**Master's Thesis**

# A Kleene Theorem for Weighted Forest Automata

Frederic Dörband
August 6, 2019, Dresden

Technische Universität Dresden
Professur für Grundlagen der Programmierung

Supervisor: Prof. Dr.-Ing. habil. Heiko Vogler

# Aufgabenstellung

## „A Kleene Theorem for Weighted Forest Automata“

Technische Universität Dresden
Fakultät Informatik

| | |
|---|---|
| Student: | Frederic Dörband |
| Geburtsdatum: | 25. April 1995 |
| Matrikelnummer: | 3947968 |
| Studiengang: | Master Informatik |
| Immatrikulationsjahr: | 2017 |
| Modul: | Masterarbeit |
| Studienleistung: | Masterarbeit |
| Umfang: | 22 Wochen, 29 CP |
| Beginn am: | 24. Juni 2019 |
| Einzureichen am: | 25. November 2019 |
| Verantw. Hochschullehrer/ Betreuer: | Prof. Dr.-Ing. habil. Heiko Vogler |

Das sogenannte *Kleene*-Theorem für Wortsprachen ist ein zentrales Resultat in der theoretischen Informatik. Es beschreibt den Zusammenhang zwischen (endlichen) Automaten und rationalen Ausdrücken [Kle56]. Eine natürliche Verallgemeinerung des *Kleene*-Theorems befasst sich mit *erkennbaren Baumsprachen* [TW68]. Weiterhin wurden sowohl für Wortsprachen als auch für Baumsprachen Verallgemeinerungen des *Kleene*-Theorems für die jeweiligen *gewichteten* Automaten und *gewichteten* rationalen Ausdrücke bewiesen ([Sch61] für Wortsprachen und [DPV05] für Baumsprachen).

Etwa 30 Jahre nach [Kle56], kam zudem die Theorie der Wälder (auch: Magmoide) und Waldsprachen auf [AD78], [AD79]. Eine Verallgemeinerung des *Kleene*-Theorems für Waldsprachen wurde bereits bewiesen [Str09]. Aktuell bleibt eine offene Frage, ob ein *Kleene*-Resultat für gewichtete Waldsprachen formuliert und bewiesen werden kann. In Anlehnung an [Sch61] und [DPV05] stellt sich diese Frage besonders für den Fall von Semiring-gewichteten Waldsprachen.

Die Aufgaben des Studenten im Rahmen der Masterarbeit sind folgende:

- Er soll eine Definition für gewichtete Waldautomaten (WTA) finden und nutzen, um eine Beschreibung der Klasse der erkennbaren gewichteten Waldsprachen (LWTA) zu erhalten. Diese Beschreibung soll auch einen Vergleich zu der Klasse der erkennbaren gewichteten Baumsprachen umfassen.
- Er soll rationale gewichtete Waldsprachen definieren. Hierbei sollen nach Möglichkeit die klassischen rationalen Operationen (endliche Sprachen, Summe, Konkatenation, Kleene-Stern) auf den Waldfall gehoben werden.
- Er soll zeigen, dass LWTA unter den eingeführten rationalen Operationen abschließt.
- Er soll zeigen, dass jede erkennbare gewichtete Waldsprache auch rational ist. Dies soll durch die Analyse von WTA zu rationalen Waldausdrücken geschehen.

Die Arbeit muss den üblichen Standards wie folgt genügen. Die Arbeit muss in sich abgeschlossen sein und alle nötigen Definitionen und Referenzen enthalten. Die Urheberschaft von Inhalten – auch die eigene – muss klar erkennbar sein. Fremde Inhalte, z.B. Algorithmen, Konstruktionen, Definitionen, Ideen, etc., müssen durch genaue Verweise auf die entsprechende Literatur kenntlich gemacht werden. Lange wörtliche Zitate sollen vermieden werden. Gegebenenfalls muss erläutert werden, inwieweit und zu welchem Zweck fremde Inhalte modifiziert wurden. Die Struktur der Arbeit muss klar erkenntlich sein, und der Leser soll gut durch die Arbeit geführt werden. Die Darstellung aller Begriffe und Verfahren soll mathematisch formal fundiert sein. Für jeden wichtigen Begriff sollen Erläuterungen und Beispiele angegeben werden, ebenso für die Abläufe der beschriebenen Verfahren sowie Konstruktionen. Wo es angemessen ist, sollen Illustrationen die Darstellung vervollständigen. Schließlich sollen alle Lemmata und Sätze möglichst lückenlos bewiesen werden. Die Beweise sollen leicht nachvollziehbar dokumentiert sein.

Dresden, 20. Juni 2019

| | |
|---|---|
| Unterschrift von Heiko Vogler | Unterschrift von Frederic Dörband |

**Literatur**


[AD78] A. Arnold und M. Dauchet. „Théorie des magmoïdes. I“. In: *RAIRO Inform. Théor.* 12.3 (1978), S. 235–257, v. ISSN: 0399-0540. DOI: `10.1051/ita/1978120302351`. URL: `https://doi.org/10.1051/ita/1978120302351`.

[AD79] A. Arnold und M. Dauchet. „Théorie des magmoïdes. II“. In: *RAIRO Inform. Théor.* 13.2 (1979), S. 135–154. ISSN: 0399-0540. DOI: `10.1051/ita/1979130201351`. URL: `https://doi.org/10.1051/ita/1979130201351`.

[DPV05] M. Droste, C. Pech und H. Vogler. „A Kleene theorem for weighted tree automata“. In: *Theory Comput. Syst.* 38.1 (2005), S. 1–38. ISSN: 1432-4350. DOI: `10.1007/s00224-004-1096-z`. URL: `https://doi.org/10.1007/s00224-004-1096-z`.

[Kle56] S. C. Kleene. „Representation of events in nerve nets and finite automata“. In: *Automata studies*. Annals of mathematics studies, no. 34. Princeton University Press, Princeton, N. J., 1956, S. 3–41.

[Sch61] M. P. Schützenberger. „On the definition of a family of automata“. In: *Information and Control* 4 (1961), S. 245–270. ISSN: 0890-5401.

[Str09] L. Straßburger. „A Kleene theorem for forest languages“. In: *Language and automata theory and applications*. Bd. 5457. Lecture Notes in Comput. Sci. Springer, Berlin, 2009, S. 715–727. DOI: `10.1007/978-3-642-00982-2_61`. URL: `https://doi.org/10.1007/978-3-642-00982-2_61`.

[TW68] J. W. Thatcher und J. B. Wright. „Generalized finite automata theory with an application to a decision problem of second-order logic“. In: *Math. Systems Theory* 2 (1968), S. 57–81. ISSN: 0025-5661. DOI: `10.1007/BF01691346`. URL: `https://doi.org/10.1007/BF01691346`.

## Selbstständigkeitserklärung

Hiermit versichere ich, die vorliegende Arbeit selbstständig, ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der von mir angegebenen Quellen angefertigt zu haben. Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche gekennzeichnet.

Die Arbeit wurde noch keiner Prüfungsbehörde in gleicher oder ähnlicher Form vorgelegt.

Dresden, den 06. August 2019

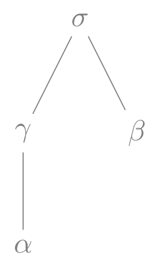
# Contents

# Chapter 1: Introduction

## 1.1 Motivation

Formal languages are studied intensively in theoretical computer science. A prominent class of formal languages is that of string languages over some alphabet $\Sigma$. A *word* over $\Sigma$ is a finite sequence over $\Sigma$ and a string language is a set of words. Another important class of formal languages is that of tree languages over a *ranked* alphabet $\Sigma$. A *tree* over $\Sigma$ is a well-bracketed string over $\Sigma$, usually depicted by a graph theoretical tree.

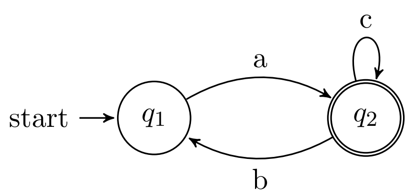
$aababcba$




A word and a tree. The tree corresponds to $\sigma(\gamma(\alpha), \beta)$.

Usually, formal languages that are interesting from a theoretical point of view are infinite sets. However, one wants to use finite formalisms to represent these languages. This is done in order to allow computers to execute algorithms on these infinite languages despite having only finite storage and finite processing speed.

Well-studied finite formalisms representing formal languages include grammars, automata and rational expressions. In the string case, these amount to the following concepts. Grammars are essentially finite sets of rules that generate parts of words. The language generated by a grammar is the set of words that can be generated by iteratively applying rules.[1] Automata are essentially graphs with start and end nodes where a word is accepted if one can reach a final node from a start node while processing a symbol of the word for every edge that is used. Rational expressions, on the other side, define a class of languages as the smallest class that contains finite languages and is closed under certain operations. These operations usually include union, concatenation, and Kleene star.




$ac^*(bac^*)^*$

An automaton and a rational expression.

[1] In this thesis however, grammars do not play any further role, whence we will stop mentioning them in the remainder of this chapter.

Automata and rational expressions can also be introduced for tree languages. However, the formalisms are mathematically more tedious. The concatenation operation of strings, for example, can not simply be lifted to trees, as trees do not have one specific leaf node. Therefore, tree concatenation is realized by a family of operators (see [9, Definition 2.27.]) and hence also the Kleene star operation becomes more cumbersome to deal with.

This mathematical pain goes away when one considers forests (finite tuples of trees) and formalizes automata and rational expressions for them. Each forest consists of a certain number, $m$, of trees and each tree may contain variables from an $n$-element set $X_n$. In this case, we say that the forest is an $(m, n)$-forest. Now, given an $(m, n)$-forest $\xi_1$ and an $(n, k)$-forest $\xi_2$, we can easily define $\xi_1 \cdot \xi_2$ as the forest obtained from $\xi_1$ by replacing each occurrence of the variable $x_i$ by the $i$-th tree in $\xi_2$. This is more aligned with the corresponding theory for string languages and hence is of high theoretical interest.

In all of these cases (strings, trees, and forests), there is a so-called *Kleene result*, which states that the class of languages accepted by automata equals the class of languages generated by rational expressions. In the string case, Kleene first proved this result in 1956 in [14]. Twelve years later, Thatcher and Wright [19] proved the statement in the tree case. The forest case was proven in 2009 by Straßburger [18].

