\mathsf{E} = \mathsf{W_E} \subseteq \mathsf{W}$. □

Corollary 3.3. *Given two pseudovarieties* V *and* W*, the following conditions are equivalent:*

(a) $\mathsf{V_E} = \mathsf{W_E}$*;*
(b) $\mathsf{EV} = \mathsf{EW}$*;*
(c) $\mathsf{V_E} \subseteq \mathsf{W} \subseteq \mathsf{EV}$.

Proof. (a) ⇒ (b) From (a) and property (6) of Lemma 3.1, it follows that $\mathsf{EV} = \mathsf{E}(\mathsf{V_E}) = \mathsf{E}(\mathsf{W_E}) = \mathsf{EW}$.

(b) ⇒ (c) The second inclusion in (c) follows from $\mathsf{W} \subseteq \mathsf{EW} = \mathsf{EV}$. To show the first inclusion, we recall that, by properties (4)–(6) of Lemma 3.1, $((\mathsf{EV})_\mathsf{E})_\mathsf{E} = (\mathsf{EV})_\mathsf{E}$ and $\mathsf{E}((\mathsf{EV})_\mathsf{E}) = \mathsf{EV} = \mathsf{EW}$. Hence, by Corollary 3.2, $(\mathsf{EV})_\mathsf{E} \subseteq \mathsf{W}$. Moreover, also by property (5) of Lemma 3.1, $\mathsf{V_E} = (\mathsf{EV})_\mathsf{E} \subseteq \mathsf{W}$.

(c) ⇒ (a) Applying the operator $\_\mathsf{E}$ to (c), by properties (2), (4) and (5) of Lemma 3.1, we obtain $\mathsf{V_E} = (\mathsf{V_E})_\mathsf{E} \subseteq \mathsf{W_E} \subseteq (\mathsf{EV})_\mathsf{E} = \mathsf{V_E}$ and, therefore, $\mathsf{W_E} = \mathsf{V_E}$. □

In other words, given a pseudovariety V, the equations $\mathsf{X_E} = \mathsf{V_E}$ and $\mathsf{EX} = \mathsf{EV}$ in the variable X are equivalent and the class of its solutions is the interval $[\mathsf{V_E}, \mathsf{EV}]$.

It is natural to ask for which pseudovarieties V, $\mathsf{V_E}$ is equal to V. As an obvious example, for every pseudovariety V of bands, since its semigroups consist only of idempotents, we have $\mathsf{V_E} = \mathsf{V}$. But, there are pseudovarieties that do not satisfy the equality $\mathsf{V} = \mathsf{V_E}$. Let us see some examples:

Example 3.4. For every pseudovariety H of groups, we have $\mathsf{H_E} = \mathsf{I}$, where $\mathsf{I} = [\![x = y]\!]$ is the trivial pseudovariety.

Example 3.5. It is well known that LG is the class of all finite semigroups such that all idempotents are $\mathcal{J}$-equivalent and, therefore, they are in the minimal ideal of the semigroup. So we have $(\mathsf{LG})_\mathsf{E} = \langle\langle E(S)\rangle \mid S \in \mathsf{LG}\rangle \subseteq \mathsf{CS} \subsetneq \mathsf{LG}$.

Example 3.6. It is well known that $\mathsf{RB} \vee \mathsf{G} = \mathsf{CS} \cap \mathsf{O}$, where $\mathsf{O} = [\![(x^\omega y^\omega)^\omega = x^\omega y^\omega]\!]$ is the class of all finite orthodox semigroups. So we have $(\mathsf{RB} \vee \mathsf{G})_\mathsf{E} = (\mathsf{CS} \cap \mathsf{O})_\mathsf{E} = \mathsf{RB}$.

The notion of E-*local* pseudovariety, introduced in [10], enables us to determine $(\mathsf{DO})_{\mathsf{E}}$ and $(\mathsf{DH})_{\mathsf{E}}$, as we see in the following examples. Recall that a pseudovariety V is E-*local* if it satisfies the following property: given $S \in \mathsf{S}$, $\langle E(S)\rangle \in \mathsf{V}$ if and only if $\langle E(D)\rangle \in \mathsf{V}$, for every regular $\mathcal{D}$-class D of S.

Example 3.7. Let $S \in \mathsf{DH}$. Since every regular $\mathcal{D}$-class D of S is a group, it follows that $\langle E(D)\rangle$ is trivial and, therefore, $\langle E(D)\rangle \in \mathsf{J}$. Since J is E-local (see [10, Example 3.6]), we have $\langle E(S)\rangle \in \mathsf{J}$. Hence $\mathsf{J} \subseteq \mathsf{DH} \subseteq \mathsf{EJ}$, where the first inclusion is trivial. Thus it follows from Corollary 3.3 that $(\mathsf{DH})_{\mathsf{E}} = \mathsf{J}_{\mathsf{E}} = \mathsf{J}$, where the last equality follows from Corollary 3.14, which is proved below.

Example 3.8. We observe that $\mathsf{DO} \subseteq \mathsf{EDA}$. Indeed, for $S \in \mathsf{DO}$ and a regular $\mathcal{D}$-class D of S, $\langle E(D)\rangle \in \mathsf{DA}$. Since DA is E-local (see [10, Proposition 3.5]), we have $\langle E(S)\rangle \in \mathsf{DA}$. Hence $S \in \mathsf{EDA}$. Since $\mathsf{DA} \subseteq \mathsf{DO} \subseteq \mathsf{EDA}$, it follows from Corollary 3.3 that $(\mathsf{DO})_{\mathsf{E}} = (\mathsf{DA})_{\mathsf{E}} = \mathsf{DA}$, where the last equality follows from Corollary 5.6 which is established in Section 5.

In an attempt to identify the pseudovarieties which are generated by their idempotent-generated elements, we present some results in the following sections. We start by suggesting, as an easy exercise, the result from Howie [6] which states that any finite semigroup embeds into a finite regular idempotent-generated semigroup, so that, in particular, we have the following result.

Proposition 3.9 (cf. Howie [6])**.** $\mathsf{S}_{\mathsf{E}} = \mathsf{S}$.

On the other hand, Pin [15] and Straubing [18] obtained the following representation theorems for $\mathcal{R}$-trivial monoids and $\mathcal{J}$-trivial monoids, respectively.

Theorem 3.10 (cf. Pin [15, Theorem IV.3.6])**.** *A finite monoid is $\mathcal{R}$-trivial if and only if it is a submonoid of $\mathcal{E}_X$, the submonoid consisting of all contractive total transformations of some finite chain X.*

Theorem 3.11 (cf. Straubing [18])**.** *A finite monoid is $\mathcal{J}$-trivial if and only if it divides $\mathcal{C}_X$, the submonoid of all order-preserving and contractive transformations of some finite chain X.*

Combining these theorems with the following results about idempotent-generated subsemigroups of total transformations due, respectively, to Laradji and Umar [8] and to Howie [7], we obtain Corollary 3.14.

Theorem 3.12 (cf. Laradji and Umar [8, Theorem 1.3])**.** *The monoid $\mathcal{E}_X$ is idempotent-generated.*

Theorem 3.13 (cf. Howie [7, Theorem 3.2])**.** *The monoid $\mathcal{C}_X$ is idempotent-generated.*

Corollary 3.14. *The equality $\mathsf{V}_{\mathsf{E}} = \mathsf{V}$ holds if V is any of the pseudovarieties $\mathsf{R}, \mathsf{L}, \mathsf{J}$.*

Similarly, using embeddings into idempotent-generated semigroups of the same type from Petrich and Reilly [14] and from Pastijn and Yan [12] concerning, respectively, completely simple semigroups and completely regular semigroups, we obtain the following results.

Proposition 3.15 (cf. Petrich and Reilly [14, cf. Lemma III.2.11])**.** $(\mathsf{CS})_{\mathsf{E}} = \mathsf{CS}$.

Proposition 3.16 (cf. Pastijn and Yan [12])**.** $(\mathsf{CR})_{\mathsf{E}} = \mathsf{CR}$.

Using Proposition 3.15 we may establish an equality in Example 3.5, as we see below.

Example 3.17. By Proposition 3.15, property (2) of Lemma 3.1 and since $\mathsf{CS} \subseteq \mathsf{LG}$, it follows that $\mathsf{CS} = (\mathsf{CS})_{\mathsf{E}} \subseteq (\mathsf{LG})_{\mathsf{E}}$. Thus, and by Example 3.5, we have $(\mathsf{LG})_{\mathsf{E}} = \mathsf{CS}$.

In Section 4 we return to these last two results and we show how to prove them using the general embedding of Petrich [13]. In Section 5, we prove that the pseudovarieties R, L and J are fixed points of the $\_{\mathsf{E}}$ operator by a different approach, namely by using implicit operations. While using transformation semigroups the number of idempotent generators of the idempotent-generated semigroup depends on the cardinality of the embedded semigroup, in this method the number of idempotent generators of the idempotent-generated semigroup is controlled by the number of generators of the embedded semigroup. However, for the case $\mathsf{V} = \mathsf{R}$, the first method enables us to show that, in fact, there exists an embedding of an $\mathcal{R}$-trivial semigroup into an idempotent-generated $\mathcal{R}$-trivial semigroup. In the other cases, we just prove a division property. The second method is also used to prove the equality for the pseudovariety DA, while we do not know how to apply the first method.

4. The Petrich embedding into a Rees matrix semigroup

Based on the embedding of Pastijn and Yan [12] (see also Pastijn [11]), Petrich [13] constructed an embedding of a semigroup S into an idempotent-generated semigroup in terms of a Rees matrix semigroup over S^1. This embedding preserves some properties of the initial semigroup and it is this peculiarity that enables us to give an alternative proof of the results from Petrich and Reilly [14, Lemma III.2.11] and Pastijn and Yan [12], respectively, that every semigroup of CS and CR embeds into a finite idempotent-generated semigroup of the same type.

We briefly recall the construction of this embedding. Let S be a semigroup (not necessarily finite). We consider the Rees matrix semigroup $\Phi S = \mathcal{M}(S^1, S^1, \Sigma; Q)$, with $\Sigma = \{\sigma, \tau\}$, where σ and τ are two distinct symbols that are not in S, and $Q = (q_{\alpha s})$ is the sandwich matrix with entries

$$q_{\sigma s} = 1, \quad q_{\tau s} = s \quad (s \in S^1).$$

The mapping

$$\varphi_S : s \to (1, s, \sigma) \quad (s \in S)$$

embeds S into ΦS, although it is not the unique embedding from S into ΦS. Petrich determined the set of idempotents of ΦS, which is

$$E(\Phi S) = \{(s,t,\sigma) \in \Phi S \mid t \in E(S^1)\} \cup \{(s,t,\tau) \in \Phi S \mid t = tst\} \tag{4.1}$$

and described Green's relations on ΦS as follows.

Lemma 4.1 (Petrich [13, Lemma 4.3])**.** *Let (s,t,α), $(u,v,\beta) \in \Phi S$. Then:*

1. *$(s,t,\alpha)\ \mathcal{L}\ (u,v,\beta)$ if and only if $t\ \mathcal{L}\ v$ and $\alpha = \beta$;*
2. *$(s,t,\alpha)\ \mathcal{R}\ (u,v,\beta)$ if and only if $s = u$ and $t\ \mathcal{R}\ v$;*
3. *$(s,t,\alpha)\ \mathcal{H}\ (u,v,\beta)$ if and only if $s = u$, $t\ \mathcal{H}\ v$, and $\alpha = \beta$;*
4. *$(s,t,\alpha)\ \mathcal{D}\ (u,v,\beta)$ if and only if $t\ \mathcal{D}\ v$.*

Thus, ΦS has the same number of $\mathcal{D}$-classes as S^1 and each $\mathcal{D}$-class D' of ΦS, which corresponds to a $\mathcal{D}$-class D of S, has the following number respectively of $\mathcal{L}$-classes and $\mathcal{R}$-classes: $2 \cdot |\mathcal{L}\text{-classes of } D|$ and $|S^1| \cdot |\mathcal{R}\text{-classes of } D|$. Each $\mathcal{H}$-class of ΦS has the same number of elements of the corresponding $\mathcal{H}$-class in S^1.

It is obvious that if S is a finite semigroup, so is ΦS. Note that ΦS is generated by the set of idempotents $\{(s, 1, \sigma) \mid s \in S^1\} \cup \{(1, 1, \tau)\}$, which gives an immediate proof of the result from Howie (see Proposition 3.9). Petrich also showed that this embedding preserves other properties of S (see [13, Theorem 5.4]). In particular, he proves the following corollary.

Corollary 4.2 (cf. Petrich [13, Theorem 5.4])**.** *Every semigroup of* $\bar{\mathsf{H}}$*, where* H *is a pseudovariety of groups, embeds into an idempotent-generated semigroup of* $\bar{\mathsf{H}}$*.*

Choosing some specific subsemigroups of ΦS, we can prove the following results:

Proposition 4.3. *Every semigroup of* CS *embeds into an idempotent-generated semigroup of* CS*.*

Proof. Note that, if $S \in \mathsf{CS}$, then ΦS has at most two $\mathcal{D}$-classes, the one corresponding to D, D', and the other corresponding to the neutral element added to S. If we show that D' is generated by its idempotents, then it suffices to consider the embedding $\varphi'_S : s \mapsto (1, s, \sigma)$ from S into the semigroup D'.

Let $(1, s, \sigma)$, with $s \in S$, be any element of $\varphi_S(S)$ and let $e \in E(S)$ be such that $e \mathrel{\mathcal{H}} s$. Then one can compute $(1, e, \tau) \cdot (s, e, \sigma) = (1, s, \sigma)$ with $(1, e, \tau), (s, e, \sigma) \in E(D')$. Hence, the group $\mathcal{H}$-class $\{(1, s, \sigma) \mid s \in S\}$ is contained in $\langle E(D')\rangle$. Since all $\mathcal{H}$-classes of $\langle E(D')\rangle$ have the same number of elements, we conclude that $\langle E(D')\rangle = D'$. Hence S embeds into $\langle E(D')\rangle$, which is an idempotent-generated completely simple semigroup. □

Proposition 4.4. *Every semigroup of* CR *embeds into an idempotent-generated semigroup of* CR*.*

Proof. Let $S \in \mathsf{CR}$. We want to determine an idempotent-generated completely regular subsemigroup of ΦS where S embeds. Let H be an $\mathcal{H}$-class of S. Since H is a group, then the $\mathcal{H}$-classes of ΦS of the form $\{s\} \times H \times \{\sigma\}$, with $s \in S^1$, and $\{1\} \times H \times \{\tau\}$ are groups. Let t be any element of S and H_t be the $\mathcal{H}$-class of S containing this element. We observe that the $\mathcal{H}$-classes of the form $\{s\} \times H_t \times \{\tau\}$, with $s \geq_{\mathcal{J}} t$, are groups. Let $e \in E(S)$ be such that $e \mathrel{\mathcal{H}} t$. Then there exist $x, y \in S$ such that $e = xsy = exsye$. Hence $e \mathrel{\mathcal{L}} sye$ and, since $L_e \cap R_{sye}$ is a group, we have that $esye \mathrel{\mathcal{H}} e$ and $es \mathrel{\mathcal{R}} e$. By Green's Lemma, it follows that

$$\begin{aligned} \mu_s : H_t &\to H_{ts} \\ u &\mapsto us \end{aligned}$$

is a bijection. Let $v \in H_t$ be such that vs is the idempotent of H_{ts}. Since $vsv \mathrel{\mathcal{H}} v$ and $\mu_s(vsv) = vsvs = vs = \mu_s(v)$, then $vsv = v$ and (s, v, τ) is the idempotent of $\{s\} \times H_t \times \{\tau\}$.