Kleene's result has been extended in serveral directions. Enriching the algebra generating the formal languages was one direction. Another extension was to consider weighted languages. That is, instead of speaking about presence or absence of transitions in an automaton or words in a language, one assigns to each transition (and word) a weight from a weight space $S$. Therefore, a weighted (string) language is not simply a set of words, but rather a map from the set of all words to $S$. This can be done for tree languages as well. A pleasant fact is that rational expressions do not differ significantly between the unweighted and weighted case. The closure properties stay almost the same: finite sets become *finitely supported* weighted languages, union becomes sum and multiplication with a scalar, and concatenation and Kleene star stay the same (of course with different definitions than in the unweighted case).

In both the weighted string and tree case, Kleene-like results have been proven. For strings, this was done by Schützenberger in [15] (see also [8]). The tree case was proven by Droste, Pech, and Vogler [6] in 2005. However, a Kleene-like result for weighted forest languages has not been proven.

In this thesis, we introduce weighted forest automata and rational weighted forest expressions and prove a Kleene-like result in this case. We do this by combining the proof of the weighted tree case [6] with the proof of the unweighted forest case [18].

## 1.2 Structure of the Thesis

In this section, we outline the contents of this thesis.

Chapter 1 was a brief introduction to the idea behind trees, forests, and weighted tree (and forest) languages. Moreover, we presented a timeline of Kleene results. This chapter demonstrates the relevance of Kleene results in recent scientific research and embeds our general case of weighted forest languages into the existing literature on the topic.

Chapter 2 is an introduction to the necessary mathematical groundwork. We recall notions such as semirings, ranked alphabets, trees, and weighted languages. Furthermore, we give a definition for weighted tree automata and thereby make our hand-wavey explanations from Chapter 1 precise.

In Chapter 3, we introduce the notion of weighted forest automata (WFA), prove a decomposition result, and introduce normal forms of WFA. The decomposition result is a property of weighted forest languages that Straßburger already observed in [18]. He showed that (unweighted) recognizable forest languages are cartesian products of (unweighted) recognizable tree languages. A similar result holds for the weighted case. The normal forms are inspired by [6] and are mainly used as a simplifying tool for later proofs.

Chapters 4 and 5 contain the proof of our Kleene-like result. In Chapter 4, we prove the different closure properties of the class of recognizable weighted forest languages. This is done by explicit constructions. We refer the reader to the introduction of Chapter 4 for a more detailed insight into the proofs and their execution. In essence, Chapter 5 introduces rational weighted forest expressions and then proves that every recognizable weighted forest language can be generated by such an expression.

In Chapter 6, we wrap up this thesis and give an outview on possible future research in the field of weighted forest languages.

# Chapter 2: Preliminaries

In this chapter, we introduce the mathematical tools that lay the groundwork for our upcoming studies and results. We start by introducing elementary formalisms from mathematics and then recall the notions of trees and weighted tree languages (over semirings). We cite corresponding introductory papers, whenever we define more complex concepts.

## 2.1 Notations and Basic Definitions

We use the conventional set-theoretic approach to mathematics after Zermelo and Fraenkel with the axiom of choice. For a set $M$, we denote the **cardinality of** $M$ by $\#M$.

The **set of positive integers** is denoted by $\mathbb{N}_+ := \{1, 2, 3, ...\}$. The **set of nonnegative integers** is denoted by $\mathbb{N}_0 := \mathbb{N}_+ \cup \{0\}$. If not stated differently we write $\mathbb{N}$ for $\mathbb{N}_+$. We abbreviate $[n, m] := \{n, n+1, ..., m\}$ and $[n] := [1, n]$ for any $n, m \in \mathbb{N}$ with $n \leq m$. Moreover we denote $[0] := \emptyset$.

We denote the (**contravariant**) **composition** of relations $\rho \subseteq A \times B$ and $\pi \subseteq B \times C$ by $\pi \circ \rho \subseteq A \times C$ or in some cases simply $\pi\rho$.

We quantify over multiple maps $f_1 \colon A \longrightarrow B, \ldots, f_n \colon A \longrightarrow B$, using the shorthand notation $f_1, \ldots, f_n \colon A \longrightarrow B$. Moreover, we identify constant maps with their unique image. If there exists a bijective map $f \colon A \longrightarrow B$, we write $A \cong B$.

A **semiring** (c.f. [12] and [13]) is an algebraic structure $(S, +, \cdot, 0, 1)$ such that $(S, +, 0)$ is a commutative monoid, $(S, \cdot, 1)$ is a monoid, for every $a, b, c \in S$ it holds that $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(b + c) \cdot a = b \cdot a + c \cdot a$, and for any $a \in S$ it holds that $a \cdot 0 = 0 = 0 \cdot a$. Due to this last property, some literature calls 0 **absorbing** (see [6]) whereas other literature says that 0 **annihilates** $S$ (see [3, after Proposition 1.11]). Note that it is customary to write $ab$ instead of $a \cdot b$.

Given a finite set $I$ and a family $(a_i \in S \mid i \in I)$, we denote the sum of all elements $a_i$ as $\sum_{i \in I} a_i$. If $S$ is a commutative semiring, we analogously denote the product of all elements $a_i$ as $\prod_{i \in I} a_i$.

Note that $\mathbb{N}_0$ acts on any semiring $S$. That is, we can define the map $\cdot \colon \mathbb{N}_0 \times S \longrightarrow S$ where for every $n \in \mathbb{N}_0$ and $a \in S$ the value $n \cdot a$ is defined inductively as follows.

$$0_{\mathbb{N}_0} \cdot a := 0_S \qquad \text{and} \qquad (k+1) \cdot a := k \cdot a + a \quad \text{for every } k \in \mathbb{N}_0$$

We illustrate semirings by listing some important examples. More involved lists of examples can be found in [12], [13], and [6].

1. The **Boolean semiring** is $\mathbb{B} = (\{0, 1\}, \vee, \wedge, 0, 1)$ with logical disjunction and logical conjunction. Weighing an automaton (or any formalism in natural langugage processing, for that matter) with $\mathbb{B}$ is equivalent to considering the unweighted case.

2. The **semiring of natural numbers** is $\mathbb{N} = (\mathbb{N}, +, \cdot, 0, 1)$ with the standard addition and multiplication.

3. The **tropical semiring** is $Trop = (\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$ with the minimum operation (recall that $\mathbb{N}$ is well-ordered) and addition.

4. Every **ring** is a semiring. In particular every field is a semiring.

Let $\Sigma$ be a set. We call $\Sigma$ an **alphabet**, if $\#\Sigma \in \mathbb{N}$, i.e. $\Sigma$ is a finite and nonempty set. $\Sigma^*$ denotes the **set of words over** $\Sigma$, i.e. finite ordered sequences of elements from $\Sigma$. The **length of** $\omega \in \Sigma^*$ is denoted $|\omega|$. Let $\Sigma^n := \{\omega \in \Sigma^* \mid |\omega| = n\}$ for $n \in \mathbb{N}_0$ and $\varepsilon$ be the unique element of $\Sigma^0$. For $\sigma \in \Sigma, \omega \in \Sigma^*$, we denote the **number of occurrences of** $\sigma$ **in** $\omega$ as $|\omega|_\sigma$.

## 2.2 Trees

Let $\Sigma$ be an alphabet and $r\colon \Sigma \longrightarrow \mathbb{N}_0$ a map. We call the pair $(\Sigma, r)$ a **ranked alphabet** and for any $a \in \Sigma$, $r(a)$ is the **rank of** $a$. If $r$ is clear from the context, we will withhold $r$ and simply write $\Sigma$ instead of $(\Sigma, r)$.

Let $\Sigma$ be a ranked alphabet and $k \in \mathbb{N}_0$. We define $\Sigma^{(k)} := \{a \in \Sigma \mid r(a) = k\} = r^{-1}(\{k\})$. Since $\Sigma$ is finite and nonempty, there exists $\text{maxrk}(\Sigma) := \max\ r(\Sigma)$ called the **maximal rank of** $\Sigma$.

We fix the sets $X := \{x_i \mid i \in \mathbb{N}\}$ and for any $n \in \mathbb{N}$, $X_n := \{x_i \mid i \in [n]\}$. These will be used as sets of variables in trees and are assumed to be disjoint from any other occuring set.

Let $\Sigma$ be a ranked alphabet and $A$ a set. Then the **set of trees over** $\Sigma$ **indexed by A**, abbreviated by $T_\Sigma(A)$, is the smallest set $T \subseteq (\Sigma \cup A \cup C)^*$ (where $C$ consists of open and closed round brackets and the comma), such that

$$A \subseteq T$$

and for any $k \in \mathbb{N}_0$, $\sigma \in \Sigma^{(k)}$, and $t_1, ...t_k \in T$, also

$$\sigma(t_1, ..., t_k) \in T.$$

Moreover $T_\Sigma := T_\Sigma(\emptyset)$ and for each $\alpha \in \Sigma^{(0)}$, we identify $\alpha$ with $\alpha()$.

A **tree language** is a set of trees $L \subseteq T_\Sigma$.

A tree $t \in T_\Sigma(X)$ is called **linear**, if for every $i \in \mathbb{N}$ $x_i$ occurs at most once in $t$. The **leaves** of $t$ are the positions of $t$ that do not have any successors.

Let $l \in \mathbb{N}_0, t \in T_\Sigma(X_n)$ and $s_1, ..., s_n \in T_\Sigma(A)$. Define the **substitution of** $s_1, \ldots, s_n$ **into** $t$ inductively by:

$$t[s_1, ..., s_n] := s_i$$

whenever $t = x_i,\ i \in [n]$ and

$$t[s_1, \ldots, s_n] := \sigma(t_1[s_1, ..., s_n], ..., t_k[s_1, ..., s_n])$$

whenever $t = \sigma(t_1, ..., t_k)$ for some $k \in \mathbb{N}_0, \sigma \in \Sigma^{(k)}$, and $t_1, ..., t_k \in T_\Sigma$.

Define $\tilde{T}_\Sigma(X_n)$ as the set of trees $t \in T_\Sigma(X_n)$ such that the left-to-right sequence of variables in $t$ is $x_1 \dots x_n$ (where $n \in \mathbb{N}_0$).