Now, we observe that if an $\mathcal{H}$-class of ΦS of the form $\{s\} \times H_t \times \{\tau\}$ is a group, then $s \geq_{\mathcal{J}} t$. Thus, if there exists $u \in H_t$ such that (s, u, τ) is an idempotent, then $usu = u$, and so $s \geq_{\mathcal{J}} u \mathrel{\mathcal{H}} t$. So, we have identified all maximal subgroups of ΦS.

We consider the subsemigroup T of ΦS generated by the following idempotents:

$$\begin{aligned} &(s, t, \sigma) \ \text{ with } t \in E(S^1) \text{ and } s \geq_{\mathcal{J}} t; \\ &(s, t, \tau) \ \text{ with } t = tst. \end{aligned}$$

Basically, we chose all idempotents of the $\mathcal{R}$-classes whose $\mathcal{H}$-classes are groups. Thus, the product of any two idempotents on a same $\mathcal{D}$-class is also on this $\mathcal{D}$-class and, specifically, on a $\mathcal{H}$-class containing an idempotent of the set of generators. Let us see what is the product of two idempotents of the set of generators that are not in the same $\mathcal{D}$-class. Let (s,t,α) and (u,v,β) be two such idempotents. We have $(s,t,\alpha)\cdot(u,v,\beta) = (s,tq_{\alpha u}v,\beta)$. As $s \geq_{\mathcal{J}} t$, then $s \geq_{\mathcal{J}} tq_{\alpha u}v$. Hence this product is in an $\mathcal{H}$-class that contains an idempotent of the set of generators of T.

Note that the $\mathcal{H}$-classes of the form $\{1\}\times H\times\{\sigma\}$, where H is an $\mathcal{H}$-class of S^1 are in T. In fact, given $a \in S^1$, we have $(1,a,\sigma) = (1,e,\tau)\cdot(a,e,\sigma)$, where $e \in E(S^1)$ is such that $e\ \mathcal{H}\ a$, and $(1,e,\tau)$ and (a,e,σ) are idempotents of T. It follows that T is the subsemigroup of ΦS consisting of the $\mathcal{R}$-classes of ΦS whose $\mathcal{H}$-classes are groups.

Hence T is a completely regular semigroup and $\varphi''_S : s \mapsto (1,s,\sigma)$ is an embedding of S into T. □

In the above proof, we may reduce the choice of the idempotents and we may consider the subsemigroup of T generated by the following idempotents:

$$(1,e,\tau),\ (a,e,\sigma)\ \text{ with } a \in S \text{ and } e \in E(S) \text{ such that } e\mathcal{H}a.$$

This subsemigroup is also a union of $\mathcal{R}$-classes of ΦS whose $\mathcal{H}$-classes are groups and the $\mathcal{H}$-classes of the form $\{1\}\times H\times\{\sigma\}$, where H is an $\mathcal{H}$-class of S^1, are also in this subsemigroup. However, to simplify the proof, we considered the subsemigroup T.

Example 4.5. Consider the completely regular semigroup

$$S = \langle a,b,c,d \mid a^3 = a, b^2 = b, c^3 = c, d^2 = d, ab = ba, cb = bc, ada = a,$$
$$ac = ca = bd = db = cd = dc = 0\rangle.$$

We present in Figure 1 a $\mathcal{D}$-class of ΦS to illustrate the distribution of the idempotents. We also observe that T consists of the $\mathcal{R}$-classes of ΦS whose $\mathcal{H}$-classes are groups, as we have mentioned previously.

Corollary 4.6. *The pseudovarieties* $\bar{\mathsf{H}}$*,* CS *and* CR *satisfy the equality* $\mathsf{V}_{\mathsf{E}} = \mathsf{V}$.

When we work with the pseudovarieties DS and DA, and since the regular $\mathcal{D}$-classes of the semigroups of these pseudovarieties are completely simple semigroups, one may ask whether the construction used in Proposition 4.4 may lead to a proof of existence of an embedding from every semigroup of any of these pseudovarieties into an idempotent-generated semigroup of the same pseudovariety. However, in the following example, we observe that this is not the case.

Example 4.7. Consider the semigroup $S = \langle a \mid a^3 = a^2\rangle$. We look at the subsemigroup T of ΦS generated by the idempotents of the same type as those of Proposition 4.4 (see Figure 2). Note that neither the element $(1,a,\sigma)$ nor any element of the $\mathcal{D}$-class D_a is in T. We have to choose the idempotent $(a,1,\sigma)$ to be a generator of T, but, in that case, T is no longer an element of DS (and, consequently, of DA).

We end this subsection with no answer for the question: *Does* $\mathsf{V}_{\mathsf{E}} = \mathsf{V}$ *for any of the pseudovarieties* DS *or* DA*?* In the following section we see that, in fact, the pseudovariety DA satisfies such equality.

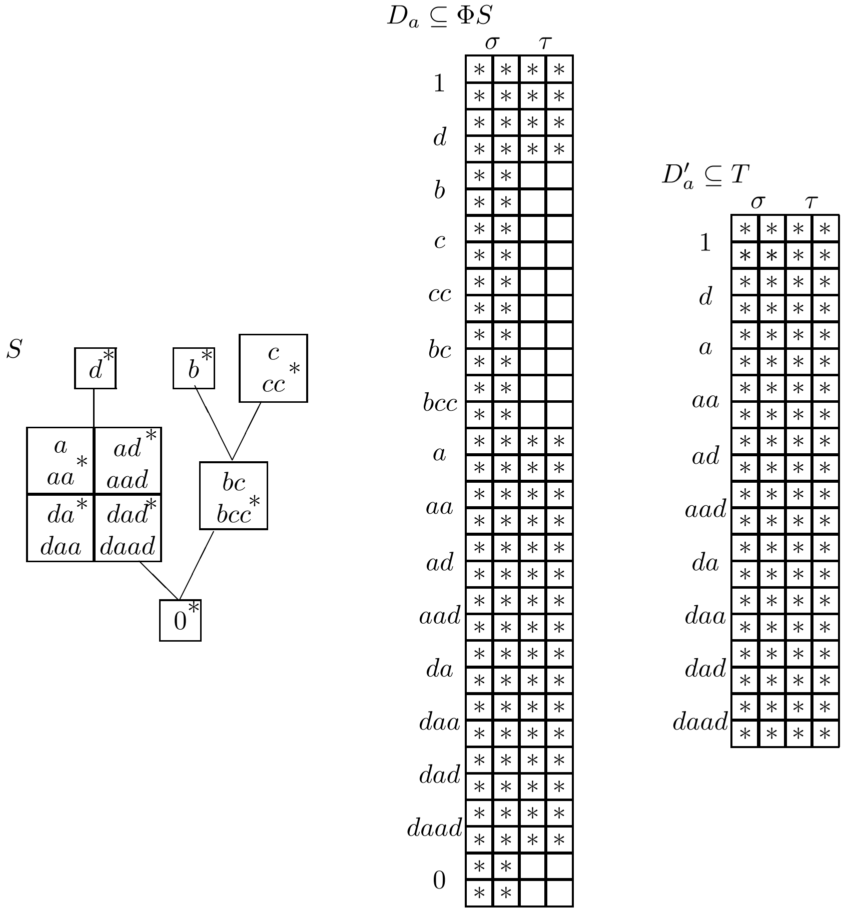


FIGURE 1. The Petrich embedding for a completely regular semigroup

5. REPRESENTATIONS BY IMPLICIT OPERATIONS

We refer the reader to [2] for detailed information about profinite semigroups and to standard references for the basics of topology. By an *embedding* of topological semigroups we mean a semigroup homomorphism that is simultaneously a homeomorphism with the image subspace. A *clopen* subset of a topological space is one that is simultaneously closed and open.

Theorem 5.1. *Let* V *be a pseudovariety such that, for every* n*, there exists* m *such that* $\overline{\Omega}_n\mathsf{V}$ *embeds in* $\overline{\langle X\rangle}$ *for some* $X \subseteq E(\overline{\Omega}_m\mathsf{V})$*. Then* $\mathsf{V}_{\mathsf{E}} = \mathsf{V}$*.*

Proof. Let V be a pseudovariety satisfying the above conditions. Let $S \in \mathsf{V}$ and $\varphi : \overline{\Omega}_n\mathsf{V} \longrightarrow S$ be a continuous surjective homomorphism. Let $\mu_{\mathsf{V}} : \overline{\Omega}_n\mathsf{V} \to \overline{\langle X\rangle} \subseteq$

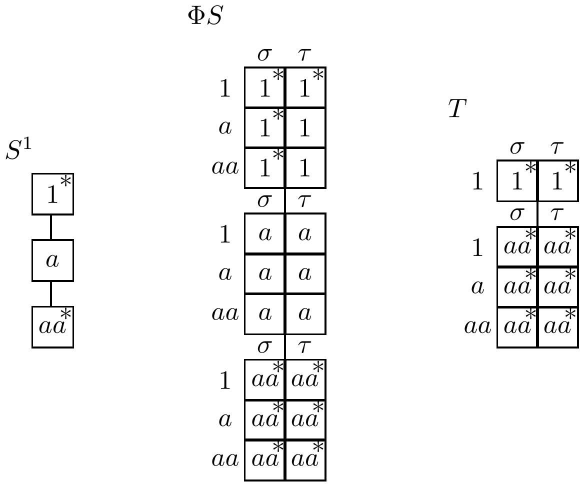


FIGURE 2. The Petrich embedding for a monogenic monoid

$\overline{\Omega}_m\mathsf{V}$ be an embedding, with $X \subseteq E(\overline{\Omega}_m\mathsf{V})$. We consider the following diagram:

$$\begin{tikzcd}
\overline{\Omega}_n\mathsf{V} \arrow[r, "\mu_{\mathsf{V}}"] \arrow[d, "\varphi"'] & T \arrow[r, hook] \arrow[d, "\phi_{|T}"] & \overline{\Omega}_m\mathsf{V} \arrow[d, "\phi"] \\
S \in \mathsf{V} & \phi(T) \subseteq \langle E(F)\rangle & F \in \mathsf{V}
\end{tikzcd}$$

where T is the image of μ_{V}.

We claim that there is a family of clopen subsets $(U_s)_{s\in S}$ of $\overline{\Omega}_m\mathsf{V}$, pairwise disjoint, such that $U_s \cap T = \mu_{\mathsf{V}}(\varphi^{-1}(s))$. We proceed to prove the claim. For each $s \in S$, let $A_s = \mu_{\mathsf{V}}(\varphi^{-1}(s))$ and $A'_s = T\backslash A_s$. Since $\{s\}$ is a clopen subset of S, then $\varphi^{-1}(s)$ is a clopen subset of $\overline{\Omega}_n\mathsf{V}$ and $\mu_{\mathsf{V}}(\varphi^{-1}(s))$ is a clopen subset of T. Since A_s and A'_s are closed sets of the closed subspace T of $\overline{\Omega}_m\mathsf{V}$, A_s and A'_s are closed subsets of $\overline{\Omega}_m\mathsf{V}$. Moreover, since $\overline{\Omega}_m\mathsf{V}$ is compact and 0-dimensional, then A_s and A'_s are separated by two disjoint clopen sets V_s and V'_s, respectively. We choose an arbitrary ordering for the elements of S: $s_1, \dots, s_{|S|}$. Let $U_{s_1} = V_{s_1}$ and recursively, for $i = 2, \dots, |S|$, let $U_{s_i} = V_{s_i} \setminus (\bigcup_{j<i} U_{s_j})$. Note that, for every i, U_{s_i} is also a clopen subset of $\overline{\Omega}_m\mathsf{V}$ and $(U_s)_{s\in S}$ is a family satisfying the claim.

Since $\overline{\Omega}_m\mathsf{V}$ is a pro-V semigroup and, for all $s \in S$, U_s is a clopen subset, there exists a continuous homomorphism $\phi_s : \overline{\Omega}_m\mathsf{V} \to F_s$ with $F_s \in \mathsf{V}$ such that $U_s = \phi_s^{-1}(\phi_s(U_s))$ (cf. [1, 2]). Let $\phi : \overline{\Omega}_m\mathsf{V} \to F = \prod_s F_s$ be the continuous homomorphism such that $\phi = (\phi_s)_{s\in S}$. Then $U_s = \phi^{-1}(\phi(U_s))$ for all $s \in S$. We consider the diagram

$$\begin{tikzcd}
\overline{\Omega}_n\mathsf{V} \arrow[d, "\varphi"'] \arrow[dr, "\phi\circ\mu_{\mathsf{V}}"] & & \\
S & \phi(T) \arrow[l, dashed, "\rho"] \arrow[r, hook] & \langle E(F)\rangle.
\end{tikzcd}$$

We show that S is a homomorphic image of $\phi(T)$, more precisely that, there exists $\rho : \phi(T) \to S$ such that the diagram commutes. It suffices to show that, for $w, z \in \overline{\Omega}_n\mathsf{V}$, if $(\phi \circ \mu_{\mathsf{V}})(w) = (\phi \circ \mu_{\mathsf{V}})(z)$, then $\varphi(w) = \varphi(z)$. Let $s_1 = \varphi(w)$, $s_2 = \varphi(z)$ and suppose that $s_1 \neq s_2$. Since $U_{s_1} \cap U_{s_2} = \emptyset$, we have that $\phi(U_{s_1}) \cap \phi(U_{s_2}) = \emptyset$. Now, $\mu_{\mathsf{V}}(w) \in \mu_{\mathsf{V}}(\varphi^{-1}(s_1)) = U_{s_1} \cap T$ and, therefore, $\phi(\mu_{\mathsf{V}}(w)) \in \phi(U_{s_1} \cap T)$. Similarly, we obtain $\phi(\mu_{\mathsf{V}}(z)) \in \phi(U_{s_2} \cap T)$. It follows that $\phi(\mu_{\mathsf{V}}(w)) \neq \phi(\mu_{\mathsf{V}}(z))$.

We conclude that S divides $\langle E(F)\rangle$. Since $F \in \mathsf{V}$, it follows that $\langle E(F)\rangle \in \mathsf{V}_{\mathsf{E}}$ and, therefore, $S \in \mathsf{V}_{\mathsf{E}}$. This shows that $\mathsf{V} \subseteq \mathsf{V}_{\mathsf{E}}$, while the reverse inclusion is always verified. □

From Theorem 5.1, to conclude that $\mathsf{V}_{\mathsf{E}} = \mathsf{V}$, it suffices to exhibit an embedding $\mu_{\mathsf{V}} : \overline{\Omega}_n\mathsf{V} \to \overline{\langle X\rangle}$ with $X \subseteq E(\overline{\Omega}_m\mathsf{V})$, for every integer n. We do not know if, conversely, such an embedding always exists in case $\mathsf{V}_{\mathsf{E}} = \mathsf{V}$.