Furthermore we define for any $n \in \mathbb{N}_0$ the maps

$$\begin{aligned} \mathrm{pos}\colon T_\Sigma(X_n) &\longrightarrow \mathcal{P}(\mathbb{N}^*), \\ \mathrm{size}\colon T_\Sigma(X_n) &\longrightarrow \mathbb{N}_0, \\ \mathrm{ht}\colon T_\Sigma(X_n) &\longrightarrow \mathbb{N}_0, \end{aligned}$$

where for every $i \in [n]$ we define

$$\mathrm{pos}(x_i) = \{\varepsilon\}, \qquad \mathrm{size}(x_i) = 0, \qquad \mathrm{ht}(x_i) = 1,$$

and for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $t_1, \dots, t_k \in T_\Sigma(X_n)$

$$\begin{aligned} \mathrm{pos}(\sigma(t_1, \dots, t_k)) &= \{\varepsilon\} \cup \bigcup_{i \in [k]} i \cdot \mathrm{pos}(t_i), \\ \mathrm{size}(\sigma(t_1, \dots, t_k)) &= 1 + \sum_{i \in [k]} \mathrm{size}(t_i), \\ \mathrm{ht}(\sigma(t_1, \dots, t_k)) &= 1 + \max_{i \in [k]} \mathrm{ht}(t_i). \end{aligned}$$

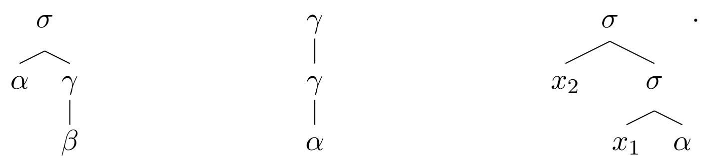
Moreover for every $t \in T_\Sigma(X_n)$ and $w \in \mathrm{pos}(t)$ we define $\mathrm{lab}(t, w)$ as the symbol from $\Sigma \cup X_n$ standing in $t$ at position $w$.

Note that considering $A$ as a set of nullary symbols, we identify the sets $T_\Sigma(A)$ and $T_{\Sigma \cup A}$.

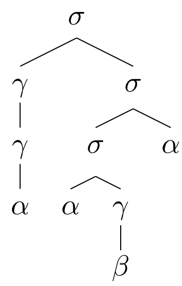
**Example 2.2.1.** Let $\Sigma := \{\sigma^{(2)}, \gamma^{(1)}, \beta^{(0)}, \alpha^{(0)}\}$ and consider the trees

$$\begin{aligned} t_1 &:= \sigma(\alpha, \gamma(\beta)) && \in T_\Sigma, \\ t_2 &:= \gamma(\gamma(\alpha)) && \in T_\Sigma, \text{ and} \\ t_3 &:= \sigma(x_2, \sigma(x_1, \alpha)) && \in T_\Sigma(X_2) \end{aligned}$$

which can be visualized as follows from left to right




It holds that $\mathrm{pos}(t_1) = \{\varepsilon, 1, 2, 21\}$, $\mathrm{ht}(t_2) = 3$, and $\mathrm{size}(t_3) = 3$. Moreover, the substitution of $t_1$ and $t_2$ into $t_3$, $t_3[t_1, t_2]$, is



,

or formally, $t_3[t_1, t_2] = \sigma\Big(\gamma(\gamma(\alpha)), \sigma\big(\sigma(\alpha, \gamma(\beta)), \alpha\big)\Big)$. ■

## 2.3 Weighted Languages and Weighted Tree Automata

Unweighted tree automata are introduced in [9] and [10]. As we solely deal with weighted automata throughout this thesis, we do not dive into details about the unweighted case. The weighted case has been introduced in [4], [16], and [17]. However, for weighted tree automata we use the notation given in [6].

Let $M$ be a set and $S$ a semiring. We call a map $\varphi\colon M \longrightarrow S$ a **weight map on** $M$ (compare to the definition of *formal power series* in [7, Ch.1 Sec.3]). Given two weight maps on $M$, say $\varphi$ and $\psi$, we denote their **sum** by $\varphi + \psi\colon M \longrightarrow S$, given for every $m \in M$ by

$$(\varphi + \psi)(m) \coloneqq \varphi(m) + \psi(m).$$

Moreover we denote the (**Hadamard**) **product** of $\varphi$ and $\psi$ by $\varphi \odot \psi\colon M \longrightarrow S$, given for every $m \in M$ by

$$(\varphi \odot \psi)(m) \coloneqq \varphi(m) \cdot \psi(m).$$

If $S$ is commutative, $I$ is a finite index set, and $(\varphi_i\colon M \longrightarrow S \mid i \in I)$ a family of weight maps on $M$. We denote by $\sum_{i\in I} \varphi_i$ and by $\prod_{i\in I} \varphi_i$ the sum and the Hadamard product of the $\varphi_i$, respectively. Both the sum and the product can happen in arbitrary order, as $S$ is commutative.

For a weight map $\varphi\colon M \longrightarrow S$ and $a \in S$ we denote the **scalar multiple** $a\varphi\colon M \longrightarrow S$, given for every $m \in M$ by

$$(a\varphi)(m) \coloneqq a \cdot \varphi(m).$$

The **support** of $\varphi\colon M \longrightarrow S$ is $\operatorname{supp}(\varphi) \coloneqq \{m \in M \mid \varphi(m) \neq 0\}$.

Let $m \in M$. We denote by $\mathbb{1}_m$ the weight map on $M$ given for any $m' \in M$ by

$$\mathbb{1}_m(m') = \begin{cases} 1 & , m = m' \\ 0 & , m \neq m' \end{cases}.$$

If $M = \Sigma^*$ for some finite set $\Sigma$, we call a weight map on $M$ a **weighted language over** $\Sigma$. If $\Sigma$ is a ranked alphabet and $M = T_\Sigma$, we call a weight map on $M$ a **weighted tree language over** $\Sigma$.

Let $S$ now be commutative. A **weighted tree automaton over** $S$ (short: WTA) is a tuple $\mathscr{A} = (Q, \Sigma, \delta, F)$, where $Q$ is a finite set, $\Sigma$ is a ranked alphabet such that $\Sigma \cap Q = \emptyset$, $F \subseteq Q$ (called the **final states**), and $\delta = (\delta_\sigma \mid \sigma \in \Sigma)$ is a family (so called **state behaviors with costs**) of maps of the form $\delta_\sigma \colon Q^k \times Q \longrightarrow S$ for $\sigma \in \Sigma^{(k)}$, $k \geq 0$.

Let $\mathscr{A}$ be a WTA and $t \in T_\Sigma(Q)$. A **run of** $\mathscr{A}$ **on** $t$ is a map $\rho\colon \operatorname{pos}(t) \longrightarrow Q$ such that $\rho(w) = \operatorname{lab}(t, w)$ for any $w \in \operatorname{pos}_Q(t)$. We say $\rho$ **ends in** $q \in Q$ if $\rho(\varepsilon) = q$. The set of runs of $\mathscr{A}$ on $t$ ending in $q$ is denoted $R_\mathscr{A}(t, q)$.

Let $\rho$ be a run of $\mathscr{A}$ on $t$ and $w \in \operatorname{pos}(t)$. The **cost of** $w$ **in** $t$ **under** $\rho$ is defined by

$$c_\mathscr{A}(\rho, t, w) \coloneqq \begin{cases} \delta_\sigma(\rho(w1), \ldots, \rho(wk), \rho(w)) & \text{, if } \operatorname{lab}(t, w) = \sigma \in \Sigma^{(k)}, k \geq 0 \\ 1 & \text{, if } \operatorname{lab}(t, w) \in Q \end{cases}$$

and the **cost of** $\rho$ **for** $t$ is defined by

$$c_\mathscr{A}(\rho, t) = \prod_{w \in \operatorname{pos}(t)} c_\mathscr{A}(\rho, t, w).$$

The **weighted tree language accepted by** $\mathscr{A}$, denoted $\mathscr{L}(\mathscr{A})\colon T_\Sigma \longrightarrow S$, is defined for any $t \in T(\Sigma)$ by

$$\mathscr{L}(\mathscr{A})(t) \coloneqq \sum_{f \in F} \sum_{\rho \in R_\mathscr{A}(t,f)} c_\mathscr{A}(\rho, t).$$

A weighted tree language $\varphi\colon T_\Sigma \longrightarrow S$ is called **recognizable weighted tree languages over** $\Sigma$ **and** $S$ if there exists a WTA $\mathscr{A}$ such that $\mathscr{L}(\mathscr{A}) = \varphi$. The class of recognizable tree series over $\Sigma$ and $S$ is denoted by $\operatorname{REC}(T_\Sigma, S)$. Note that this also defines the class $\operatorname{REC}(T_{\Sigma \cup X_n}, S)$.

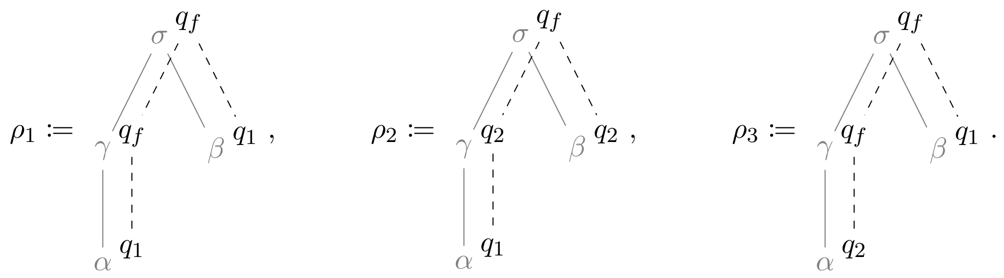
**Example 2.3.1.** Let $\Sigma \coloneqq \{\sigma^{(2)}, \gamma^{(1)}, \beta^{(0)}, \alpha^{(0)}\}$ and consider the WTA $\mathscr{A} = (Q, \Sigma, \delta, F)$ where

$$Q = \{q_1, q_2, q_f\} \text{ and } F = \{q_f\}.$$

Moreover, $\delta$ is 0 except in the cases

$$\begin{aligned} \delta_\alpha(q_1) = \delta_\alpha(q_2) = \delta_\beta(q_1) &= 1 \\ \forall q \in Q\colon \delta_\gamma(q, q_f) &= 1 \\ \forall q, q' \in Q\colon \delta_\sigma(q, q', q_f) &= 1. \end{aligned}$$

We consider the tree $t = \sigma(\gamma(\alpha), \beta)$. The following dashed trees illustrate exemplary runs of $\mathscr{A}$ on $t$.