For $\mathsf{V} \in \{\mathsf{R}, \mathsf{L}, \mathsf{J}, \mathsf{DA}\}$, we consider the unique continuous homomorphism μ_{V} such that

$$\begin{array}{rccc} \mu_{\mathsf{V}} : & \overline{\Omega}_n\mathsf{V} & \to & \overline{\Omega}_{n+1}\mathsf{V} \\ & x_i & \mapsto & x_i^{\omega}y^{\omega}, \end{array}$$

where y is a new variable and we prove that μ_{V} is an embedding. In each case, we depend heavily on a suitable representation of the profinite semigroup $\overline{\Omega}_n\mathsf{V}$.

Let us start with the pseudovariety R. We use the representation of implicit operations over R by means of labeled ordinals due to Almeida and Weil [3]. We briefly recall it. Let A be an alphabet with $|A| = n$ and let $\mathbf{rLO}(A)$ be the set of reduced A-labeled ordinals. Recall that an *A-labeled ordinal* is a pair (α, l), where α is an ordinal and $l : \alpha \to A$ is a labeling function. The *content* of (α, l), $c(\alpha, l)$, is the range of l. The *cumulative content* of a limit ordinal $\beta \leq \alpha$, $\overleftarrow{c}(\beta)$, is the set of all letters $a \in A$ such there exists a sequence $(\gamma_k)_k$ of ordinals with $\cup_k \gamma_k = \beta$, $\gamma_k < \beta$ and $l(\gamma_k) = a$ for all k. An A-labeled ordinal (α, l) is said to be *reduced* if $l(\beta) \notin \overleftarrow{c}(\beta)$ for each limit ordinal $\beta < \alpha$.

Let $(\alpha, l) \in \mathbf{rLO}(A)$. For each $a \in A$, let γ_a be the smallest ordinal such that $\gamma_a < \alpha$ and $l(\gamma_a) = a$ (i.e., γ_a is the position of the first occurrence of a). We set $\gamma_a = 0$ if $l(\gamma) \neq a$, for all $\gamma < \alpha$. Let $\alpha_1 = \max\{\gamma_a \mid a \in A\}$ (i.e., the first occurrence of the last appearing letter) and let β_1 be such that $\alpha = \alpha_1 + 1 + \beta_1$, with $(\alpha_1, l_1), (\beta_1, m_1) \in \mathbf{rLO}(A)$, $l_1 = l_{|\alpha_1}$ and $m_1(\gamma) = l(\alpha_1 + 1 + \gamma)$ where $\gamma < \beta_1$. We call the equality $\alpha = \alpha_1 + 1 + \beta_1$ the *left basic partition* of (α, l). We iterate this process while $\beta_i \neq 0$. Let $\beta_0 = \alpha$, $m_0 = l$ and $\beta_i = \alpha_{i+1} + 1 + \beta_{i+1}$ with α_{i+1} and β_{i+1} constructed in the same way. While $\beta_i \neq 0$, we obtain ordinals (α_i, l_i) and (β_i, m_i), with $i \geq 1$, where $l_{i+1} = m_{i_{|\alpha_{i+1}}}$ and $m_{i+1}(\gamma) = m_i(\alpha_{i+1} + 1 + \gamma)$ where $\gamma < \beta_{i+1}$. Almeida and Weil showed that $\alpha = \sum_{i \geq 1}(\alpha_i + 1)$ and they define the product in $\mathbf{rLO}(A)$ that follows. Let $(\alpha, l), (\beta, m) \in \mathbf{rLO}(A)$. If α is not a limit ordinal, then

$$(\alpha, l)(\beta, m) = (\alpha + \beta, l') \tag{5.1}$$

where $l'(\gamma) = l(\gamma)$ if $\gamma < \alpha$ and $l'(\alpha + \gamma) = m(\gamma)$ if $\gamma < \beta$. If α is a limit ordinal, then we write $\beta = \beta_1 + \beta_2$ where β_1 is the smallest ordinal such that $m(\beta_1) \notin \overleftarrow{c}(\alpha)$. The product is given by

$$(\alpha, l)(\beta, m) = (\alpha + \beta_2, l') \tag{5.2}$$

where $l'(\gamma) = l(\gamma)$ if $\gamma < \alpha$ and $l'(\alpha + \gamma) = m(\beta_1 + \gamma)$ if $\gamma < \beta_2$. Almeida and Weil proved that $\mathbf{rLO}(A)$ equipped with this operation is isomorphic to $\overline{\Omega}_n\mathsf{R}$.

Proposition 5.2. *The function $\mu_{\mathsf{R}} : \overline{\Omega}_n\mathsf{R} \to \overline{\Omega}_{n+1}\mathsf{R}$ is an embedding.*

Proof. We denote by $\psi_A : \overline{\Omega}_n\mathsf{R} \to \mathbf{rLO}(A)$ the isomorphism defined by Almeida and Weil [3], where $|A| = n$. Let $B = \{a, b : a \in A\}$ with $b \notin A$. We consider the following diagram

$$\begin{tikzcd}
\overline{\Omega}_n\mathsf{R} \arrow[r, "\mu_{\mathsf{R}}"] \arrow[d, "\psi_A"'] & \overline{\Omega}_{n+1}\mathsf{R} \arrow[d, "\psi_B"] \\
\mathbf{rLO}(A) \arrow[r, dashed, "\nu"'] & \mathbf{LO}(B)
\end{tikzcd}$$

with ν defined as follows:

$$\begin{aligned}
\nu : \mathbf{rLO}(A) &\to \mathbf{LO}(B) \\
(\alpha, l) &\mapsto ((\omega+\omega)\alpha, l')
\end{aligned}$$

where $\mathbf{LO}(B)$ is the set of B-labeled ordinals and

$$\begin{aligned}
l' : (\omega+\omega)\alpha &\to B \\
\beta &\mapsto \begin{cases} l(\gamma) & \text{if } \beta = (\omega+\omega)\gamma + k \text{ with } \gamma < \alpha,\ k \in \omega \\ b & \text{if } \beta = (\omega+\omega)\gamma + \omega + k \text{ with } \gamma < \alpha,\ k \in \omega. \end{cases}
\end{aligned}$$

We prove that the diagram commutes, i.e., that $\nu = \psi_B \circ \mu_{\mathsf{R}} \circ \psi_A^{-1}$. Let $(\alpha = \sum_{i\geq 1}(\alpha_i + 1), l) \in \mathbf{rLO}(A)$. Since ψ_A^{-1}, ψ_B and μ_{R} are homomorphisms, we proceed by induction on $|c(\alpha_i, l_i)|$, which is finite and less than $|c(\alpha, l)|$, and we obtain:

$$\begin{tikzcd}
\prod_{i\geq 1}(\psi_A^{-1}(\alpha_i, l_i)a_i) \arrow[r, mapsto, "\mu_{\mathsf{R}}"] & \prod_{i\geq 1}\big(\mu_{\mathsf{R}}(\psi_A^{-1}(\alpha_i, l_i))a_i^{\omega}b^{\omega}\big) \arrow[d, mapsto, "\psi_B"] \\
(\alpha = \sum_{i\geq 1}(\alpha_i + 1), l) \arrow[u, mapsto, "\psi_A^{-1}"] & (\delta, m) = \sum_{i\geq 1}(((\omega+\omega)\alpha_i, l_i') + (\omega+\omega, g_i))
\end{tikzcd}$$

where $a_i = m_{i-1}(\alpha_i)$ and

$$\begin{aligned}
g_i : \omega + \omega &\to B \\
\beta &\mapsto \begin{cases} a_i & \text{if } \beta < \omega \\ b & \text{if } \beta = \omega + \gamma \text{ with } \gamma < \omega. \end{cases}
\end{aligned}$$

We want to prove that $(\delta, m) = \nu(\alpha, l) = ((\omega+\omega)\alpha, l')$. Indeed, we have

$$\begin{aligned}
\delta &= \sum_{i\geq 1}((\omega+\omega)\alpha_i + (\omega+\omega)) \\
&= (\omega+\omega)\alpha_1 + (\omega+\omega) + (\omega+\omega)\alpha_2 + (\omega+\omega) + \cdots \\
&= (\omega+\omega)(\alpha_1 + 1 + \alpha_2 + 1 + \cdots) \\
&= (\omega+\omega)\sum_{i\geq 1}(\alpha_i + 1) \\
&= (\omega+\omega)\alpha,
\end{aligned}$$

where the third equality follows from [17, Exercise 1.41], and

$$m : (\omega+\omega)\alpha \to B$$

$$\beta \mapsto \begin{cases} l'_i(\gamma) & \text{if } \beta = (\omega+\omega)(\sum_{j=1}^{i-1}(\alpha_j+1)) + \gamma \\ & \quad \text{with } \gamma < (\omega+\omega)\alpha_i \\ g_i(\gamma) & \text{if } \beta = (\omega+\omega)(\sum_{j=1}^{i-1}(\alpha_j+1)) + (\omega+\omega)\alpha_i + \gamma \\ & \quad \text{with } \gamma < (\omega+\omega), \end{cases}$$

where we set $\sum_{j=1}^{0}(\alpha_j+1) = 0$. In the first case, it follows that

$$\begin{aligned} m(\beta) &= l'_i(\gamma) \\ &= \begin{cases} l'_i(\delta) & \text{if } \gamma = (\omega+\omega)\delta + k, \text{ with } \delta < \alpha_i, k \in w \\ b & \text{if } \gamma = (\omega+\omega)\delta + \omega + k, \text{ with } \delta < \alpha_i, k \in w \end{cases} \\ &= \begin{cases} l(\alpha_1 + 1 + \cdots + \alpha_{i-1} + 1 + \delta) & \text{if } \gamma = (\omega+\omega)\delta + k \\ b & \text{if } \gamma = (\omega+\omega)\delta + \omega + k \end{cases} \\ &= \begin{cases} l(\eta) & \text{if } \beta = (\omega+\omega)\eta + k \\ b & \text{if } \beta = (\omega+\omega)\eta + \omega + k \end{cases} \\ &= l'(\beta) \end{aligned}$$

where $\eta = \sum_{j=1}^{i-1}(\alpha_j+1) + \delta$. In the second case, for $\eta = \sum_{j=1}^{i-1}(\alpha_j+1) + \alpha_i$, we have

$$\begin{aligned} m(\beta) &= g_i(\gamma) \\ &= \begin{cases} a_i & \text{if } \gamma < w \\ b & \text{if } \gamma = w + \delta \text{ with } \delta < w \end{cases} \\ &= \begin{cases} l(\alpha_1 + 1 + \cdots + \alpha_{i-1} + 1 + \alpha_i) & \text{if } \beta = (\omega+\omega)\eta + \delta \\ b & \text{if } \beta = (\omega+\omega)\eta + \omega + \delta \end{cases} \\ &= \begin{cases} l(\eta) & \text{if } \beta = (\omega+\omega)\eta + \delta \\ b & \text{if } \beta = (\omega+\omega)\eta + \omega + \delta \end{cases} \\ &= l'(\beta) \end{aligned}$$

and, therefore, $m = l'$. It follows that the diagram commutes and ν is a homomorphism from $\mathbf{rLO}(A)$ into $\mathbf{rLO}(B)$, where the product involved is the one defined by formulas (5.1) and (5.2). Thus, μ_{R} is injective if and only if ν is injective. Let (α, l) and (β, m) be reduced labeled ordinals such that $\nu(\alpha, l) = \nu(\beta, m)$. By [17, Exercise 3.41], we have

$$(\omega+\omega)\alpha = (\omega+\omega)\beta \Longrightarrow \alpha = \beta$$

and

$$l' = m' \Longrightarrow l(\gamma) = m(\gamma) \text{ for all } \gamma < \alpha \Longrightarrow l = m.$$

Hence $(\alpha, l) = (\beta, m)$ and ν is injective. □

The dual result for the pseudovariety L follows by duality.

Proposition 5.3. *The function $\mu_{\mathsf{L}} : \overline{\Omega}_n\mathsf{L} \to \overline{\Omega}_{n+1}\mathsf{L}$ is an embedding.* □

Now, we consider the pseudovariety J of $\mathcal{J}$-trivial semigroups. We use the representation by canonical form of implicit operations over J obtained by the first author [1, Section 8.2]. Consider the variety $\mathcal{V}$ of type $(2,1)$ defined by the set of identities

$$\Sigma = \{(xy)z = x(yz), (xy)^\omega = (yx)^\omega = (x^\omega y^\omega)^\omega, x^\omega x = x^\omega = xx^\omega, (x^\omega)^\omega = x^\omega\}.$$

We may reduce any term in the variables $x_1, x_2, \ldots$ using the following Noetherian and confluent system of *reduction rules*:

- **(rr1):** to eliminate parentheses concerning the application of the operation of multiplication;
- **(rr2):** to substitute any subterm of the form t^ω by u^ω, where u is the product, in increasing order of the indices, of the variables occurring in t;
- **(rr3):** to absorb in factors of the form u^ω any adjacent factors in which only occur variables of u.

A term of $\mathcal{V}$ is called a *word* if it does not involve the unary operation $\_^\omega$, and it is called *idempotent* if it is of the form t^ω, for some term t. The *content* $c(t)$ of a term t is the set of variables occurring in t. The factorization in *canonical form* of a term t is $t = t_1 \cdots t_n$, where:

- **(cf1):** each t_i is a word or an idempotent;
- **(cf2):** each idempotent t_i is of the form u^ω, where u is a product of variables with the indices in strictly increasing order;
- **(cf3):** given two consecutive idempotents t_i and t_{i+1}, the sets $c(t_i)$ and $c(t_{i+1})$ are incomparable;
- **(cf4):** two consecutive terms t_i and t_{i+1} are not both words;
- **(cf5):** if t_i is a word and t_{i+1} is an idempotent, then the last letter of t_i is not in $c(t_{i+1})$;
- **(cf6):** if t_{i+1} is a word and t_i is an idempotent, then the first letter of t_{i+1} is not in $c(t_i)$.

Let $F_n\mathcal{V}$ be the $\mathcal{V}$-free algebra on $\{x_1, \ldots, x_n\}$. The semigroup $\overline{\Omega}_n\mathsf{J}$ may be seen as an algebra of type $(2,1)$, where all elements are constructed using the operations of multiplication and omega power and the variables $\{x_1, \ldots, x_n\}$. Then we have a natural surjective homomorphism

$$\begin{aligned} \psi : F_n\mathcal{V} &\to \overline{\Omega}_n\mathsf{J} \\ x_i &\mapsto x_i \quad (i = 1, \ldots, n), \end{aligned}$$

and [1, Theorem 8.2.7] establishes that ψ is, in fact, an isomorphism. We are now able to prove the desired proposition.

Proposition 5.4. *The function* $\mu_\mathsf{J} : \overline{\Omega}_n\mathsf{J} \to \overline{\Omega}_{n+1}\mathsf{J}$ *is an embedding.*

Proof. By the above, to show that μ_J is injective is equivalent to establishing that

$$\begin{aligned} \nu : F_n\mathcal{V} &\to F_{n+1}\mathcal{V} \\ x_i &\mapsto x_i^\omega y^\omega \end{aligned}$$

is injective. Let $w, z \in F_n\mathcal{V}$ be such that $\nu(w) = \nu(z)$ and let $w = w_1 \cdots w_m$ and $z = z_1 \cdots z_n$ be the respective factorizations in canonical form. We want to determine the factorizations in canonical form of $\nu(w)$ and $\nu(z)$. Let $i \in \{1, \ldots, m\}$. Suppose that w_i is a word, i.e., $w_i = x_{i_1} \cdots x_{i_k}$. Then $\nu(w_i) = x_{i_1}^\omega y^\omega \cdots x_{i_k}^\omega y^\omega$. Note that this factorization is in canonical form, because it is a product such that two consecutive idempotents have incomparable contents. Suppose now that w_i is an idempotent, i.e., $w_i = (x_{i_1} \cdots x_{i_l})^\omega$ with $i_1 < \cdots < i_l$. Then $\nu(w_i) = (x_{i_1}^\omega y^\omega \cdots x_{i_l}^\omega y^\omega)^\omega = (x_{i_1} \cdots x_{i_l} y)^\omega$ applying the reduction rule (rr2). Note that $i_1 < \cdots < i_l < y$ (assuming that the new letter y is larger than any of the others) and, therefore, the last presented factorization of $\nu(w_i)$ is in canonical form. Therefore, an idempotent of $F_n\mathcal{V}$ has as image an idempotent of $F_{n+1}\mathcal{V}$ and a word of length k has as image a product of $2k$ idempotents of $F_{n+1}\mathcal{V}$, in canonical form.

Consider now the product w_iw_j with $j = i + 1$. Note that, by definition of canonical form, w_i and w_j are not both words. Suppose that w_i is a word and w_j is an idempotent. Then

$$\nu(w_iw_j) = x_{i_1}^\omega y^\omega \cdots x_{i_k}^\omega y^\omega \cdot (x_{j_1} \cdots x_{j_l} y)^\omega = x_{i_1}^\omega y^\omega \cdots x_{i_k}^\omega (x_{j_1} \cdots x_{j_l} y)^\omega$$

applying the reduction rule (rr3). By hypothesis $x_{i_k} \notin c(w_j)$ and we conclude that the last factorization of $\nu(w_iw_j)$ is in canonical form. If w_i is an idempotent and w_j is a word, or if both w_i and w_j are idempotents, then we have, respectively, the following canonical forms for w_iw_j:

$$\nu(w_iw_j) = (x_{i_1} \cdots x_{i_k} y)^\omega x_{j_1}^\omega y^\omega \cdots x_{j_l}^\omega y^\omega$$

and

$$\nu(w_iw_j) = (x_{i_1} \cdots x_{i_k} y)^\omega (x_{j_1} \cdots x_{j_l} y)^\omega.$$

Let $\nu(w) = \bar{w}_1 \cdots \bar{w}_{m'}$ and $\nu(z) = \bar{z}_1 \cdots \bar{z}_{n'}$ be the factorizations in canonical form of $\nu(w)$ and $\nu(z)$, respectively. Since $\nu(w) = \nu(z)$, by [1, Theorem 8.2.8] we have $m' = n'$ and $\bar{w}_i = \bar{z}_i$, for all i. Three cases can occur for each factor $\bar{w}_i$: $\bar{w}_i = x_j^\omega$, $\bar{w}_i = y^\omega$ or $\bar{w}_i = (x_{j_1} \cdots x_{j_l} y)^\omega$, for some j, l. Note that the content of the idempotent in the last case has cardinal bigger than 1, whether in the other cases is 1. We recover w as follows. In the first two cases, we substitute $\bar{w}_i$ by x_j and by 1, respectively. In the last case, we substitute $\bar{w}_i$ by $(x_{j_1} \cdots x_{j_l})^\omega$. It is easy to see that the canonical forms of w and z are recovered and they are equal. It follows that $w = z$ and ν is injective. $\square$

Finally, we treat the case of DA using the representation of implicit operations over DA by means of labeled orderings obtained by Moura [9], which is similar to the case of the pseudovariety R. So, we omit most details and we refer the reader to [9] as needed. In that paper, we proved that there is a bijection between the free profinite semigroup over DA, $\overline{\Omega}_A\mathsf{DA}$, and the set of all reduced A-labeled $*$-linear orderings, $\mathbf{rLO}^*(A)$.

Proposition 5.5. *The function $\mu_{\mathsf{DA}} : \overline{\Omega}_n\mathsf{DA} \to \overline{\Omega}_{n+1}\mathsf{DA}$ is an embedding.*

Proof. Since $\overline{\Omega}_A\mathsf{DA}$ and $\mathbf{rLO}^*(A)$ are isomorphic, it suffices to prove that the following mapping is injective:

$$\begin{aligned} \nu : \mathbf{rLO}^*(A) &\to \mathbf{rLO}^*(B) \\ (o, l) &\mapsto ((\omega + \omega^*)2o, l'). \end{aligned}$$

By [9], $\nu(o, l)$ is constructed from (o, l) in the following way: each position of o is replaced by the ordering $(\omega + \omega^*)2$ and, if this position is labeled $a \in A$, the label of each position on the resulting ordering is a or b, depending on whether the position is in the first or second term of the form $\omega + \omega^*$. Thus, given two consecutive positions $p < q$ from $\nu(o, l)$, one and only one of the following cases can occur: $l'(p) = l'(q) = a$, $l'(p) = l'(q) = b$, $l'(p) = a$ and $l'(q) = b$, or $l'(p) = b$ and $l'(q) = a$, for some $a \in A$. In the first three cases, both positions are in the same interval $(\omega+\omega^*)2$ of o, resulting from the replacement of a position of o labeled a, for some $a \in A$. In the fourth case, the positions are in consecutive intervals corresponding to the replacement of consecutive positions of o. We split $(\omega+\omega^*)2o$ in intervals I_p, $p \in o$, of the form $(\omega+\omega^*)2$ and that are maximal for the following condition: $I_p = I_{p,1} \cup I_{p,2}$ where $I_{p,1}$ is an interval whose elements are labeled with the same letter of A and $I_{p,2}$ is an interval whose elements are labeled with b. It follows that $(\omega+\omega^*)2o = \bigcup_{p\in o} I_p$ and $l'(\bar{p}) = l(p)$, for all $\bar{p} \in I_{p,1}$, with $I_p = I_{p,1} \cup I_{p,2}$ satisfying the above condition. Thus, we may recover (o,l) from $((\omega+\omega^*)2o, l')$ considering the ordering of such intervals with the labeling function that labels each interval by a if the set of labels of its positions is $\{a, b\}$. □

Combining Propositions 5.2, 5.3, 5.4, and 5.5 with Theorem 5.1, we obtain the following result, of which only the case of DA was not already proved by the alternative approach referred in Section 3.

Corollary 5.6. *The pseudovarieties* R, L, J *and* DA *satisfy the equation* $\mathsf{V_E} = \mathsf{V}$.

It remains an open problem whether the pseudovariety DS satisfies the equality $\mathsf{V_E} = \mathsf{V}$. This motivates the study of the free profinite semigroup over DS, for which no representation result is currently known.

6. Ranks

For a pseudovariety V we consider the following associated parameters:

- $\operatorname{rank}\mathsf{V}$ is the least positive integer n such that V is defined by a set of pseudoidentities on at most n variables, unless there is no such n, in which case we let $\operatorname{rank}\mathsf{V} = \infty$;
- for a positive integer n, $\mathsf{V}(n)$ denotes the pseudovariety generated by all n-generated members of V, that is the class of all finite continuous homomorphic images of $\overline{\Omega}_n\mathsf{V}$;
- the *generator rank* of V, denoted $\operatorname{grank}\mathsf{V}$, is the least positive integer n such that $\mathsf{V} = \mathsf{V}(n)$, unless there is no such n, in which case we set $\operatorname{grank}\mathsf{V} = \infty$;
- the *idempotent generator rank* of V, denoted $\operatorname{idgrank}\mathsf{V}$, is the least positive integer n such that V is generated by its members which are generated by at most n idempotents, unless there is no such n, in which case we put $\operatorname{idgrank}\mathsf{V} = \infty$.

The following are simple observations concerning these parameters.

Lemma 6.1. *Let* V *be a pseudovariety and* n *a positive integer. Denote by* Σ_n *the set of all pseudoidentities in at most* n *variables which are valid in* V. *Then the following hold:*

(1) $\operatorname{rank}\mathsf{V} \le n$ *if and only if* $\mathsf{V} = [\![\Sigma_n]\!]$;
(2) $\operatorname{grank}\mathsf{V} \le n$ *if and only if* $\mathsf{V} = \mathsf{V}(n)$;
(3) $\operatorname{idgrank}\mathsf{V} \le n$ *implies* $\mathsf{V} = \mathsf{V_E}(n)$;
(4) $\operatorname{grank}\mathsf{V} \le \operatorname{idgrank}\mathsf{V}$.

Lemma 6.2. *Let* $x_1, \ldots, x_n$ *be* n *distinct variables and consider the word* $u_i = x_1 \cdots x_{i-1}x_{i+1} \cdots x_n$. *Then the pseudoidentity*

$$(u_1^\omega \cdots u_n^\omega)^{\omega+1} = u_1^\omega \cdots u_n^\omega \tag{6.1}$$

holds in $\mathsf{DS}(n-1)$ *but not in* $\mathsf{J}(n)$.

Proof. Ordering the variables by increasing order of the indices, the canonical form of the implicit operation over J determined by the left side of the pseudoidentity (6.1) is $(x_1 \cdots x_n)^\omega$ while the right side is already in canonical form. By [1, Theorem 8.2.7] it follows that $\overline{\Omega}_n\mathsf{J}$ fails (6.1), whence so does $\mathsf{J}(n)$.

Let $\varphi : \overline{\Omega}_n\mathsf{S} \to \overline{\Omega}_{n-1}\mathsf{DS}$ be any continuous homomorphism. We need to show that φ identifies the two sides of (6.1), that is that $\varphi(u_1^\omega \cdots u_n^\omega)$ is regular. Now, by [1, Theorems 8.1.10 and 8.2.7], $\varphi(u_1^\omega \cdots u_n^\omega)$ is regular if and only if it has the same content as some $\varphi(u_i^\omega)$. Since $c\big(\varphi(u_1^\omega \cdots u_n^\omega)\big)$ contains at most $n-1$ variables, there is an index $i \in \{1, \ldots, n\}$ such that $c\big(\varphi(u_1^\omega \cdots u_n^\omega)\big) = \bigcup_{j\in\{1,\ldots,n\}\setminus\{i\}} c\big(\varphi(x_j)\big)$, whence $c\big(\varphi(u_1^\omega \cdots u_n^\omega)\big) = c\big(\varphi(u_i^\omega)\big)$. □

The following result is an immediate application of Lemma 6.2.

Proposition 6.3. *Every pseudovariety in the interval* $[\mathsf{J}, \mathsf{DS}]$ *has infinite grank.* □

In contrast, the pseudoidentity definitions of the pseudovarieties $\mathsf{J}, \mathsf{R}, \mathsf{L}, \mathsf{DA}, \mathsf{DG}, \mathsf{DO}, \mathsf{DS}$ given at the end of Section 2 show that the rank of each of them is two. Indeed, the smallest pseudovarieties defined by one-variable pseudoidentities containing the first four, respectively the last three, of these are respectively A and S.

There are pseudovarieties whose generator rank is smaller than its rank. As an example, we consider the pseudovariety $\mathsf{V} = \mathsf{J}(2)$. It is obvious that $\mathsf{V} = \mathsf{V}(2)$ and so the generator rank of V is two (note that, for example, the semigroup $\langle e, f \mid e^2 = e, f^2 = f, fe = 0\rangle$ is in $\mathsf{J}(2)\setminus\mathsf{J}(1)$). Since, by Proposition 6.3, J has infinite generator rank, it follows that $\mathsf{J}(2) \subsetneq \mathsf{J}$. Now, by Lemma 6.4 below, we conclude that $\mathsf{J}(2)$ has rank bigger than two.

Lemma 6.4. *Let* Σ_2 *be the set of all pseudoidentities in at most two variables which are valid in* $\mathsf{J}(2)$*. Then* $[\![\Sigma_2]\!] = \mathsf{J}$.

Proof. Let $u, v \in \overline{\Omega}_2\mathsf{S}$ be such that the pseudoidentity $u = v$ holds in $\mathsf{J}(2)$. For every semigroup $S \in \mathsf{J}$ and every continuous homomorphism $\varphi : \overline{\Omega}_2\mathsf{S} \to S$, the elements $\varphi(u)$ and $\varphi(v)$ belong to a two-generated subsemigroup of S, which in turn is in $\mathsf{J}(2)$. Thus we have the equality of $\varphi(u)$ and $\varphi(v)$ and so the pseudoidentity $u = v$ holds in J.

For the direct inclusion, it suffices to note that, if $u = v$ is a pseudoidentity in at most two variables which is valid in J, then it is obviously valid in $\mathsf{J}(2)$, and so it belongs to Σ_2. Since $\operatorname{rank} \mathsf{J} = 2$, it follows that $[\![\Sigma_2]\!] \subseteq \mathsf{J}$. □

At this point, we do not know what is the rank of the pseudovariety $\mathsf{J}(2)$ not even if it is finite. Of course, if $\mathsf{J}(2)$ is finitely based, then it has finite rank; but we also do not know if that is the case.

For the idgrank, we can prove the following results.

Lemma 6.5. *We define, recursively, two sequences of implicit operations as follows: for* $n \geq 3$*, we put*

$$\begin{aligned}
v_3 &= (x_1x_2)^\omega x_3 \cdot x_1^2 \cdot x_3(x_1x_2)^\omega \\
w_3 &= (x_1x_2)^\omega x_3 \cdot x_1 \cdot x_3(x_1x_2)^\omega \\
v_{n+1} &= (x_1 \cdots x_n)^\omega x_{n+1} \cdot v_n \cdot x_{n+1}(x_1 \cdots x_n)^\omega \\
w_{n+1} &= (x_1 \cdots x_n)^\omega x_{n+1} \cdot w_n \cdot x_{n+1}(x_1 \cdots x_n)^\omega.
\end{aligned}$$

(a) If S is a semigroup from DA *generated by $n \geq 2$ idempotents, then S satisfies the pseudoidentity $v_{n+1} = w_{n+1}$.*
(b) The pseudovariety $\mathsf{J}(n)$ *fails the pseudoidentity $v_{n+1} = w_{n+1}$.*

Proof. (a) By hypothesis, there exists some continuous homomorphism $\pi : \overline{\Omega}_n\mathsf{DA} \to S$ that maps the free generators x_i to idempotents.

We proceed by induction on n. Given $\varphi : \overline{\Omega}_{n+1}\mathsf{S} \to S$, we must show that $\varphi(v_{n+1}) = \varphi(w_{n+1})$. Since π is onto and $\overline{\Omega}_n\mathsf{S}$ is free profinite, φ factors through π, say as $\varphi = \pi \circ \psi$ for some continuous homomorphism $\psi : \overline{\Omega}_{n+1}\mathsf{S} \to \overline{\Omega}_n\mathsf{DA}$.

At the basis of the induction, let us consider first the case $n = 2$. If $\psi(x_1x_2)$ has full content, then $\varphi\big((x_1x_2)^\omega\big)$ belongs to the minimum ideal of S. Since this ideal is a rectangular band and both $\varphi(v_3)$ and $\varphi(w_3)$ are $\mathcal{L}$ and $\mathcal{R}$-below $\varphi\big((x_1x_2)^\omega\big)$, they are both equal to this idempotent. Otherwise, $\psi(x_1x_2)$ only involves one of the free generators of $\overline{\Omega}_n\mathsf{DA}$ and so $\varphi(x_1)$ is an idempotent, in which case the equality $\varphi(v_3) = \varphi(w_3)$ is trivially verified.