By definition, $c_{\mathscr{A}}(\rho_1, t, \varepsilon) = c_{\mathscr{A}}(\rho_1, t, 1) = c_{\mathscr{A}}(\rho_1, t, 2) = c_{\mathscr{A}}(\rho_1, t, 11) = 1$ and hence $c_{\mathscr{A}}(\rho_1, t) = 1$. The same holds for $\rho_3$. In $\rho_2$, however, we have that $c_{\mathscr{A}}(\rho_2, t, 1) = 0$ and hence $c_{\mathscr{A}}(\rho_2, t) = 0$.

One easily sees that the only runs of $\mathscr{A}$ on $t$ that have non-vanishing cost are $\rho_1$ and $\rho_3$. Therefore, we obtain

$$\mathscr{L}(\mathscr{A})(t) = c_{\mathscr{A}}(\rho_1, t) + c_{\mathscr{A}}(\rho_3, t) = 1 + 1.$$

It now depends on the choice of the semiring $S$, what $1 + 1$ evaluates to. In the boolean semiring, $S = \mathbb{B}$, we have $\mathscr{L}(\mathscr{A})(t) = 1$. In the natural semiring, $S = \mathbb{N}_0$, we have $\mathscr{L}(\mathscr{A})(t) = 2$.

For arbitrary semirings $S$, one can now check that for every tree $t' \in T_\Sigma$ with $\mathrm{ht}(t) > 1$, it holds that

$$\mathscr{L}(\mathscr{A})(t') = (2 \cdot 1)^{\# \operatorname{pos}_\alpha(t)}.$$

In $S = \mathbb{B}$, this evaluates to 1 (therefore, $\mathscr{L}(\mathscr{A})$ is the characteristic weight map for trees of height greater than 1) and in $S = \mathbb{N}_0$, this evaluates to $2^{\# \operatorname{pos}_\alpha(t)}$. ■

# Chapter 3: Forests and Weighted Forest Automata

In this chapter we define *forests* over $\Sigma$. In essence, a forest over $\Sigma$ is a tuple of trees over $\Sigma$ and the set of such forests will be denoted $T(\Sigma)$. We can allow a certain number of variables to occur in forests and hence have an option to vertically concatenate matching forests (if the first one uses $n$ variables and the second one consists of $n$ trees, copy the $i$-th tree of the second forest into every occurrence of $x_i$ in the first forest). Moreover, we can horizontally concatenate two forests. These operations turn $T(\Sigma)$ into a so-called *magmoid* (see [5], [1], and [2]).

We continue by introducing weighted forest automata (WFA), which are very similar to weighted tree automata. Essentially, a WFA $\mathscr{A}$ consists of a set of states and a family of transition weights. Given a forest $\xi$, a *run* of $\mathscr{A}$ on $\xi$ labels each position in $\xi$ with some state of the automaton. The weight of a run is simply the product of the occurring transition weights. We sum up the weights of all runs of $\mathscr{A}$ on $\xi$ to obtain the weight of $\xi$ in the language of $\mathscr{A}$, denoted $\mathscr{L}(\mathscr{A})(\xi)$. To our knowledge, WFA have not been introduced in the literature before[2]. We introduce different semantics for WFA and show their equivalence.

In section 3.3, we prove the central theorem of the theory of recognizable weighted forest languages. Namely, that each such weighted forest language can be decomposed into a "horizontal concatenation" of weighted *tree* languages.

We wrap up this chapter by introducing two useful normal forms for weighted forest automata. These are inspired by normal forms of weighted tree automata from [6].

Throughout this chapter, let $\Sigma$ be a ranked alphabet.

## 3.1 Forests

**Definition 3.1.1.** Let $m, n \in \mathbb{N}_0$. We define the set

$$T(\Sigma)^m_n \coloneqq \{n\} \times T_\Sigma(X_n)^m,$$

of **$(m,n)$-forests** over $\Sigma$. The values $m$ and $n$ are called the **upper** and **lower rank** of an $(m,n)$-forest, respectively. The set of all **forests** over $\Sigma$ is then defined as

$$T(\Sigma) \coloneqq \bigcup_{m,n \in \mathbb{N}_0} T(\Sigma)^m_n.$$

For notational convenience, we introduce the two sets

$$T(\Sigma)^m = \bigcup_{n \in \mathbb{N}_0} T(\Sigma)^m_n \qquad \text{and} \qquad T(\Sigma)_n = \bigcup_{m \in \mathbb{N}_0} T(\Sigma)^m_n.$$

Moreover, we will denote forests using angle brackets, to aid readability of examples. ■

[2] However, unweighted forest automata have been studied in [18].

**Remark 3.1.2.** First note that the sets in $(T(\Sigma)_n^m \mid m, n \in \mathbb{N}_0)$ are pairwise disjoint. This is an immediate consequence of the definition, as forests with different numbers of variables have different first components and forests with different numbers of trees have different tuple sizes. ■

**Definition 3.1.3.** Let $m, n \in \mathbb{N}_0$ and $\xi = \langle n, t_1, \ldots, t_m \rangle \in T(\Sigma)_n^m$. We define the maps

$$\begin{aligned} \text{pos}\colon T(\Sigma)_n^m &\longrightarrow \mathcal{P}(\mathbb{N} \times \mathbb{N}^*), \\ \text{size}\colon T(\Sigma)_n^m &\longrightarrow \mathbb{N}_0, \\ \text{ht}\colon T(\Sigma)_n^m &\longrightarrow \mathbb{N}_0, \end{aligned}$$

where

$$\begin{aligned} \text{pos}(\xi) &= \bigcup_{i=1}^{n} \{i\} \times \text{pos}(t_i), \\ \text{size}(\xi) &= \sum_{i \in [m]} \text{size}(t_i), \\ \text{ht}(\xi) &= \max\{\text{ht}(t_i) \mid i \in [m]\}, \end{aligned}$$

and call $\text{pos}(\xi)$ the **set of positions of** $\xi$.

Moreover, we define for every $w = (i, w') \in \text{pos}(\xi)$ the **label of** $\xi$ **at position** $w$ as

$$\text{lab}(\xi, w) = \text{lab}(t_i, w').$$

Let $l_1, \ldots, l_n \in \mathbb{N}_0$ such that for any $i \in [n]$, the variable $x_i$ occurs in $\xi$ exactly $l_i$ times. Given some $r \in \mathbb{N}_0$ and a family of trees

$$(\zeta_j^i \in T_\Sigma(X_r) \mid i \in [n], j \in [l_i]),$$

we define

$$\xi[\forall i \in [n] : x_i \leftarrow \zeta_1^i, \ldots, \zeta_{l_i}^i] \in T(\Sigma)_r^m$$

as the forest obtained from $\xi$ by replacing the $j$-th occurence[3] of $x_i$ by $\zeta_j^i$ for all $i \in [n]$ and $j \in [l_i]$.

If no variable occurs in $\xi$, this simply interprets $\xi$ as an element of $T(\Sigma)_r^m$. ■

**Definition 3.1.4.** Let $S$ be a commutative semiring and $m, n \in \mathbb{N}_0$. A weight map over $T(\Sigma)_n^m$ is called **weighted (*m*,*n*)-forest language**.

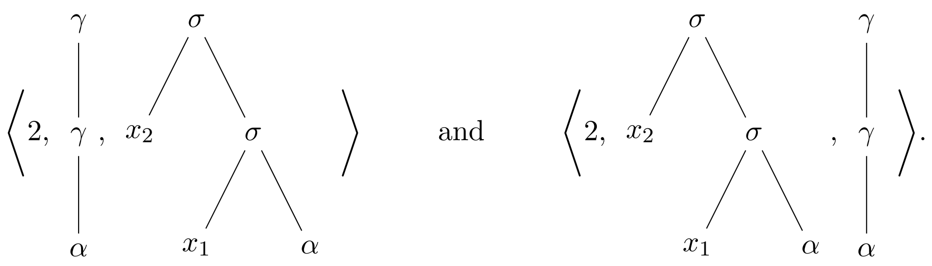
A weighted $(m,n)$-forest language $\tau$ is called **rectangular** if there exist weighted $(1,n)$-forest languages $\tau_1, \ldots, \tau_m$ such that for any $\langle n, t_1, \ldots, t_m \rangle \in T(\Sigma)_n^m$ it holds that

$$\tau(\langle n, t_1, \ldots, t_m \rangle) = \tau_1(\langle n, t_1 \rangle) \cdot \ldots \cdot \tau_m(\langle n, t_m \rangle).$$

In this case, we call the $\tau_1, \ldots, \tau_m$ the **rectangular components of** $\tau$. ■

[3] Occurences are counted with respect to the left-to-right order on the leafs of $\xi$.

**Example 3.1.5.** We continue Example 2.2.1. Consider the set $F \subseteq T(\Sigma)_2^2$ consisting of the forests



That is, $F = \{\langle 2, t_2, t_3\rangle, \langle 2, t_3, t_2\rangle\}$.

The weighted $(2,1)$-forest language $\mathbb{1}_F$ that maps forests from $F$ to 1 and everything else to 0 is not rectangular if $0 \neq 1$. This can be easily seen as follows.

Assume there exist weighted $(1,2)$-forest languages $\varphi$ and $\psi$, which are rectangular components of $\mathbb{1}_F$. The equality

$$\varphi(t_2) \cdot \psi(t_3) = \mathbb{1}_F(\langle 2, t_2, t_3\rangle) = 1$$

shows that $\varphi(t_2)$ and $\psi(t_3)$ are invertible with respect to multiplication in $S$. The same holds for $\varphi(t_3)$ and $\psi(t_2)$. But then

$$\varphi(t_2) \cdot \psi(t_2)$$

is invertible as well and hence not 0. This contradicts the definition of $\mathbb{1}_F$.