For the general case $n > 2$, similarly, if $\psi(x_1 \cdots x_n)$ has full content, then $\varphi(v_{n+1}) = \varphi(w_{n+1})$. Otherwise, let $T = \overline{\langle x_1, \ldots, x_n\rangle}$ so that $\varphi(T)$ is a semigroup of DA generated by at most $n-1$ idempotents. By induction hypothesis, it satisfies the pseudoidentity $v_n = w_n$, whence $\varphi(v_n) = \varphi(w_n)$. Taking into account the definition of v_{n+1} and w_{n+1}, we conclude that $\varphi(v_{n+1}) = \varphi(w_{n+1})$.

(b) Let $\varphi = \overline{\Omega}_{n+1}\mathsf{S} \to \overline{\Omega}_n\mathsf{J}$ be the continuous homomorphism that fixes x_1 and sends each other x_i to x_{i-1}. Consider the factorizations of v_{n+1} and w_{n+1} in idempotents and maximal explicit factors between them which results from the recursive definition of these implicit operations. Then a straightforward induction shows that both these factorizations and the result of applying φ to each factor (and eliminating the repetition of x_1 within each ω-power) are in canonical form. Hence $\varphi(v_{n+1}) \neq \varphi(w_{n+1})$ by [1, Theorem 8.2.7]. Since $\overline{\Omega}_n\mathsf{J}$ is residually finite, this shows that there is some member of $\mathsf{J}(n)$ that fails $v_{n+1} = w_{n+1}$. □

In view of the definitions, Lemma 6.5 yields the following result.

Proposition 6.6. *The inequality* $\operatorname{idgrank}\mathsf{V}(n) > n$ *holds for every pseudovariety* V *in the interval* $[\mathsf{J}, \mathsf{DA}]$. □

Combining Propositions 6.6, 5.4, 5.2, 5.3, and 5.5 yields the following result.

Corollary 6.7. *The equality* $\operatorname{idgrank}\mathsf{V}(n) = n+1$ *holds for every pseudovariety* V *in* $\{\mathsf{J}, \mathsf{R}, \mathsf{L}, \mathsf{DA}\}$. □

We finish the paper with a brief comparison of the results obtained here for the equality $\mathsf{V} = \mathsf{V}_\mathsf{E}$, with $\mathsf{V} \in \{\mathsf{J}, \mathsf{R}, \mathsf{L}\}$, and the results that follow from the work of Straubing [18] and Howie [7], and of Pin [15] and Laradji and Umar [8], respectively.

Straubing showed that an n-element $\mathcal{J}$-trivial monoid divides $\mathcal{C}_{n+1}$, and Pin proved that a finite $\mathcal{R}$-trivial monoid with cardinal n embeds into $\mathcal{E}_n$.

On the other hand, several works deal with the ranks and idempotent ranks of various finite transformation semigroups. Recall that the *rank* of a finite semigroup is the minimum number of generators, and the *idempotent rank* of an idempotent-generated finite semigroup is the minimum number of idempotent generators. Gomes and Howie [5] showed that the rank and idempotent rank of the subsemigroup of $\mathcal{T}_n$ consisting of all full transformations with range less than n are both equal to $n(n-1)/2$. The rank and idempotent rank of the subsemigroup of all contractive finite full transformations are both equal to $n+1$, as showed by Umar [19]. Finally, Laradji and Umar [8] proved that the rank and idempotent rank of the subsemigroup of all contractive and order-preserving finite full transformations are both equal to n. We refer the reader to [20, 8] for detailed information on this topic.

Therefore, in the results quoted in Section 3 that follow from the works of these authors, the idempotent rank of the idempotent-generated semigroup is related with the cardinality of the embedded semigroup. In this section together with Section 5, we proved that any finite semigroup of J, R, L or DA with rank n divides an idempotent-generated semigroup of the same type with idempotent rank at most $n+1$ (cf. Corollary 6.7). So, here the control on the number of generators is related with the number of generators of the embedded semigroup, which may be much smaller than its cardinality.

Acknowledgments. The work of both authors was supported by *Fundação para a Ciência e a Tecnologia (FCT)* through the *Centro de Matemática da Universidade do Porto (CMUP)* and also through the project PTDC/MAT/65481/2006, which is partly funded by the European Community Fund FEDER. The work of the second author was also supported by the *FCT PhD Grant* SFRH/BD/19720/2004.

References

[1] J. Almeida, *Finite Semigroups and Universal Algebra*, World Scientific, Singapore, 1994. English translation.
[2] ———, *Profinite semigroups and applications*, in Structural Theory of Automata, Semigroups, and Universal Algebra, V. B. Kudryavtsev and I. G. Rosenberg, eds., vol. 207 of NATO Science Series II: Mathematics, Physics and Chemistry, New York, 2005, Springer, 1–45. Proceedings of the NATO Advanced Study Institute on Structural Theory of Automata, Semigroups and Universal Algebra, Montréal, Québec, Canada, 7-18 July 2003.
[3] J. Almeida and P. Weil, *Free profinite* R*-trivial monoids*, Int. J. Algebra and Comput. **7** (1997) 625–671.
[4] S. Eilenberg, *Automata, Languages and Machines*, vol. B, Academic Press, New York, 1976.
[5] G. Gomes and J. Howie, *On the ranks of certain finite semigroups of transformations*, Math. Proc. Cambridge Phil. Soc. **101** (1987) 395–403.
[6] J. M. Howie, *The subsemigroup generated by the idempotents of a full transformation semigroup*, J. London Math. Soc. **41** (1966) 707–716.
[7] ———, *Products of idempotents in certain semigroups of order-preserving transformations*, Proc. Edinburgh Math. Soc. **17** (1971) 223–236.
[8] A. Laradji and A. Umar, *On certain finite semigroups of order-decreasing transformations I*, Semigroup Forum **69** (2004) 184–200.
[9] A. Moura, *Representations of the free profinite object over* DA, Tech. Rep. CMUP 2009-26, Univ. of Porto, 2009. `http://cmup.fc.up.pt/cmup/v2/include/filedb.php?id=275&table=publicacoes&field=file`.
[10] ———, E*-local pseudovarieties*, Tech. Rep. CMUP 2009-45, Univ. of Porto, 2009. `http://cmup.fc.up.pt/cmup/v2/include/filedb.php?id=296&table=publicacoes&field=file`.
[11] F. Pastijn, *Embbeding semigroups in semibands*, Semigroup Forum **14** (1977) 247–263.
[12] F. Pastijn and X. Yan, *Varieties of semigroups and varieties of completely regular semigroups closed for certain extensions*, J. Algebra **163** (1994) 777–794.
[13] M. Petrich, *Embbeding semigroups into idempotent generated ones*, Monatsh. Math. **141** (2004) 315–322.
[14] M. Petrich and N. R. Reilly, *Completely Regular Semigroups*, Canadian Math. Soc. Series of Monographs and Advanced Texts 23, Wiley, New York, 1999.
[15] J.-E. Pin, *Varieties of Formal Languages*, Plenum, London, 1986. English translation.
[16] J. Reiterman, *The Birkhoff theorem for finite algebras*, Algebra Universalis **14** (1982) 1–10.
[17] J. G. Rosenstein, *Linear Orderings*, Academic Press, New York, 1982.
[18] H. Straubing, *On finite J-trivial monoids*, Semigroup Forum **19** (1980) 107–110.

[19] A. Umar, *On the semigroups of order-decreasing finite full transformations*, Proc. Roy. Soc. Edinburgh Sect. A **120** (1992) 129–142.

[20] ———, *On the ranks of certain finite semigroups of order-decreasing transformations*, PortugaliæMath. **53** (1996) 23–32.

(J. Almeida) Centro de Matemática e Departamento de Matemática Pura, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

E-mail address: `jalmeida@fc.up.pt`

(A. Moura) Instituto Superior de Engenharia do Porto/LEMA and Centro de Matemática da Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

E-mail address: `aim@isep.ipp.pt/amoura@fc.up.pt`

Paper 4

E-LOCAL PSEUDOVARIETIES

A. MOURA

Abstract. Generalizing a property of the pseudovariety of all aperiodic semigroups observed by Tilson, we call E-*local* a pseudovariety V which satisfies the following property: for a finite semigroup, the subsemigroup generated by its idempotents belongs to V if and only if so do the subsemigroups generated by the idempotents in each of its regular $\mathcal{D}$-classes. In this paper, we present several sufficient or necessary conditions for a pseudovariety to be E-local or for a pseudoidentity to define an E-local pseudovariety. We also determine several examples of the smallest E-local pseudovariety containing a given pseudovariety.

1. Introduction

The motivation for this work came from an exercise suggested by Pin [5] about a result from Tilson [8]. With the aim of finding a method for computing the complexity of a finite semigroup in terms of the structure of its subsemigroups, Tilson started by establishing a useful method for computing the group-complexity of a finite semigroup with at most two non-zero $\mathcal{D}$-classes. This led him to prove the following result: given a finite semigroup S, the subsemigroup $\langle E(S)\rangle$ is aperiodic if and only if, for every regular $\mathcal{D}$-class D of S, the subsemigroup $\langle E(D)\rangle$ is aperiodic.

As a consequence of the work of Fitz-Gerald [4], we have that a regular semigroup is orthodox if and only if the product of idempotents of every regular $\mathcal{D}$-class of S is idempotent. Thus, it suffices to analyze the property of the product of idempotents to be an idempotent on every regular $\mathcal{D}$-class to conclude the property for an arbitrary product of idempotents.

Much work has been done on the structure of idempotent-generated semigroups. So, it becomes interesting to determine the pseudovarieties V satisfying the following property: given $S \in \mathsf{S}$, $\langle E(S)\rangle \in \mathsf{V}$ if and only if $\langle E(D)\rangle \in \mathsf{V}$, for each regular $\mathcal{D}$-class D of S. We call *E-local* a pseudovariety with this property.

This paper is a contribution towards the complete characterization of E-local pseudovarieties. We start by recalling, in Section 2, some basics of the theory of pseudovarieties of semigroups, in particular, some results concerning the block operator $\mathsf{B}\_$ and the idempotent-generated subsemigroup of a semigroup. Section 3 concerns the study of E-local pseudovarieties: we observe some properties and examples, we present several sufficient conditions for a pseudovariety to be E-local and we show that, in certain cases, these conditions are also necessary, and we introduce a new operator, $\_^{\mathsf{E}}$, where V^{E} is the smallest E-local pseudovariety containing a pseudovariety V. Finally, in Section 4, we present some more necessary or sufficient conditions for a pseudoidentity to define an E-local pseudovariety.

2000 *Mathematics Subject Classification.* 20M07.

Key words and phrases. Finite semigroup; pseudovariety; idempotent-generated; regular D-class; operator; E-local.

2. Preliminaries

We briefly recall some basics of the theory of pseudovarieties of semigroups. We recommend [1, 5, 7] for a better understanding of this area.

Let S be a semigroup. We denote by $E(S)$ the set of idempotents of S and by $\langle E(S)\rangle$ the subsemigroup of S generated by $E(S)$. More generally, $\langle X\rangle$ denotes the subsemigroup of S generated by $X \subseteq S$. In case S is finite, s^ω denotes the unique idempotent in the subsemigroup generated by a given $s \in S$.

Let S be a finite semigroup and let D be a regular $\mathcal{D}$-class of S. Consider the equivalence relation $\sim$ on the set of group elements of a regular $\mathcal{D}$-class D of S defined in the following way: given two group elements a and b of D, $a \sim b$ if and only if there exists an *idempotent-chain* $e_0, e_1, \ldots, e_{n-1}, e_n$ such that $a \mathrel{\mathcal{H}} e_0$, $b \mathrel{\mathcal{H}} e_n$, and either $e_i \mathrel{\mathcal{R}} e_{i-1}$ or $e_i \mathrel{\mathcal{L}} e_{i-1}$, for all $i \in \{1, \ldots, n\}$. A *block of* D is the Rees quotient of the subsemigroup of S generated by a $\sim$-class modulo the ideal consisting of the elements that are not in D. The *blocks of* S are the blocks of its regular $\mathcal{D}$-classes.

A class of finite semigroups that is closed under taking subsemigroups, homomorphic images and finite direct products is called a *pseudovariety* and generally denoted by V. For example, S denotes the pseudovariety of all finite semigroups.

We may construct new pseudovarieties from known ones by applying operators to pseudovarieties. In this paper, we use the following operators on pseudovarieties:

- EV consists of all $S \in \mathsf{S}$ such that $\langle E(S)\rangle \in \mathsf{V}$;
- DV consists of all $S \in \mathsf{S}$ such that, for every regular $\mathcal{D}$-class D of S, $D \in \mathsf{V}$;
- for a pseudovariety H of groups, $\bar{\mathsf{H}}$ consists of all $S \in \mathsf{S}$ such that every subgroup G of S belongs to H;
- BV consists of all $S \in \mathsf{S}$ such that, for every block B of S, $B \in \mathsf{V}$;
- $\mathsf{V_E}$ is the pseudovariety generated by the idempotent-generated semigroups of V.

The last operator was introduced in Almeida and Moura [2] and we refer the reader to that paper as needed, but we opt to present here an easy lemma that will be used frequently in this paper:

Lemma 2.1 (Almeida and Moura [2])**.** *The operator* $\_\mathsf{E}$ *has the following properties, where* V *and* W *are arbitrary pseudovarieties:*

(1) $\mathsf{V} \subseteq \mathsf{W}$ *implies* $\mathsf{V_E} \subseteq \mathsf{W_E}$*;*
(2) $(\mathsf{V} \cap \mathsf{W})_\mathsf{E} \subseteq \mathsf{V_E} \cap \mathsf{W_E}$*;*
(3) $(\mathsf{V_E})_\mathsf{E} = \mathsf{V_E}$*;*
(4) $(\mathsf{EV})_\mathsf{E} = \mathsf{V_E}$*;*
(5) $\mathsf{E}(\mathsf{V_E}) = \mathsf{EV}$.

The main aim of our study is the characterization of the E-local pseudovarieties. For this purpose, we need some results concerning idempotent-generated subsemigroups and blocks of such subsemigroups.

Lemma 2.2. *For every pseudovariety* V, $\mathsf{BBV} = \mathsf{BV}$.

Lemma 2.3. *Let* $S \in \mathsf{S}$ *and* $X \subseteq E(S)$*. Then* $\langle E\langle X\rangle\rangle = \langle X\rangle$.

To prove that the idempotent-generated subsemigroup of a regular semigroup is also regular, Fitz-Gerald [4] uses a technique that consists in writing a product of idempotents of $\langle E(S)\rangle$ as a product of idempotents of $\langle E(D)\rangle$, for a regular $\mathcal{D}$-class

D of S. As a consequence, we have the following lemma whose statement and proof may be found in [7, Lemma 4.13.1], for example. It enables us to easily conclude the statement presented in the introduction that a regular semigroup S is orthodox if and only if the product of idempotents of every regular $\mathcal{D}$-class of S is idempotent.

Lemma 2.4. *Let S be a semigroup and let $s \in \langle E(S)\rangle$ be an element of a regular $\mathcal{J}$-class J of S. Then, there exists an idempotent-chain $e_1, e_2, \ldots, e_m \in E(J)$ such that $s = e_1e_2\cdots e_m$. Hence $\langle E(S)\rangle \cap J = \langle E(J)\rangle \cap J$.*

Corollary 2.5. *Every finite semigroup S has the following properties:*

(1) *Let a be a regular element of $\langle E(S)\rangle$. Then a is in a block of D_a, where D_a is the regular $\mathcal{D}$-class of S containing a.*
(2) *Let B be a block of $\langle E(S)\rangle$. Then B is also a block of $\langle E(D)\rangle$, for some regular $\mathcal{D}$-class D of S.*
(3) *Given $X \subseteq E(S)$, the regular $\mathcal{D}$-classes of $\langle X\rangle$ have only one block.*

Proof. (1) and (2) follow immediately from Lemma 2.4 and from the definition of block of S. Now, by Lemma 2.3 and by (1), we have that every regular element of $\langle E\langle X\rangle\rangle = \langle X\rangle$ is in a block of $\langle E\langle X\rangle\rangle = \langle X\rangle$ and we have (3). $\square$

Lemma 2.4 also enables us to prove that the product of the operators $\mathsf{E}\_$ and $\mathsf{B}\_$ is an idempotent operator, as we see in the following corollary.

Corollary 2.6. *For every pseudovariety* V, $\mathsf{EBEBV} = \mathsf{EBV}$.

Proof. Let $S \in \mathsf{EBEBV}$, i.e., for every block B' of $\langle E(B)\rangle$, with B a block of $\langle E(S)\rangle$, $B' \in \mathsf{V}$. By Lemma 2.4, $B = \langle E(B)\rangle$. Moreover, the blocks of $\langle E(B)\rangle = B$ are B and the trivial semigroup, the last one if B is not a $\mathcal{D}$-class. Thus, for every block B of $\langle E(S)\rangle$, $B \in \mathsf{V}$, i.e., $S \in \mathsf{EBV}$. The converse follows from $\mathsf{E}\_$ and $\mathsf{B}\_$ being increasing operators. $\square$

A *pseudoidentity* is a formal equality $u = v$, where $u, v \in \overline{\Omega}_A\mathsf{S}$, the set of A-ary implicit operations. We say that $S \in \mathsf{V}$ satisfies $u = v$, and we write $S \models u = v$, if $u_S = v_S$. Recall that an A-ary operation $u_S : S^A \to S$ has the following property: for every homomorphism $\varphi : S \to T$, with $S, T \in \mathsf{V}$, the following diagram commutes:

$$\begin{tikzcd} S^A \arrow[r, "u_S"] \arrow[d, "\varphi^A"'] & S \arrow[d, "\varphi"] \\ T^A \arrow[r, "u_T"'] & T \end{tikzcd}$$

Reiterman's Theorem [6] says that every pseudovariety is defined by some set of finitary pseudoidentities, in the sense that it is the class of finite semigroups satisfying this set of pseudoidentities. The converse of the theorem is easily verified.

In this paper, we use, in particular, the pseudovarieties that we list below together with some corresponding bases of pseudoidentities defining them:

$\mathsf{I} = [\![x = y]\!]$	trivial semigroups;
$\mathsf{J} = [\![(xy)^\omega x = (xy)^\omega = y(xy)^\omega]\!]$	$\mathcal{J}$-trivial semigroups;
$\mathsf{R} = [\![(xy)^\omega x = (xy)^\omega]\!]$	$\mathcal{R}$-trivial semigroups;
$\mathsf{L} = [\![y(xy)^\omega = (xy)^\omega]\!]$	$\mathcal{L}$-trivial semigroups;

$\mathsf{A} = [\![x^{\omega+1} = x^\omega]\!]$	aperiodic (or $\mathcal{H}$-trivial) semigroups;
$\mathsf{G} = [\![x^\omega = 1]\!]$	groups;
$\mathsf{LG} = [\![(x^\omega y)^\omega x^\omega = x^\omega]\!]$	local groups;
$\mathsf{CR} = [\![x^{\omega+1} = x]\!]$	completely regular semigroups;
$\mathsf{CS} = [\![x^{\omega+1} = x, (xyx)^\omega = x^\omega]\!]$	completely simple semigroups;
$\mathsf{RB} = [\![x^2 = x, xyx = x]\!]$	rectangular bands;
$\mathsf{LZ} = [\![xy = x]\!]$	left-zero semigroups;
$\mathsf{DA} = [\![((xy)^\omega x)^2 = (xy)^\omega x]\!]$	regular $\mathcal{D}$-classes are aperiodic semigroups;
$\mathsf{DG} = [\![(xy)^\omega = (yx)^\omega]\!]$	regular $\mathcal{D}$-classes are groups;
$\mathsf{DO} = [\![(xy)^\omega (yx)^\omega (xy)^\omega = (xy)^\omega]\!]$	regular $\mathcal{D}$-classes are orthodox semigroups;
$\mathsf{DS} = [\![((xy)^\omega x)^{\omega+1} = (xy)^\omega x]\!]$	regular $\mathcal{D}$-classes are semigroups.

3. E-local pseudovarieties

We start this section by observing some properties of E-local pseudovarieties and several examples of pseudovarieties having this property. In particular, we prove, in Example 3.8, the result of Tilson [8] that the pseudovariety A is E-local. After that, we present several sufficient conditions for a pseudovariety to be E-local and we show that these conditions are also necessary in case the pseudovariety is contained in EDS. We finish with the introduction of the operator $\_^{\mathsf{E}}$, where V^{E} denotes the smallest E-local pseudovariety containing V.

3.1. **Properties and examples.** We start by noting that the property of being E-local is preserved under intersection. Next, we relate the E-locality of V, EV and V_{E}.

Lemma 3.1. *Let V be a pseudovariety and let $S \in \mathsf{S}$. The following conditions are equivalent:*

(1) *for every regular $\mathcal{D}$-class D in S, $\langle E(D)\rangle \in \mathsf{V}$;*
(2) *for every regular $\mathcal{D}$-class D in S, $\langle E(D)\rangle \in \mathsf{EV}$;*
(3) *for every regular $\mathcal{D}$-class D in S, $\langle E(D)\rangle \in \mathsf{V}_{\mathsf{E}}$.*

Proof. (3) ⇒ (1) ⇒ (2): This follows immediately from $\mathsf{V}_{\mathsf{E}} \subseteq \mathsf{V} \subseteq \mathsf{EV}$.

(2) ⇒ (3): Note that $\langle E(D)\rangle \in \mathsf{EV}$ if and only if $\langle E\langle E(D)\rangle\rangle = \langle E(D)\rangle \in \mathsf{V}$, by Lemma 2.3. By the same lemma and by the definition of V_{E}, we deduce that $\langle E\langle E(D)\rangle\rangle = \langle E(D)\rangle \in \mathsf{V}_{\mathsf{E}}$. □

Similarly, we may prove the following lemma:

Lemma 3.2. *The following conditions are equivalent for every pseudovariety V and every finite semigroup S:*

(1) *$\langle E(S)\rangle \in \mathsf{V}$;*
(2) *$\langle E(S)\rangle \in \mathsf{EV}$;*
(3) *$\langle E(S)\rangle \in \mathsf{V}_{\mathsf{E}}$.* □

The equivalence of E-locality for the pseudovarieties V, EV, and V_{E} follows directly from the previous lemmas.

Corollary 3.3. *Let V be a pseudovariety. The following conditions are equivalent:*

(1) V *is* E*-local;*
(2) EV *is* E*-local;*
(3) $\mathsf{V_E}$ *is* E*-local.*

The properties of the operator $\_\mathsf{E}$ (see Lemma 2.1) together with the previous corollary enable us to identify intervals consisting of E-local pseudovarieties.

Proposition 3.4. *Let* V *be an* E*-local pseudovariety. Then any pseudovariety* $\mathsf{U} \in [\mathsf{V_E}, \mathsf{EV}]$ *is* E*-local.*

Proof. Applying Lemma 2.1, we obtain $\mathsf{EV} = \mathsf{E}(\mathsf{V_E}) \subseteq \mathsf{EU} \subseteq \mathsf{EV}$. The result now follows from Corollary 3.3. □

In an attempt to identify all E-local pseudovarieties, we start by determining several families of pseudovarieties satisfying this property.

Proposition 3.5. *Let* V *and* H*, with* $\mathsf{H} \subseteq \mathsf{G}$*, be pseudovarieties. Then:*
(1) BV *is* E*-local;*
(2) DV *is* E*-local;*
(3) $\bar{\mathsf{H}}$ *is* E*-local.*

Proof. (1) follows directly from item (2) of Corollary 2.5.

(2) By items (2) and (3) from Corollary 2.5, we have that, for every regular $\mathcal{D}$-class D of $\langle E(S)\rangle$, there exists a regular $\mathcal{D}$-class D' of S such that D is a $\mathcal{D}$-class of $\langle E(D')\rangle$. Let S be a semigroup such that, for every regular $\mathcal{D}$-class D, $\langle E(D)\rangle \in \mathsf{DV}$. Then, every regular $\mathcal{D}$-class of $\langle E(D)\rangle$ is a semigroup in V. It follows that every regular $\mathcal{D}$-class of $\langle E(S)\rangle$ is a semigroup in V.

(3) Let $S \in \mathsf{S}$ be such that, for every regular $\mathcal{D}$-class D, $\langle E(D)\rangle \in \bar{\mathsf{H}}$. Let T be a subgroup of $\langle E(S)\rangle$. By Lemma 2.4, $T \subseteq \langle E(D_T)\rangle$, where D_T is the $\mathcal{D}$-class of S containing T. Hence $T \in \mathsf{H}$ and $\langle E(S)\rangle \in \bar{\mathsf{H}}$. □

Example 3.6. Since $\mathsf{J} = \mathsf{DI}$ (see Pin [5, Proposition III.4.1]), it follows from Proposition 3.5 that J is E-local.

Example 3.7. Since $\mathsf{R} = \mathsf{DLZ}$ (see Pin [5, Proposition III.4.1]), it follows from Proposition 3.5 that R is E-local.

Example 3.8. To conclude the result from Tilson [8], it suffices to note that $\mathsf{A} = \bar{\mathsf{I}}$. So, the conclusion that the pseudovariety is E-local follows immediately from Proposition 3.5.

3.2. **Characterizations.** The properties of the operators $\mathsf{E}\_$ and $\mathsf{B}\_$ are useful to obtain the following sufficient and equivalent conditions for a pseudovariety to be E-local.

Proposition 3.9. *The following conditions are equivalent:*
(1) $\mathsf{EV} = \mathsf{EBEV}$*;*
(2) $\mathsf{EV} = \mathsf{BEV}$*;*
(3) *there exists* W *such that* $\mathsf{EV} = \mathsf{BW}$*;*
(4) *there exists* W *such that* $\mathsf{EV} = \mathsf{EBW}$*;*
(5) *there exists* W *such that* $(\mathsf{EBW})_\mathsf{E} \subseteq \mathsf{V} \subseteq \mathsf{EBW}$*;*
(6) *the interval* $[\mathsf{V_E}, \mathsf{EV}]$ *has a fixed point for the operator* B*;*
(7) $\mathsf{B}(\mathsf{V_E}) \subseteq \mathsf{EV}$*;*
(8) $\mathsf{BV} \subseteq \mathsf{EV}$.

Proof. (1) $\Leftrightarrow$ (2): Since $\mathsf{EV} \subseteq \mathsf{BEV} \subseteq \mathsf{EBEV}$, if $\mathsf{EV} = \mathsf{EBEV}$, then $\mathsf{EV} = \mathsf{BEV}$ and we have the direct implication. The converse follows immediately by applying the idempotent operator E to equality (2).

(2) $\Leftrightarrow$ (3): The direct implication is trivial. Conversely, if $\mathsf{EV} = \mathsf{BW}$ for some W, then applying the idempotent operator B, we obtain $\mathsf{BEV} = \mathsf{BW} = \mathsf{EV}$.

(4) $\Leftrightarrow$ (5): Given W such that $\mathsf{EV} = \mathsf{EBW}$, it follows, by Lemma 2.1, that $(\mathsf{EBW})_{\mathsf{E}} = (\mathsf{EV})_{\mathsf{E}} = \mathsf{V}_{\mathsf{E}} \subseteq \mathsf{V} \subseteq \mathsf{EV} = \mathsf{EBW}$. Conversely, if $(\mathsf{EBW})_{\mathsf{E}} \subseteq \mathsf{V} \subseteq \mathsf{EBW}$ for some W, then applying the increasing operator E, we obtain, by the same lemma, $\mathsf{E}((\mathsf{EBW})_{\mathsf{E}}) = \mathsf{EBW} \subseteq \mathsf{EV} \subseteq \mathsf{EEBW} = \mathsf{EBW}$, i.e., $\mathsf{EV} = \mathsf{EBW}$.

(2) $\Rightarrow$ (6) $\Rightarrow$ (4) $\Rightarrow$ (1): The first implication is trivial. For the second one, applying the increasing operator E to $\mathsf{V}_{\mathsf{E}} \subseteq \mathsf{W} = \mathsf{BW} \subseteq \mathsf{EV}$, we obtain, by Lemma 2.1, the equality $\mathsf{EV} = \mathsf{EBW}$. Finally, if the equality $\mathsf{EV} = \mathsf{EBW}$ holds for some W, then $\mathsf{EBEV} = \mathsf{EBEBW} = \mathsf{EBW} = \mathsf{EV}$, by Corollary 2.6.

(6) $\Leftrightarrow$ (7): Let W be a fixed point of the operator B in $[\mathsf{V}_{\mathsf{E}}, \mathsf{EV}]$. It follows immediately that $\mathsf{B}(\mathsf{V}_{\mathsf{E}}) \subseteq \mathsf{BW} = \mathsf{W} \subseteq \mathsf{EV}$. Conversely, assuming (7) we have $\mathsf{V}_{\mathsf{E}} \subseteq \mathsf{B}(\mathsf{V}_{\mathsf{E}}) \subseteq \mathsf{EV}$ and $\mathsf{B}(\mathsf{V}_{\mathsf{E}})$ is a fixed point for the operator B.

(2) $\Rightarrow$ (8) $\Rightarrow$ (7): For the first implication, assuming (2) and applying the operator B to the inclusion $\mathsf{V} \subseteq \mathsf{EV}$, we obtain $\mathsf{BV} \subseteq \mathsf{BEV} = \mathsf{EV}$. The second implication follows immediately from $\mathsf{B}(\mathsf{V}_{\mathsf{E}}) \subseteq \mathsf{BV} \subseteq \mathsf{EV}$. □

Theorem 3.10. *Let* V *be a pseudovariety satisfying the conditions of Proposition 3.9. Then* V *is* E*-local.*

Proof. It follows immediately from Proposition 3.5, Corollary 3.3 and item (2) from Proposition 3.9. □

It is natural to ask whether the conditions of Proposition 3.9 are also necessary. The answer is affirmative when the pseudovariety is contained in EDS.