Of course, we can find a superset $F' \supsetneq F$ such that $\mathbb{1}_{F'}$ is rectangular, namely

$$F' := F \cup \{\langle 2, t_2, t_2\rangle, \langle 2, t_3, t_3\rangle\}.$$

The rectangular components of $\mathbb{1}_{F'}$ are both equal to $\mathbb{1}_{\{t_2,t_3\}}$. ■

**Remark 3.1.6.** Note that rectangular components of a weighted forest language are only unique up to a scalar. ■

**Definition 3.1.7.** We define the **vertical concatenation** as the partial binary operation $\cdot$ on $T(\Sigma)$, given for every $\langle n, u_1, \ldots, u_m\rangle \in T(\Sigma)_n^m$ and $\langle l, v_1 \ldots, v_n\rangle \in T(\Sigma)_l^n$ by

$$\langle n, u_1, \ldots, u_m\rangle \cdot \langle l, v_1 \ldots, v_n\rangle = \langle l, w_1, \ldots, w_m\rangle \in T(\Sigma)_l^m,$$

where for any $i \in [m]$ we define

$$w_i := u_i[v_1, \ldots, v_n].$$

Note that the lower rank of the first operand and the upper rank of the second operand have to be equal in order for vertical concatenation to be defined.

We moreover define the operation $\times : T(\Sigma) \times T(\Sigma) \longrightarrow T(\Sigma)$ via

$$\langle n, u_1, \ldots, u_m\rangle \times \langle n', v_1, \ldots, v_{m'}\rangle = \langle n + n', u_1, \ldots, u_m, w_1, \ldots, w_{m'}\rangle,$$

where for any $i \in [m']$ we define

$$w_i := v_i[x_{n+1}, \ldots, x_{n+n'}].$$

We call $\times$ the **horizontal concatenation**. ■

**Example 3.1.8.** We continue Example 2.2.1. First note that for any $k \le l \in \mathbb{N}_0$, we can interpret any tree over $k$ variables as a tree over $l$ variables, and formally have $T_\Sigma(X_k) \subseteq T_\Sigma(X_l)$. This shows the following inclusions.

$$\begin{aligned} \xi_1 &:= \langle 3, t_2, t_3 \rangle && \in T(\Sigma)^2_3, \\ \xi_2 &:= \langle 1, \alpha, \gamma(x_1), \alpha \rangle && \in T(\Sigma)^3_1. \end{aligned}$$

It holds that $\mathrm{pos}(\xi_2) = \{(1,\varepsilon),(2,\varepsilon),(2,1),(3,\varepsilon)\}$, $\mathrm{size}(\xi_2) = 3$, and $\mathrm{ht}(\xi_2) = 2$. The vertical concatenation of $\xi_1$ and $\xi_2$ is

$$\xi_1 \cdot \xi_2 = \Big\langle 1, \gamma(\gamma(\alpha)), \sigma(\gamma(x_1), \sigma(\alpha, \alpha)) \Big\rangle \in T(\Sigma)^2_1,$$

and their direct sum is

$$\xi_1 \times \xi_2 = \Big\langle 4, t_2, t_3, \alpha, \gamma(x_4), \alpha \Big\rangle \in T(\Sigma)^5_4.$$

■

This demonstrates the basic idea behind the defined forest operations. Next we introduce the respective neutral elements.

**Remark 3.1.9.** First note that the set $T(\Sigma)^0_n$ where $n \in \mathbb{N}_0$ is a singleton, namely

$$T(\Sigma)^0_n = \{\langle n \rangle\}.$$

Moreover the element $\mathrm{o} := \langle 0 \rangle$ is neutral with respect to $\times$, as immediately follows from the definition. For $k \in \mathbb{N}, i \in [k]$, the elements

$$\pi^k_i := \langle k, x_i \rangle \in T(\Sigma)^1_k,$$

act as projections from the left on $T(\Sigma)^k$, that is, for any $k \in \mathbb{N}, i \in [k]$, and $\langle l, t_1, \ldots, t_k \rangle \in T(\Sigma)^k_l$ (where $l \in \mathbb{N}_0$),

$$\pi^k_i \cdot \langle l, t_1, \ldots, t_k \rangle = \langle l, t_i \rangle \in T(\Sigma)^1_l.$$

A special case is $k = i = 1$. We define $\mathtt{1} \coloneqq \pi_1^1 = \langle 1, x_1 \rangle$, and its finite horizontal multiples

$$\mathtt{1}_m \coloneqq \bigtimes_{i=1}^{m} \mathtt{1} = \langle m, x_1, \ldots, x_m \rangle.$$

This yields that for any $m \in \mathbb{N}, k, l \in \mathbb{N}_0$, $\langle k, t_1, \ldots, t_m \rangle \in T(\Sigma)_k^m$, and $\langle m, s_1, \ldots, s_l \rangle \in T(\Sigma)_m^l$,

$$\mathtt{1}_m \cdot \langle k, t_1, \ldots, t_m \rangle = \langle k, x_1[t_1, \ldots, t_m], \ldots, x_m[t_1, \ldots, t_m] \rangle = \langle k, t_1, \ldots, t_m \rangle,$$
$$\langle m, s_1, \ldots, s_l \rangle \cdot \mathtt{1}_m = \langle m, s_1[x_1, \ldots, x_m], \ldots, s_l[x_1, \ldots, x_m] \rangle = \langle m, s_1, \ldots, s_l \rangle.$$

Therefore, the $\mathtt{1}_m$ can be seen as the neutral elements for vertical concatenation. ■

**Remark 3.1.10.** Up to some degree, this Master's Thesis is related to the preceding Bachelor's Thesis [5]. To pinpoint this connection, we briefly introduce the concept of magmoids at this point. This interesting algebraic characterization of the structure of forests was presented in [1] and [2].

In essence, a **magmoid** is a tuple $(\mathbb{M}, \cdot, \times, \mathtt{0}, \mathtt{1})$, where $\mathbb{M}$ is a biranked set (that is, $\mathbb{M}$ is partitioned into (possibly empty) sets $\mathbb{M}_n^m$ for $m, n \in \mathbb{N}$), $\cdot$ is a partial associative binary operation on $\mathbb{M}$, defined if and only if the lower rank of the first operand equals the upper rank of the second operand, $\times$ is an associative binary operation on $\mathbb{M}$, $\times$ distributes over $\cdot$, and the elements $\mathtt{1}_m = \bigoplus_{i=1}^m \mathtt{1}$ and $\mathtt{0}$ are neutral with respect to $\cdot$ and $\times$ respectively.

To see that $(T(\Sigma), \cdot, \times, \mathtt{0}, \mathtt{1})$ satisfies these properties, we have to show associativity of both operations and distributivity of $\times$ over $\cdot$. Associativity of $\cdot$ has been proven in [11, Proposition 2.4] and associativity of $\times$ follows easily by directly evaluating both sides of the corresponding equation. Distributivity has been proven in [5, Lemma 9], again by a direct evaluation of the two sides of the corresponding equation.

For a detailed definition and more examples, see [5]. ■

## 3.2 Recognizable Weighted Forest Languages

In this subchapter, $S$ denotes a commutative semiring.

**Definition 3.2.1.** Let $m, n \in \mathbb{N}_0$. A **weighted $(m,n)$-forest automaton** (short: $(m,n)$-WFA) is a tuple $\mathscr{A} = (Q, \Sigma, S, I, F, E)$, where $Q$ is a finite set (of **states**), $\Sigma$ is a ranked alphabet with $\Sigma \cap Q = \emptyset$, $I = (I_1, \ldots, I_n)$ for some $I_1, \ldots, I_n \colon Q \longrightarrow S$ (called the **leaf weights**), $F = F_1 \times \cdots \times F_m$ for some $F_1, \ldots, F_m \subseteq Q$ (called the **root states**), and $E = (E_k \mid k \geq 0)$, where for any $k \geq 0$

$$E_k \colon Q^k \times \Sigma^{(k)} \times Q \longrightarrow S.$$

Each map $E_k$ is called **state transition weight**. Furthermore, $S$ is sometimes called **weight space**. In this thesis, we omit the brackets of the state tuples. ■

**Definition 3.2.2.** Let $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be an $(m,n)$-WFA. We lift the state transition weights to forests and state weights as follows. Let $k \in \mathbb{N}_0$. We define

$$E^{\mathscr{A}}_{k,0} \colon (S^Q)^k \times T(\Sigma)^0_k \times Q^0 \longrightarrow S$$

as the constant map $1 \in S$. Moreover we define the family of maps

$$\left(E^{\mathscr{A}}_{k,k'} \colon (S^Q)^k \times T(\Sigma)^{k'}_k \times Q^{k'} \longrightarrow S \mid k' \in \mathbb{N}, k \in \mathbb{N}_0\right)$$

by simultaneous induction.

The case $k' = 1$ is given for any $\omega_1, \ldots, \omega_k \colon Q \longrightarrow S$, $q \in Q$, and $\langle k, t_1 \rangle \in T(\Sigma)^1_k$ inductively on the structure of $t_1$. If $t_1 = x_i$ for some $i \in [k]$, define

$$E^{\mathscr{A}}_{k,1}((\omega_1, \ldots, \omega_k), \langle k, t_1 \rangle, q) = \omega_i(q).$$

If $t_1 = \alpha$ for some $\alpha \in \Sigma^{(0)}$, define

$$E^{\mathscr{A}}_{k,1}((\omega_1, \ldots, \omega_k), \langle k, t_1 \rangle, q) = E_0(\alpha, q).$$

Furthermore, if $t_1 = \sigma(\xi_1, \ldots, \xi_s)$ for some $s \geq 0$, $\sigma \in \Sigma^{(s)}$, and $\xi_1, \ldots, \xi_s \in T_\Sigma(X_k)$, define

$$\begin{aligned} &E^{\mathscr{A}}_{k,1}((\omega_1, \ldots, \omega_k), \langle k, t_1 \rangle, q) \\ &= \sum_{p_1, \ldots, p_s \in Q} E^{\mathscr{A}}_{k,s}((\omega_1, \ldots, \omega_k), \langle k, \xi_1, \ldots, \xi_s \rangle, (p_1, \ldots, p_s)) E_s(p_1, \ldots, p_s, \sigma, q). \end{aligned}$$

For $k' > 1$, $\omega_1, \ldots, \omega_k \colon Q \longrightarrow S$, $q_1, \ldots, q_{k'} \in Q$, and $\xi \coloneqq \langle k, t_1, \ldots, t_{k'} \rangle \in T(\Sigma)^{k'}_k$, then ultimately define

$$E^{\mathscr{A}}_{k,k'}((\omega_1, \ldots, \omega_k), \xi, (q_1, \ldots, q_{k'})) = \prod_{i=1}^{k'} E^{\mathscr{A}}_{k,1}((\omega_1, \ldots, \omega_k), \langle k, t_i \rangle, q_i).$$

We will again omit the brackets around tuples of states and sets of states. ■

**Definition 3.2.3.** Let $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be an $(m,n)$-WFA. The **weighted forest language accepted by** $\mathscr{A}$, denoted $\mathscr{L}(\mathscr{A})$, is the weighted forest language

$$\mathscr{L}(\mathscr{A}) \colon T(\Sigma)^m_n \longrightarrow S,$$

given for any $\xi \in T(\Sigma)^m_n$ by

$$\mathscr{L}(\mathscr{A})(\xi) \coloneqq \sum_{f \in F} E^{\mathscr{A}}_{n,m}(I_1, \ldots, I_n, \xi, f),$$

where $I = (I_1, \ldots, I_n)$.