Theorem 3.11. *If* $\mathsf{V} \subseteq \mathsf{EDS}$ *is an* E*-local pseudovariety, then items* (1)-(8) *from Proposition 3.9 hold.*

Proof. Let $\mathsf{V} \subseteq \mathsf{EDS}$ be an E-local pseudovariety. We prove that $\mathsf{BEV} \subseteq \mathsf{EV}$, as the converse inclusion is trivial. Let $S \in \mathsf{BEV}$, i.e., for every block B of S, $\langle E(B)\rangle \in \mathsf{V}$. Using the E-locality of V, to prove that $S \in \mathsf{EV}$, i.e., $\langle E(S)\rangle \in \mathsf{V}$, it suffices to show that, for every regular $\mathcal{D}$-class D of S, $\langle E(D)\rangle \in \mathsf{V}$. Let D be a regular $\mathcal{D}$-class of S. Using again the E-locality of V and Lemma 2.3, we prove that, for every regular $\mathcal{D}$-class D' of $\langle E(D)\rangle$, $\langle E(D')\rangle \in \mathsf{V}$. Recall that, by item (3) from Corollary 2.5, D' has a unique block, B', and is itself a semigroup, since $\mathsf{V} \subseteq \mathsf{EDS}$. Hence $D' = B'$, for some block B' of $\langle E(D)\rangle$. Moreover, there exists a block B of S such that $B' \leq B$. Thus $\langle E(D')\rangle = \langle E(B')\rangle \leq \langle E(B)\rangle \in \mathsf{V}$ (in fact, the equality holds). Hence $\langle E(S)\rangle \in \mathsf{V}$ and $S \in \mathsf{EV}$. □

The previous theorem was initially announced without the hypothesis $\mathsf{V} \subseteq \mathsf{EDS}$, but there was a flaw in the proof which was detected by an anonymous referee. The example at the end of the paper shows that an extra hypothesis is needed.

Corollary 3.12. *Let* $\mathsf{V} \subseteq \mathsf{EDS}$ *be such that* $\mathsf{V} = \mathsf{EV}$*. Then* V *is* E*-local if and only if there exists* $\mathsf{W} \subseteq \mathsf{CS}$ *such that* $\mathsf{V} = \mathsf{BW}$*.*

Proof. Suppose that V is E-local. Since $\mathsf{V} = \mathsf{EV}$ and by Theorem 3.11 and item (2) from Proposition 3.9, we have $\mathsf{V} = \mathsf{EV} = \mathsf{BEV} = \mathsf{BV}$. Given $S \in \mathsf{V} = \mathsf{BV}$, we have that every block B of S is such that $B \in \mathsf{V} \subseteq \mathsf{EDS}$ and, therefore, $\langle E(B)\rangle \in \mathsf{DS}$, i.e.,

$\langle E(B)\rangle \in \mathsf{CS}$. Note that $\langle E(B)\rangle$ has the same structure in $\mathcal{R}$-classes and $\mathcal{L}$-classes as B, but it may have less elements in the $\mathcal{H}$-classes. Hence $B \in \mathsf{CS}$. It follows that $B \in \mathsf{V} \cap \mathsf{CS}$ and $S \in \mathsf{B}(\mathsf{V} \cap \mathsf{CS})$. The converse follows from $\mathsf{B}\_$ being an increasing operator and from $\mathsf{V} = \mathsf{BV}$. Thus, we have $\mathsf{V} = \mathsf{BV} = \mathsf{B}(\mathsf{V} \cap \mathsf{CS})$. The converse implication follows directly from Proposition 3.5. □

We suggest, as an easy exercise, the verification of the E-locality of the pseudovarieties J, R, DS and A, for example, using the sufficient conditions of Proposition 3.9.

3.3. **The operator $\_^{\mathsf{E}}$.** Because there are pseudovarieties V which are not E-local, it is natural to consider the smallest E-local pseudovariety containing V, which we denote V^{E}.

In this subsection, we determine some pseudovarieties of the form V^{E}. For that purpose, we also use the operator $\_{\mathsf{E}}$ which is studied in detail in [2].

Proposition 3.13. *Let* $\mathsf{V} \subseteq \mathsf{CS}$ *be such that* $\mathsf{V}_{\mathsf{E}} = \mathsf{V}$. *Then* $\mathsf{V}^{\mathsf{E}} = (\mathsf{DV})_{\mathsf{E}}$.

Proof. Let $S = \langle E(S)\rangle \in \mathsf{DV}$ and let W be any E-local pseudovariety containing V. Then $\langle E(D)\rangle \in \mathsf{V}$, for every regular $\mathcal{D}$-class D of S, since D is in V and $\langle E(D)\rangle \leq D$. Since W is E-local and $\mathsf{V} \subseteq \mathsf{W}$, it follows that $S = \langle E(S)\rangle \in \mathsf{W}$. Hence $(\mathsf{DV})_{\mathsf{E}} \subseteq \mathsf{W}$ and, therefore, $(\mathsf{DV})_{\mathsf{E}} \subseteq \mathsf{V}^{\mathsf{E}}$.

For the direct inclusion, since $\mathsf{V} \subseteq \mathsf{CS}$, we have $\mathsf{V} \subseteq \mathsf{DV}$. As $\_{\mathsf{E}}$ is an increasing operator, it follows that $\mathsf{V} = \mathsf{V}_{\mathsf{E}} \subseteq (\mathsf{DV})_{\mathsf{E}}$. By Proposition 3.5, DV is E-local and so is $(\mathsf{DV})_{\mathsf{E}}$, by Corollary 3.3. This yields the inclusion $\mathsf{V}^{\mathsf{E}} \subseteq (\mathsf{DV})_{\mathsf{E}}$. □

Corollary 3.14. *The class* J *is the smallest* E*-local pseudovariety.*

Proof. Let V be an E-local pseudovariety. Since $\mathsf{I} \subseteq \mathsf{V}$, we have $\mathsf{I}^{\mathsf{E}} \subseteq \mathsf{V}^{\mathsf{E}} = \mathsf{V}$. By Proposition 3.13, we have $\mathsf{I}^{\mathsf{E}} = (\mathsf{DI})_{\mathsf{E}} = \mathsf{J}_{\mathsf{E}} = \mathsf{J}$, where the last equality follows from [2, Corollary 5.6]. Hence $\mathsf{J} \subseteq \mathsf{V}$. By Example 3.6, J is E-local, which establishes the corollary. □

Example 3.15. It follows from Corollary 3.14 that the pseudovarieties LG and CR are not E-local.

Example 3.16. By Proposition 3.13, we conclude that $(\mathsf{RB})^{\mathsf{E}} = (\mathsf{DRB})_{\mathsf{E}} = (\mathsf{DA})_{\mathsf{E}} = \mathsf{DA}$, where the last equality follows by [2, Corollary 5.6], and $(\mathsf{CS})^{\mathsf{E}} = (\mathsf{DCS})_{\mathsf{E}} = (\mathsf{DS})_{\mathsf{E}}$.

Example 3.17. By Example 3.16, we have $(\mathsf{DS})_{\mathsf{E}} = (\mathsf{CS})^{\mathsf{E}} \subseteq (\mathsf{CR})^{\mathsf{E}}$. Conversely, since $\mathsf{CR} \subseteq \mathsf{DS}$, by [2, Proposition 3.16] and Lemma 2.1, we deduce that $\mathsf{CR} = (\mathsf{CR})_{\mathsf{E}} \subseteq (\mathsf{DS})_{\mathsf{E}}$. Note that $(\mathsf{DS})_{\mathsf{E}}$ is E-local, by Proposition 3.5 and Corollary 3.3. Hence $(\mathsf{CR})^{\mathsf{E}} \subseteq (\mathsf{DS})_{\mathsf{E}}$, which establishes the equality $(\mathsf{CR})^{\mathsf{E}} = (\mathsf{DS})_{\mathsf{E}}$.

Example 3.18. Let H be a pseudovariety of groups. By [2, Example 3.7], we have $(\mathsf{DH})_{\mathsf{E}} = \mathsf{J} \subseteq \mathsf{J} \vee \mathsf{H} \subseteq \mathsf{DH}$. Since, by Proposition 3.5, DH is E-local, it follows, by Proposition 3.4, that $\mathsf{J} \vee \mathsf{H}$ is E-local. That $\mathsf{J} \vee \mathsf{H}$ is the smallest E-local pseudovariety containing H, is an immediate consequence from the fact that J is the smallest E-local pseudovariety. Thus $\mathsf{H}^{\mathsf{E}} = \mathsf{J} \vee \mathsf{H}$.

Example 3.19. Let H be a pseudovariety of groups. By Example 3.16, we have $\mathsf{DA} = (\mathsf{RB})^{\mathsf{E}} \subseteq (\mathsf{RB} \vee \mathsf{H})^{\mathsf{E}}$ and, therefore, $\mathsf{DA} \vee \mathsf{H} \subseteq (\mathsf{RB} \vee \mathsf{H})^{\mathsf{E}}$. Now, from [2, Lemma 3.1, Example 3.8, Corollary 4.2], we obtain $(\mathsf{DO} \cap \bar{\mathsf{H}})_{\mathsf{E}} \subseteq (\mathsf{DO})_{\mathsf{E}} \cap \bar{\mathsf{H}}_{\mathsf{E}} =$

$\mathsf{DA} \cap \bar{\mathsf{H}} \subseteq \mathsf{DA}$. Therefore, $(\mathsf{DO} \cap \bar{\mathsf{H}})_{\mathsf{E}} \subseteq \mathsf{DA} \subseteq \mathsf{DA} \vee \mathsf{H} \subseteq \mathsf{DO} \cap \bar{\mathsf{H}}$. As an intersection of E-local pseudovarieties is E-local, the pseudovariety $\mathsf{DO} \cap \bar{\mathsf{H}}$ is E-local. By Proposition 3.4, $\mathsf{DA} \vee \mathsf{H}$ is E-local. Thus $(\mathsf{RB} \vee \mathsf{H})^{\mathsf{E}} = \mathsf{DA} \vee \mathsf{H}$.

Example 3.20. Since $\mathsf{LG} \subseteq \mathsf{DS}$ and DS is E-local by Proposition 3.5, we have $(\mathsf{LG})^{\mathsf{E}} \subseteq \mathsf{DS}$. On the other hand, by Example 3.16 and by [2, Example 3.17], we obtain $(\mathsf{DS})_{\mathsf{E}} = (\mathsf{CS})^{\mathsf{E}} = ((\mathsf{LG})_{\mathsf{E}})^{\mathsf{E}} \subseteq (\mathsf{LG})^{\mathsf{E}}$. Thus the equality $(\mathsf{DS})_{\mathsf{E}} \subseteq (\mathsf{LG})^{\mathsf{E}} \subseteq \mathsf{DS}$ holds.

If we prove that $(\mathsf{DS})_{\mathsf{E}} = \mathsf{DS}$, we will have the equality in the previous example. This provides additional motivation for the calculation of $(\mathsf{DS})_{\mathsf{E}}$ which remains an open problem (see [2]).

We end this subsection by noting that $(\mathsf{V} \cap \mathsf{W})^{\mathsf{E}} \subseteq \mathsf{V}^{\mathsf{E}} \cap \mathsf{W}^{\mathsf{E}}$, for all pseudovarieties V and W. However, we do not know whether equality holds.

4. E-local pseudoidentities

We call E-*local* a pseudoidentity which defines an E-local pseudovariety. Note that a pseudovariety defined by a set of E-local pseudoidentities is E-local, since it is the intersection of the E-local pseudovarieties defined by each pseudoidentity of the set. We do not know whether the converse is valid.

Our results from the previous section yield several E-local pseudoidentities. However, some of the results that we obtained, like some techniques developed allow us to give a different characterization of several types of pseudoidentities with this property.

For $u \in \overline{\Omega}_A\mathsf{S}$, let $\mathrm{first}(u)$ and $\mathrm{last}(u)$ be, respectively, the first and last letters of u. We relate the E-locality of the pseudoidentities of the form $u = v$, where $\mathrm{first}(u) \neq \mathrm{first}(v)$ or $\mathrm{last}(u) \neq \mathrm{last}(v)$, with the condition $\mathsf{V} \subseteq [\![u = v]\!]$, where V is one of the pseudovarieties R, L and J. We also obtain some results concerning the pseudovariety DA.

Proposition 4.1. *The following properties are verified by every pseudoidentity* $u = v$.

(1) *If* $\mathrm{last}(u) \neq \mathrm{last}(v)$ *and* $\mathsf{R} \models u = v$, *then* $u = v$ *is* E-*local.*
(2) *If* $\mathrm{first}(u) \neq \mathrm{first}(v)$ *and* $\mathsf{L} \models u = v$, *then* $u = v$ *is* E-*local.*
(3) *If* $\mathrm{first}(u) \neq \mathrm{first}(v)$, $\mathrm{last}(u) \neq \mathrm{last}(v)$ *and* $\mathsf{J} \models u = v$, *then* $u = v$ *is* E-*local.*

Proof. Let $u = v$ be a pseudoidentity such that $\mathrm{last}(u) \neq \mathrm{last}(v)$ and suppose that $\mathsf{R} \models u = v$. We claim that $\mathsf{R} \subseteq [\![u = v]\!] \subseteq \mathsf{ER}$. So that, by Example 3.7 and by Proposition 3.4, $[\![u = v]\!]$ is E-local. The first inclusion is assumed by hypothesis. To prove the second inclusion, let S be a semigroup satisfying $u = v$ and suppose that $S \notin \mathsf{ER}$, i.e., $\langle E(S) \rangle \notin \mathsf{R}$. Then, by [1, cf. Exercise 5.2.8], there exist two distinct idempotents such that $ef = f$ and $fe = e$. Evaluating the last letter of u by e, the last letter of v by f and the other letters by e or f, we obtain $S \models e = u = v = f$, which is a contradiction.

Similarly, we obtain (2) and (3). □

Since, by [2], the pseudovarieties R, L, J and DA satisfy the equality $\mathsf{V}_{\mathsf{E}} = \mathsf{V}$, it is easy to obtain the following results:

Theorem 4.2. *Let* $u = v$ *be an arbitrary pseudoidentity.*

(1) *If* $\mathrm{first}(u) = \mathrm{first}(v)$ *and* $\mathrm{last}(u) \neq \mathrm{last}(v)$*, then* $u = v$ *is* E*-local if and only if* $\mathsf{R} \models u = v$*.*
(2) *If* $\mathrm{first}(u) \neq \mathrm{first}(v)$ *and* $\mathrm{last}(u) = \mathrm{last}(v)$*, then* $u = v$ *is* E*-local if and only if* $\mathsf{L} \models u = v$*.*
(3) *If* $\mathrm{first}(u) \neq \mathrm{first}(v)$ *and* $\mathrm{last}(u) \neq \mathrm{last}(v)$*, then* $u = v$ *is* E*-local if and only if* $\mathsf{J} \models u = v$*.*
(4) *If* $\mathrm{first}(u) = \mathrm{first}(v)$*,* $\mathrm{last}(u) = \mathrm{last}(v)$ *and* $u = v$ *is* E*-local, then* $\mathsf{DA} \models u = v$*.*

Proof. (1) Let $u = v$ be a pseudovariety such that $\mathrm{first}(u) = \mathrm{first}(v)$ and $\mathrm{last}(u) \neq \mathrm{last}(v)$. Suppose that it is E-local. Since $\mathsf{LZ} \subseteq \mathsf{CS}$ and $(\mathsf{LZ})_{\mathsf{E}} = \mathsf{LZ}$, it follows from Proposition 3.13 that $(\mathsf{LZ})^{\mathsf{E}} = (\mathsf{DLZ})_{\mathsf{E}} = \mathsf{R}_{\mathsf{E}} = \mathsf{R}$. Thus, as the pseudovariety $[\![u = v]\!]$ is E-local and it contains LZ, it also contains $(\mathsf{LZ})^{\mathsf{E}} = \mathsf{R}$. The converse follows from Proposition 4.1. Dually, we obtain (2).

(3) It follows directly from J being the smallest E-local pseudovariety (see Corollary 3.14) and from Proposition 4.1.

(4) In that case, we have $\mathsf{RB} \models u = v$, $\mathsf{RB} \subseteq \mathsf{CS}$ and $(\mathsf{RB})_{\mathsf{E}} = \mathsf{RB}$. By Proposition 3.13, we have $(\mathsf{RB})^{\mathsf{E}} = (\mathsf{DRB})_{\mathsf{E}} = (\mathsf{DA})_{\mathsf{E}} = \mathsf{DA}$. As in (1), we deduce that $\mathsf{DA} = (\mathsf{RB})^{\mathsf{E}} \subseteq [\![u = v]\!]$. □

However, we do not have a characterization of all E-local pseudoidentities of the form $u = v$, with $\mathrm{first}(u) = \mathrm{first}(v)$ and $\mathrm{last}(u) = \mathrm{last}(v)$.

Theorem 4.3 provides another sufficient condition for a pseudoidentity to be E-local that follows from Lemma 2.4.

Theorem 4.3. *Let* $u = v$ *be a pseudoidentity such that* $u, v \in \overline{\langle X\rangle}$*, where all elements of* $X \subseteq \overline{\Omega}_A\mathsf{S}$ *lie in a same regular* $\mathcal{D}$*-class of* $\overline{\Omega}_A\mathsf{S}$*. Then* $u = v$ *is* E*-local.*

Proof. Let S be a finite semigroup and suppose that $\langle E(D)\rangle \models u = v$, for each regular $\mathcal{D}$-class D of S. We want to prove that $\langle E(S)\rangle \models u = v$. Let $\varphi : \overline{\Omega}_A\mathsf{S} \to S$ be a continuous surjective homomorphism such that, for every $x \in X$, $\varphi(x) \in \langle E(S)\rangle$. Since all elements of X lie in a same regular $\mathcal{D}$-class of $\overline{\Omega}_A\mathsf{S}$, then there exists a regular $\mathcal{D}$-class D of S such that $\varphi(x) \in D$, for all $x \in X$. By Lemma 2.4, it follows that $\varphi(x) \in \langle E(D)\rangle$, for all $x \in X$. Since $u, v \in \overline{\langle X\rangle}$, it follows that $\varphi(u), \varphi(v) \in \langle E(D)\rangle$ and, by hypothesis, they are equal. Thus $\langle E(S)\rangle \models u = v$ and $u = v$ is E-local. □

Note that several pseudoidentities considered in this paper are of this form. Specifically, the pseudoidentities that we used in Section 2 to define the pseudovarieties J, R, L, A, DA, DG, DO and DS are all of this form. Another example is the pseudoidentity $(x^\omega y^\omega)^\omega = (y^\omega x^\omega)^\omega$ which defines the pseudovariety BG. As a last example, Almeida and Volkov [3] showed that, if $u_i = v_i$, with $i \in I$, is a basis of pseudoidentities for a pseudovariety of groups H, then $u_i' = v_i'$ is a basis of pseudoidentities for $\bar{\mathsf{H}}$, where u_i' and v_i' result from the substitution of each letter $x_j \in A$ of u_i and v_i by ex_je where e is a fixed idempotent in the minimum ideal of $\overline{\Omega}_A\mathsf{S}$. These pseudoidentities are also of the form of Theorem 4.3.

In the last result, we identify all E-local pseudoidentities with only one variable.

Corollary 4.4. *The* E*-local pseudoidentities in one variable are those of the form* $x^\alpha = x^\beta$*, with both* α *and* β *infinite.*

Proof. If α or β are finite and are not equal, then DA does not satisfy the pseudoidentity $u = v$ since DA contains all finite monogenic aperiodic semigroups. Thus, $u = v$ is not E-local, by item (4) of Theorem 4.2.

On the other hand, if α and β are infinite, then x^α and x^β are in a same group with neutral element x^ω. Thus, by Theorem 4.3, the pseudoidentity is E-local. □

We do not know if every E-local pseudovariety is defined by a set of pseudoidentities satisfying the condition of Theorem 4.3.

We finish the paper going back to Section 3.2, where we asked whether Theorem 3.11 can be generalized to all E-local pseudovarieties. The answer is negative as we see in the following example.

Example 4.5. Consider the pseudovariety $\mathsf{W} = [\![((xy)^\omega x(xy)^\omega)^2 = (xy)^\omega x(xy)^\omega]\!]$, which is E-local by Theorem 4.3. We show that W does not satisfy the condition $\mathsf{BV} \subseteq \mathsf{EV}$ from Proposition 3.9.

Let S be the syntactic semigroup of the regular language $(a^+b^+c^+d^+)^+a^+b^+ \cdot (a^+b^+c^+d^+)^+$, whose *egg-box* picture is presented in Figure 1. The blocks of S are the trivial semigroup and the unique block of the unique non-trivial regular $\mathcal{D}$-class of S, D, and which we denote by B. It is easy to see that $B \in \mathsf{W}$. Note that $(xy)^\omega x$ and $(xy)^\omega$ define implicitly $\mathcal{R}$-equivalent elements, the latter being an idempotent. Now, the product in B of two $\mathcal{R}$-equivalent elements of D with the second one an idempotent is an idempotent in D or is 0. On the other hand, the elements $abcdab$ and $abcd$ are $\mathcal{R}$-equivalent and the second one is an idempotent, but their product in S, $abcdababcd$, is not an idempotent (it belongs to the 0-minimal and non-regular $\mathcal{D}$-class of S). Thus $S = \langle E(S)\rangle \notin \mathsf{W}$, i.e., $S \notin \mathsf{EW}$.

Acknowledgments. The author is grateful to the anonymous referee for warning a flaw in the proof of the initially statement of Theorem 3.11. This work is part of the author's doctoral thesis, written under the supervision of Prof. Jorge Almeida, from whose advice the author has greatly benefited. This work was supported by *Fundação para a Ciência e a Tecnologia (FCT)* through the *PhD Grant* SFRH/BD/19720/2004, through the *Centro de Matemática da Universidade do Porto (CMUP)* and also through the project PTDC/MAT/65481/2006, which is partly funded by the European Community Fund FEDER.

References

[1] J. Almeida, *Finite Semigroups and Universal Algebra*, World Scientific, Singapore, 1994. English translation.

[2] J. Almeida and A. Moura, *Idempotent-generated semigroups and pseudovarieties*, Tech. Rep. CMUP 2009-42, Univ. of Porto, 2009. `http://cmup.fc.up.pt/cmup/v2/include/filedb.php?id=292&table=publicacoes&field=file`.

[3] J. Almeida and M. V. Volkov, *Profinite identities for finite semigroups whose subgroups belong to a given pseudovariety*, J. Algebra and Applications **2** (2003) 137–163.

[4] D. G. Fitz-Gerald, *On inverses of products of idempotents in regular semigrups*, J. Aust. Math. Soc. **13** (1972) 335–337.

[5] J.-E. Pin, *Varieties of Formal Languages*, Plenum, London, 1986. English translation.

[6] J. Reiterman, *The Birkhoff theorem for finite algebras*, Algebra Universalis **14** (1982) 1–10.

[7] J. Rhodes and B. Steinberg, *The q-theory of finite semigroups*, Springer, New York, 2009.

[8] B. Tilson, *Complexity of two J-class semigroups*, Advances Math. **11** (1973) 215–237.

Instituto Superior de Engenharia do Porto/LEMA and Centro de Matemática da Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

E-mail address: `aim@isep.ipp.pt/amoura@fc.up.pt`

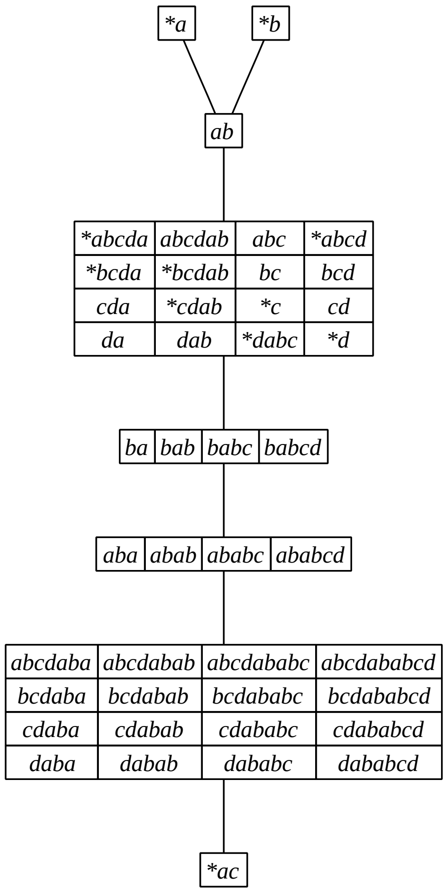


FIGURE 1. The *egg-box* picture of a semigroup in BW and not in EW

Appendix 5

THE COMPLETE PROGRAMMING OF THE DA-AUTOMATON

```python
"""
September 2009, A. Moura

This file contains the complete programming in Python of the algorithm
to compute the minimal DA−automaton presented in url{http://cmup.fc.up.pt/cmup/v2/include/filedb.php?id=276&table=publicacoes&field=file}.

It uses the module FSA−1.0 (url http://osteele.com/software/python/fsa/)
to minimize the automaton and to produce a graphical view with the
psviewer gv. If you are not interested in the visualization, you may
comment the last two lines of this file.
"""

from FSA import *
import time

start = time.time()

#functions#
class Stack:
    def __init__(self):
        self.items=[]
    def push(self,item):
        self.items.append(item)
    def pop(self):
        return self.items.pop()
#
def Parenthesis(w):
    s=Stack()
    P=range(len(w))
    for i in P:
         if w[i]=='(':
             s.push(i)
         elif w[i]==')':
             j=s.pop()
             P[i]=j
             P[j]=i
         else:
             P[i]=-1
    return P
#
def LeftLabel(w):
    s=['(',')']
    for i in range(len(w)):
        char=w[i]
        if char not in s:
            ll=i
            s+=[char]
```

```
    return ll
#
def RightLabel(w):
    s=['(',')']
    for i in range(len(w)):
        j=-1-i
        char=w[j]
        if char not in s:
            rl=j
            s+=[char]
    return len(w)+rl
#
def Factorization(w,ll,rl):
    m=-1
    P=Parenthesis(w)
    for i in range(len(P)):
        if i<ll<P[i] and i<rl<P[i]:
            m=i
            break
    if ll<rl or m!=-1:
        nu=[w,w[ll]+w[rl]]
        desc=[S0forget(w,P,ll),S1remind(w,P,ll,rl,m),S2forget(w,P,rl)]
    elif ll>rl:
        nu=[w,w[rl]+w[ll]]
        desc=[S0remind(w,P,rl),S1forget(w,P,rl,ll),S2remind(w,P,ll)]
    else:
        nu=[w,w[ll]]
        desc=[S0forget(w,P,ll),S2forget(w,P,rl)]
    return [nu,desc]
#
def S0forget(w,P,ll):
    w0=''
    for i in range(ll):
        if w[i]!='(' or P[i]<ll:
            w0+=w[i]
    return w0
#
def S2forget(w,P,rl):
    w2=''
    for i in range(rl+1,len(w)):
        if w[i]!=')' or P[i]>rl:
            w2+=w[i]
    return w2
#
def S1forget(w,P,rl,ll):
    w1=''
    for i in range(rl+1,ll):
        if (w[i]!='(' and w[i]!=')') or \
           (w[i]==')' and P[i]>rl) or \
           (w[i]=='(' and P[i]<ll):
            w1+=w[i]
    return w1
#
def S0remind(w,P,rl):
    w0=''
    for i in range(rl):
        if w[i]!='(' or P[i]<rl:
            w0+=w[i]
```

```
        else:
            for l in range(i,P[i]+1):
                w0+=w[l]
    return w0
#
def S2remind(w,P,ll):
    w2=''
    for i in range (ll+1,len(w)):
        if w[i]!=')' or P[i]>ll:
            w2+=w[i]
        else:
            for l in range(P[i],i+1):
                w2+=w[l]
    return w2
#
def S1remind(w,P,ll,rl,m):
    w1=''
    if m==-1:
        for i in range(ll+1,rl):
            if (w[i]!='(' and w[i]!=')') or \
               (w[i]==')' and P[i]>ll) or \
               (w[i]=='(' and P[i]<rl):
                w1+=w[i]
            elif w[i]==')' and P[i]<ll:
                for l in range(P[i],i+1):
                    w1+=w[l]
            else:
                for l in range (i,P[i]+1):
                    w1+=w[l]
    else:
        M=P[m]
        for i in range(ll+1,M):
             if w[i]!=')' or P[i]>ll:
                 w1+=w[i]
             else:
                 for l in range(P[i],i+1):
                     w1+=w[l]
        for i in range(m,M+1):
            w1+=w[i]
        for i in range(m+1,rl):
            if w[i]!='(' or P[i]<rl:
                w1+=w[i]
            else:
                 for l in range(i,P[i]+1):
                     w1+=w[l]
    return w1
#
def DAautomaton(input):
    e=''
    iota=input
    V=[e,input]
    E=[]
    nu=[[e,'epsilon']]
    V0=[]
    V1=[input]
    while V1!=[]:
        for w in V1:
            ll=LeftLabel(w)
```

```
            rl=RightLabel(w)
            F=Factorization(w,ll,rl)
            nu=nu+[F[0]]
            desc=F[1]
            if len(desc)==3:
                for j in range(3):
                    E+=[[w,j,desc[j]]]
                    if desc[j] not in V:
                        V+=[desc[j]]
                        V0+=[desc[j]]
            else:
                E+=[[w,0,desc[0]]]
                E+=[[w,2,desc[1]]]
                for i in range(2):
                    if desc[i] not in V:
                        V+=[desc[i]]
                        V0+=[desc[i]]
        V1=V0
        V0=[]
    A=[V,E,iota,e,nu]
    return A


#input#
input=raw_input("What is the w-word?")
#input="(a(b)ca(a))"

Gw=DAautomaton(input)

elapsed = (time.time() - start)
print "number of states of Gw=",len(Gw[0])
print "time =",elapsed,"sec."


#FSA#
def DAautomaton2fsa(A):
    V=A[0]
    E=A[1]
    nu=A[4]
    G=[]
    l=-1
    for i in range(len(E)):
        l=E[i][1]
        if l==0:
            G+=[[E[i][0],E[i][2],'0']]
        elif l==1:
            G+=[[E[i][0],E[i][2],'1']]
        else:
            G+=[[E[i][0],E[i][2],'2']]
    g=[]
    for i in range(len(V)):
        for j in range(len(G)):
            if G[j][0]==V[i]:
                G[j][0]=i
                G[j][2]+=nu[i][1]
            if G[j][1]==V[i]:
                G[j][1]=i
    v=range(len(V))
```

```
    for i in range(len(G)):
        g+=[(G[i][0],G[i][1],G[i][2])]
    a=FSA(v,[],g,v[1],[v[0]])
    amin=a.minimized()
    return [a,amin]


f=DAautomaton2fsa(Gw)
a=f[0]
amin=f[1]
a.view()
amin.view()
```

Algorithm 1