Ultimately, we define the classes

$$\mathrm{REC}(T(\Sigma)^m_n, S) \coloneqq \{\tau \colon T(\Sigma)^m_n \longrightarrow S \mid \exists \mathscr{A}\ (m,n)\text{-WFA} \colon \mathscr{L}(\mathscr{A}) = \tau\}$$

of **recognizable weighted $(m,n)$-forest languages** (for $m, n \in \mathbb{N}_0$) and the class

$$\mathrm{REC}(T(\Sigma), S) \coloneqq \bigcup_{m,n \in \mathbb{N}_0} \mathrm{REC}(T(\Sigma)^m_n, S)$$

of (all) **recognizable weighted forest languages**. ■

**Example 3.2.4.** Let $\Sigma \coloneqq \{\sigma^{(2)}, \gamma^{(1)}, \beta^{(0)}, \alpha^{(0)}\}$ and $S$ an arbitrary semiring. Consider the (2,2)-WFA $\mathscr{A} = (Q, \Sigma, S, I, F, E)$, where

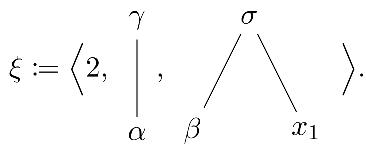
$$Q = \{q_1, q_2, f_1, f_2\}, \qquad F = \{f_1\} \times \{f_1, f_2\}, \qquad I = (\mathbb{1}_{q_1}, 0).$$

Moreover, $E$ is 0 except in the cases

$$\begin{aligned}
&E_0(\alpha, q_1) = E_0(\alpha, q_2) = E_0(\beta, q_1) = 1\\
&E_1(q, \gamma, f) = E_1(f, \gamma, f) = 1\\
&E_2(q, q', \sigma, f) = E_2(f, f, \sigma, f) = 1,\\
&E_2(q, f, \sigma, f) = E_2(f, q, \sigma, f) = 1,
\end{aligned}$$

where $q, q' \in \{q_1, q_2\}$ and $f \in \{f_1, f_2\}$.

Consider the forest



We calculate $\mathscr{L}(\mathscr{A})(\xi)$. For $f = (f_1, f_2) \in F$, it holds that

$$\begin{aligned}
E^{\mathscr{A}}_{2,2}(I, \xi, f) &= E^{\mathscr{A}}_{2,1}(I, \langle 2, \gamma(\alpha), \rangle, f_1) \cdot E^{\mathscr{A}}_{2,1}(I, \langle 2, \sigma(\beta, x_1)\rangle, f_2)\\
&= \Big( \sum_{q \in Q} E_1(q, \gamma, f_1) E^{\mathscr{A}}_{2,1}(I, \langle 2, \alpha\rangle, q) \Big) \cdot E^{\mathscr{A}}_{2,1}(I, \langle 2, \sigma(\beta, x_1)\rangle, f_2)\\
&= \Big( \sum_{q \in Q} 1 \cdot E_0(\alpha, q) \Big) \cdot E^{\mathscr{A}}_{2,1}(I, \langle 2, \sigma(\beta, x_1)\rangle, f_2)\\
&= \Big( E_0(\alpha, q_1) + E_0(\alpha, q_2) \Big) \cdot E^{\mathscr{A}}_{2,1}(I, \langle 2, \sigma(\beta, x_1)\rangle, f_2)\\
&= (2 \cdot 1) \cdot \Big( \sum_{q, q' \in Q} E_2(q, q', \sigma, f_2) E^{\mathscr{A}}_{2,2}(I, \langle 2, \beta, x_1\rangle, (q, q')) \Big)\\
&= (2 \cdot 1) \cdot \Big( \sum_{q, q' \in Q} 1 \cdot E^{\mathscr{A}}_{2,1}(I, \langle 2, \beta\rangle, q) E^{\mathscr{A}}_{2,1}(I, \langle 2, x_1\rangle, q') \Big)\\
&= (2 \cdot 1) \cdot \Big( \sum_{q, q' \in Q} E_0(\beta, q) \mathbb{1}_{q_1}(q') \Big)\\
&= (2 \cdot 1) \cdot 1 = (2 \cdot 1).
\end{aligned}$$

Repeating this process for $f = (f_1, f_1) \in F$ analogously yields

$$E^{\mathscr{A}}_{2,2}(I, \xi, f) = (2 \cdot 1).$$

Alltogether we obtain

$$\mathscr{L}(\mathscr{A})(\xi) = E^{\mathscr{A}}_{2,2}(I, \xi, (f_1, f_1)) + E^{\mathscr{A}}_{2,2}(I, \xi, (f_1, f_2)) = (2 \cdot 1) + (2 \cdot 1) = (4 \cdot 1).$$

This result evaulates in $S = \mathbb{B}$ to $\mathscr{L}(\mathscr{A})(\xi) = 1$ and in $S = \mathbb{N}_0$ to $\mathscr{L}(\mathscr{A})(\xi) = 4$. ■

**Definition 3.2.5.** Let $k, l, m, n \in \mathbb{N}_0$, $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be an $(m,n)$-WFA, $\xi = \langle l, t_1, \ldots, t_k \rangle \in T(\Sigma)^k_l$, $J_1, \ldots, J_l \colon Q \longrightarrow S$, $q_1, \ldots, q_k \in Q$, and $P \subseteq Q$. A **run of $\mathscr{A}$ on $\xi$ using $P$, starting in $J = (J_1, \ldots, J_l)$, and ending in** $q = (q_1, \ldots, q_k)$ is a map $\rho \colon \operatorname{pos}(\xi) \longrightarrow Q$ such that

$$\begin{aligned} \rho((j, \varepsilon)) &= q_j, \text{ and} \\ \rho((j, w)) &\in P, \end{aligned}$$

for every $j \in [k]$ and $w \in \operatorname{pos}(t_j)$ such that $w \neq \varepsilon$ and $\xi((j, w)) \in \Sigma$.

The set of runs of $\mathscr{A}$ on $\xi$ using $P$, starting in $J$, and ending in $q$ is denoted by $R^P_{\mathscr{A}}(J, \xi, q)$. Moreover we denote $R_{\mathscr{A}}(J, \xi, q) \coloneqq R^Q_{\mathscr{A}}(J, \xi, q)$.

If $\rho$ is such a run, we define for any $w = (i, u) \in \operatorname{pos}(\xi)$ the **cost of $w$ in $\xi$ under $\rho$** as

$$c_{\mathscr{A}}(\rho, \xi, w) \coloneqq \begin{cases} E_j(\rho((i, u1), \ldots, (i, uj)), \sigma, \rho(w)) & \text{, if } \operatorname{lab}(\xi, w) = \sigma \in \Sigma^{(j)}, j \geq 0 \\ P_i(\rho(w)) & \text{, if } \operatorname{lab}(\xi, w) = x_i, i \in [l] \end{cases}$$

and the **cost of $\xi$ under $\rho$** as

$$c_{\mathscr{A}}(\rho, \xi) \coloneqq \prod_{i=1}^{k} \prod_{u \in \operatorname{pos}(t_i)} c_{\mathscr{A}}(\rho, \xi, (i, u)),$$

which evaluates to $1 \in S$ for $m = 0$ by convention.

The **weighted forest language run-recognized by $\mathscr{A}$**, denoted $\mathscr{L}_{\text{run}}(\mathscr{A})$, is the weighted forest language

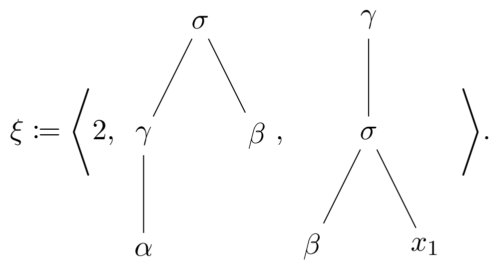
$$\mathscr{L}_{\text{run}}(\mathscr{A}) \colon T(\Sigma)^m_n \longrightarrow S,$$

given for any $\xi \in T(\Sigma)^m_n$ by

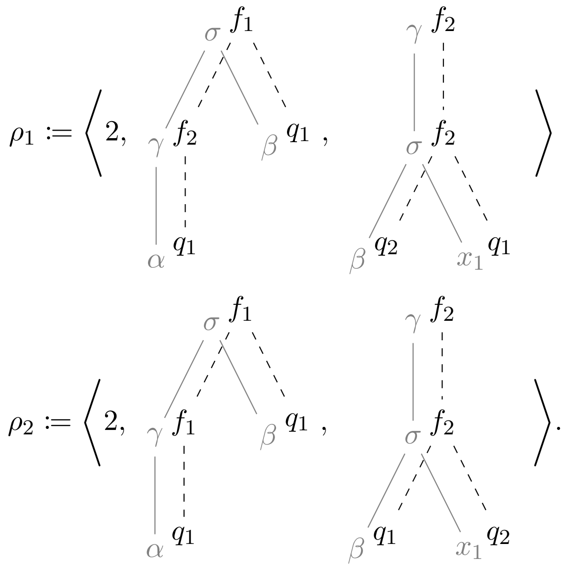
$$\mathscr{L}_{\text{run}}(\mathscr{A})(\xi) \coloneqq \sum_{q \in F} \sum_{\rho \in R_{\mathscr{A}}(I, \xi, q)} c_{\mathscr{A}}(\rho, \xi),$$

which for $m = 0$ evaluates to $1 \in S$ by convention (as there is exactly one run in this case, namely the empty set, which has cost 1). ■

**Example 3.2.6.** Consider $\Sigma$ and $\mathscr{A}$ from Example 3.2.4. Moreover, consider the forest

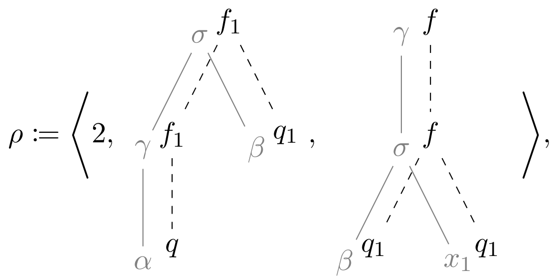



Two exemplary runs of $\mathscr{A}$ on $\xi$ are given by the following dashed trees.




One easily sees that $c_{\mathscr{A}}(\rho_1, \xi, (1, \varepsilon)) = 0$ and hence $c_{\mathscr{A}}(\rho_1, \xi) = 0$. Analogously it holds that $c_{\mathscr{A}}(\rho_2, \xi, (2, 12)) = 0$ and hence $c_{\mathscr{A}}(\rho_2, \xi) = 0$.

In fact, the only runs of $\mathscr{A}$ on $\xi$ that have non-vanishing cost are of the form




where $q \in \{q_1, q_2\}$ and $f \in \{f_1, f_2\}$. Every such run $\rho$ has cost $c_{\mathscr{A}}(\rho, \xi) = 1$ and therefore we obtain

$$\mathscr{L}(\mathscr{A})_{\text{run}}(\xi) = (4 \cdot 1).$$

■

**Remark 3.2.7.** Let $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be an $(m,n)$-WFA.

1) First let $\xi = \langle n, t_1, \ldots, t_m\rangle \in T(\Sigma)^m_n$, $P_1, \ldots, P_n \colon Q \longrightarrow S$, $q_1, \ldots, q_m \in Q$, and $\rho$ a run of $\mathscr{A}$ on $\xi$ starting in $P = (P_1, \ldots, P_n)$ and ending in $q = (q_1, \ldots, q_m)$. Define for any $i \in [m]$ the run of $\mathscr{A}$ on $\xi_i \coloneqq \langle n, t_i\rangle$ starting in $P$ and ending in $q_i$ as the restriction of $\rho$ onto $\xi_i$, denoted $\rho_i$. It holds that

$$\begin{aligned} c_{\mathscr{A}}(\rho, \xi) &= \prod_{i=1}^{m} \prod_{u \in \mathrm{pos}(t_i)} c_{\mathscr{A}}(\rho, \xi, (i, u)) = \prod_{i=1}^{m} \prod_{u \in \mathrm{pos}(t_i)} c_{\mathscr{A}}(\rho_i, \xi_i, (1, u)) \\ &= \prod_{i=1}^{m} \Big( \prod_{j=1}^{1} \prod_{u \in \mathrm{pos}(t_i)} c_{\mathscr{A}}(\rho_i, \xi_i, (j, u)) \Big) = \prod_{i=1}^{m} c_{\mathscr{A}}(\rho_i, \xi_i). \end{aligned}$$

2) Now let $\xi = \langle n, \sigma(t_1, \ldots, t_k)\rangle \in T(\Sigma)^1_n$ for some $k \geq 1$, $P_1, \ldots, P_n \colon Q \longrightarrow S$, $q \in Q$, and $\rho$ a run on $\mathscr{A}$ on $\xi$ starting in $P = (P_1, \ldots, P_n)$ and ending in $q$. Define for any $i \in [k]$ the run of $\mathscr{A}$ on $\xi_i \coloneqq \langle n, t_i\rangle$ starting in $P$ and ending in $\rho((1, i))$ as the restriction of $\rho$ onto $\xi_i$, denoted $\rho_i$. It holds that

$$\begin{aligned} c_{\mathscr{A}}(\rho, \xi) &= \prod_{u \in \mathrm{pos}(\sigma(t_1, \ldots, t_k))} c_{\mathscr{A}}(\rho, \xi, (1, u)) \\ &= \Big( \prod_{i=1}^{k} \prod_{u \in \mathrm{pos}(t_i)} c_{\mathscr{A}}(\rho, \xi, (1, iu)) \Big) c_{\mathscr{A}}(\rho, \xi, (1, \varepsilon)) \\ &= \Big( \prod_{i=1}^{k} \prod_{u \in \mathrm{pos}(t_i)} c_{\mathscr{A}}(\rho_i, \xi_i, (1, u)) \Big) c_{\mathscr{A}}(\rho, \xi, (1, \varepsilon)) \\ &= \Big( \prod_{i=1}^{k} c_{\mathscr{A}}(\rho_i, \xi_i) \Big) c_{\mathscr{A}}(\rho, \xi, (1, \varepsilon)). \end{aligned}$$

These equations will be used in later proofs. ■

**Proposition 3.2.8.** Let $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be an $(m,n)$-WFA. It holds that

$$\mathscr{L}(\mathscr{A}) = \mathscr{L}_{\mathrm{run}}(\mathscr{A}).$$

*Proof.* Let $\xi \in T(\Sigma)^m_n$. By definition we can reduce the claim to proving equation $\times$ in

$$\begin{aligned} \mathscr{L}(\mathscr{A})(\xi) &= \sum_{f \in F} E^{\mathscr{A}}_{n,m}(I_1, \ldots, I_n, \xi, f) \\ &\overset{\times}{=} \sum_{f \in F} \sum_{\rho \in R_{\mathscr{A}}(I, \xi, f)} c_{\mathscr{A}}(\rho, \xi) = \mathscr{L}_{\mathrm{run}}(\mathscr{A})(\xi), \end{aligned}$$

hence it suffices to prove

$$\forall q \in Q^m \colon E^{\mathscr{A}}_{n,m}(I_1, \ldots, I_n, \xi, q) = \sum_{\rho \in R_{\mathscr{A}}(I, \xi, q)} c_{\mathscr{A}}(\rho, \xi), \tag{1}$$

by structural induction on the structure of $\xi$. Thus, let $q \in Q^m$. Note that for $m = 0$, we have forced the desired equation by convention.

*Case 1:* Assume that $m = 1$ and $\xi = \langle n, x_i \rangle$ for some $i \in [n]$. Then,

$$\begin{aligned} E^{\mathscr{A}}_{n,1}(I_1, \ldots, I_n, \xi, q) = I_i(q) &= \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} I_i(q) \\ &= \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} c_{\mathscr{A}}(\rho, \xi, (1, \varepsilon)) = \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} c_{\mathscr{A}}(\rho, \xi), \end{aligned}$$

where we have used that there is exactly one run on $\xi$ starting in $I$ and ending in $q$ (namely the one labeling $x_i$ with $q$).

*Case 2:* Assume that $m = 1$ and $\xi = \langle n, \alpha \rangle$ for some $\alpha \in \Sigma^{(0)}$. Then,

$$\begin{aligned} E^{\mathscr{A}}_{n,1}(I_1, \ldots, I_n, \xi, q) = E_0(\alpha, q) &= \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} E_0(\alpha, q) \\ &= \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} c_{\mathscr{A}}(\rho, \xi, (1, \varepsilon)) = \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} c_{\mathscr{A}}(\rho, \xi), \end{aligned}$$

where we have used that there is exactly one run on $\xi$ starting in $I$ and ending in $q$ (namely the one labeling $\alpha$ with $q$).

*Case 3:* Assume that $m = 1$ and $\xi = \langle n, \sigma(t_1, \ldots, t_k) \rangle$ for some $\sigma \in \Sigma^{(k)}$, $k \geq 1$ such that equation (1) holds for $\xi' := \langle n, t_1, \ldots, t_k \rangle$. Then,

$$\begin{aligned} E^{\mathscr{A}}_{n,1}(I_1, \ldots, I_n, \xi, q) &= \sum_{p=(p_1,\ldots,p_k) \in Q^k} E^{\mathscr{A}}_{n,k}(I_1, \ldots, I_n, \xi', p) E_k(p, \sigma, q) \\ &\overset{\text{IH}}{=} \sum_{p=(p_1,\ldots,p_k) \in Q^k} \sum_{\rho' \in R_{\mathscr{A}}(I,\xi',p)} c_{\mathscr{A}}(\rho', \xi') E_k(p, \sigma, q) \\ &\overset{\star}{=} \sum_{p=(p_1,\ldots,p_k) \in Q^k} \sum_{\rho' \in R_{\mathscr{A}}(I,\xi',p)} c_{\mathscr{A}}(\rho, \xi), \end{aligned}$$

where $\rho$ is the extension of $\rho'$ to $\xi$, given by

$$\begin{aligned} (1, \varepsilon) &\mapsto q, \text{ and} \\ (1, iu) &\mapsto \rho'((i, u)), \text{ for } i \in [k]. \end{aligned}$$

Note that equality $\star$ therefore follows from the equations in Remark 3.2.7. To conclude this case, we show that on the right hand side of equality $\star$, $\rho$ runs over all elements of $R_{\mathscr{A}}(I, \xi, q)$ exactly once, which proves equation (1).

First note that, as in Remark 3.2.7, every $\rho \in R_{\mathscr{A}}(I, \xi, q)$ can be restricted to a run on $\xi'$ starting in $I$ and ending in $(\rho((1, 1)), \ldots, \rho((1, k))) \in Q^k$. Now of course, any two runs on $\xi'$ only extend to the same run on $\xi$, if they are already equal.

*Case 4:* Assume that $m > 1$, $\xi = \langle n, t_1, \dots, t_m \rangle$, and $q = (q_1, \dots, q_m)$ such that equation (1) holds for $\xi_i := \langle n, t_i \rangle$ for any $i \in [m]$. Then,

$$E^{\mathscr{A}}_{n,m}(I_1, \dots, I_n, \xi, q) = \prod_{i=1}^{m} E^{\mathscr{A}}_{n,1}(I_1, \dots, I_n, \xi_i, q_i) = \prod_{i=1}^{m} \sum_{\rho_i \in R_{\mathscr{A}}(I, \xi_i, q_i)} c_{\mathscr{A}}(\rho_i, \xi_i)$$

$$\overset{\star}{=} \sum_{\substack{\rho_i \in R_{\mathscr{A}}(I, \xi_i, q_i), \\ \text{for any } i \in [m]}} \prod_{i=1}^{m} c_{\mathscr{A}}(\rho_i, \xi_i) \overset{\bullet}{=} \sum_{\rho \in R_{\mathscr{A}}(I, \xi, q)} c_{\mathscr{A}}(\rho, \xi).$$

Equality $\star$ simply uses the generalized distributivity law in $S$. Equality $\bullet$ uses the first equation in Remark 3.2.7 and moreover the easy fact that a run on $\xi$ is uniquely determined by its restrictions onto the $\xi_i$.

This concludes the induction and finishes the proof. □

**Example 3.2.9.** Continuing Example 3.2.6, we apply Proposition 3.2.8 to obtain

$$\mathscr{L}(\mathscr{A})(\xi) = \mathscr{L}_{\text{run}}(\mathscr{A})(\xi) = (4 \cdot 1).$$

Let $\xi' = \langle 2, t_1, t_2 \rangle \in T(\Sigma)^2_2$ be an arbitrary forest. It is fairly easy to see (and we will prove later) that, whenever $\text{ht}(t_1) > 1$ and $\text{ht}(t_2) > 1$, we have that

$$\mathscr{L}(\mathscr{A})(\xi') = \underbrace{(2 \cdot 1)}_{\text{root states}} \cdot \underbrace{(2 \cdot 1)^{\#\, \text{pos}_\alpha(\xi')}}_{\text{counting } \alpha\text{s}} \cdot \underbrace{0^{\#\, \text{pos}_{x_2}(\xi')}}_{\text{counting } x_2\text{s}}. \tag{2}$$

If $\text{ht}(t_1) = 1$ or $\text{ht}(t_2) = 1$, we have $\mathscr{L}(\mathscr{A})(\xi') = 0$. This follows from the fact that the root states can not be reached with non-vanishing cost from a leaf in a forest. ■

## 3.3 Decomposition of Weighted Forest Automata

We first show that the recognizable weighted $(1, n)$-forest languages are the recognizable weighted tree languages with variables in $X_n$. After that, we prove that recognizable weighted $(m, n)$-forest languages are products of ($m$ many) recognizable weighted $(1, n)$-forest languages. Or in short, recognizable weighted $(m, n)$-forest languages are rectangular.

In this subchapter, $S$ denotes a commutative semiring.

**Proposition 3.3.1.** Let $n \in \mathbb{N}_0$. It holds that

$$\text{REC}(T(\Sigma)^1_n, S) \cong \text{REC}(T_\Sigma(X_n), S).$$

Recall that $\text{REC}(T_\Sigma(X_n), S) = \text{REC}(T_{\Sigma \cup X_n}, S)$. This says that the recognizable weighted (1,$n$)-forest languages are exactly the recognizable weighted tree languages over $\Sigma \cup X_n$ and $S$ (up to an identification of $T(\Sigma)^1_n$ and $T_\Sigma(X_n)$).

*Proof.* Let $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be a (1,$n$)-WFA. We define the WTA

$$\mathscr{B} \coloneqq (Q, \Sigma \cup X_n, F, \delta)$$

over $S$ via

$$\delta_\sigma(q_1, \ldots, q_k, q) \coloneqq E_k(q_1, \ldots, q_k, \sigma, q),$$

for any $q_1, \ldots, q_k, q \in Q$ and $\sigma \in \Sigma^{(k)}$, $k \geq 0$, and

$$\delta_{x_i}(q) \coloneqq I_i(q),$$

for any $q \in Q$ and $i \in [n]$.

Let $\xi = \langle n, t\rangle \in T(\Sigma)^1_n$. Using Proposition 3.2.8 and identifying $\xi$ with $t \in T_\Sigma(X_n)$, we find that only equation $\times$ in

$$\begin{aligned}\mathscr{L}(\mathscr{A})(\xi) &= \sum_{q \in F} \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} c_{\mathscr{A}}(\rho, \xi) \\ &\overset{\times}{=} \sum_{q \in F} \sum_{\rho \in R_{\mathscr{B}}(t,q)} c_{\mathscr{B}}(\rho, t) = \mathscr{L}(\mathscr{B})(\xi)\end{aligned}$$

needs to be verified. Of course, the sets $R_{\mathscr{A}}(I, \xi, q)$ and $R_{\mathscr{B}}(t, q)$ are equal, up to identification of $\xi$ and $t$. Therefore, we show that for any $q \in F$ and $\rho \in R_{\mathscr{A}}(I, \xi, q)$ (where the corresponding run of $\mathscr{B}$ on $t$ is denoted by $\rho'$), $c_{\mathscr{A}}(\rho, \xi) = c_{\mathscr{B}}(\rho', t)$ holds. This is by definition equivalent to

$$\forall w \in \mathrm{pos}(t)\colon c_{\mathscr{A}}(\rho, \xi, (1, w)) = c_{\mathscr{B}}(\rho', t, w).$$

The equality surely holds for $\mathrm{lab}(t, w) \in \Sigma$. If $\mathrm{lab}(t, w) = x_i$, we have

$$c_{\mathscr{A}}(\rho, \xi, (1, w)) = I_i(\rho((1, w))) = I_i(\rho'(w)) = c_{\mathscr{B}}(\rho', t, w),$$

by definition.

Now let $\mathscr{B} = (Q, \Sigma \cup X_n, F, \delta)$ be a WTA over $S$. We define the (1,$n$)-WFA

$$\mathscr{A} \coloneqq (Q, \Sigma, S, I, F, E),$$

where $I = (I_1, \ldots, I_n)$ for $I_i\colon Q \longrightarrow S$, which is defined as $I_i \coloneqq \delta_{x_i}$, for any $i \in [n]$, and

$$E_k(q_1, \ldots, q_k, \sigma, q) \coloneqq \delta_\sigma(q_1, \ldots, q_k, q),$$

for any $q_1, \ldots, q_k, q \in Q$, and $\sigma \in \Sigma^{(k)}$, $k \geq 0$.

We now prove $\mathscr{L}(\mathscr{A}) = \mathscr{L}(\mathscr{B})$. By the first part of this proof, we obtain an automaton $\mathscr{B}'$ such that $\mathscr{L}(\mathscr{A}) = \mathscr{L}(\mathscr{B}')$. However, the construction shows that $\mathscr{B} = \mathscr{B}'$, which proves the claim. □

**Proposition 3.3.2.** Let $m \in \mathbb{N}$, $n \in \mathbb{N}_0$, and $\tau\colon T(\Sigma)^m_n \longrightarrow S$. It holds that

$$\tau \in \mathrm{REC}(T(\Sigma)^m_n, S) \Longleftrightarrow \exists \tau_1, \dots, \tau_m \in \mathrm{REC}(T(\Sigma)^1_n, S)\colon \tau = \prod_{i=1}^{m} \Big(\tau_i \circ \pi^m_i\Big). \tag{3}$$

Here, $\pi^m_i$ denotes by abuse of notation the map

$$\pi^m_i\colon T(\Sigma)^m_n \longrightarrow T(\Sigma)^1_n,$$
$$\xi \mapsto \pi^m_i \cdot \xi.$$

*Proof.* "$\Longrightarrow$": Let $\tau \in \mathrm{REC}(T(\Sigma)^m_n, S)$. By definition, there exists an $(m,n)$-WFA $\mathscr{A} = (Q, \Sigma, S, I, (F_1, \dots, F_m), E)$ such that $\mathscr{L}(\mathscr{A}) = \tau$. We define for any $i \in [m]$ the $(1,n)$-WFA

$$\mathscr{A}_i \coloneqq (Q, \Sigma, S, I, F_i, E)$$

and claim that the weighted forest language $\tau_i \coloneqq \mathscr{L}(\mathscr{A}_i) \in T(\Sigma)^1_n$ satisfy the right hand side in (3).

To prove this claim, take $\xi \in T(\Sigma)^m_n$ and define $\xi_i \coloneqq \pi^m_i \cdot \xi$. For any $q \in F$ and $\rho \in R_{\mathscr{A}}(I, \xi, q)$, we use the notation from point 1) in Remark 3.2.7 to denote by $\rho_i$ the restriction of $\rho$ to $\xi_i$, ending in $q_i$ (and vice versa). It holds that

$$\begin{aligned}
\tau(\xi) &= \sum_{q \in F} \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} c_{\mathscr{A}}(\rho, \xi) \\
&\overset{\star_1}{=} \sum_{q \in F} \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} \prod_{i=1}^{m} c_{\mathscr{A}}(\rho_i, \xi_i) \\
&\overset{\star_2}{=} \prod_{i=1}^{m} \sum_{q_i \in F_i} \sum_{\rho_i \in R_{\mathscr{A}}(I,\xi_i,q_i)} c_{\mathscr{A}}(\rho_i, \xi_i) \\
&\overset{\star_3}{=} \prod_{i=1}^{m} \sum_{q_i \in F_i} \sum_{\rho_i \in R_{\mathscr{A}_i}(I,\xi_i,q_i)} c_{\mathscr{A}_i}(\rho_i, \xi_i) = \prod_{i=1}^{m} \tau_i(\xi_i).
\end{aligned}$$

In equation $\star_1$, we use point 1) from Remark 3.2.7. Equation $\star_2$ uses the fact that $F$ is a direct product of sets and $R_{\mathscr{A}}(I, \xi, q)$ is bijective to $\bigtimes_{i=1}^m R_{\mathscr{A}}(I, \xi_i, q_i)$, which in turn is again a direct product of sets. We can therefore apply the generalized distributivity law, which proves equation $\star_2$. In equation $\star_3$, we use that weights in $\mathscr{A}$ equal weights in $\mathscr{A}_i$ and runs on $\xi_i$ in $\mathscr{A}$ equal runs on $\xi_i$ in $\mathscr{A}_i$.

"$\Longleftarrow$": Let $\tau_1, \dots, \tau_m \in \mathrm{REC}(T(\Sigma)^1_n, S)$. By definition, there exists for any $i \in [m]$ a $(1,n)$-WFA $\mathscr{A}_i = (Q_i, \Sigma, S, I_i, F_i, E_i)$ such that $\tau_i = \mathscr{L}(\mathscr{A}_i)$. Without loss of generality, we can impose that the $(Q_i \mid i \in [m])$ are pairwise disjoint. Let moreover $I_i = (I_{i,1}, \dots, I_{i,n})$ and $E_i = (E_{i,k} \mid k \in \mathbb{N}_0)$ for any $i \in [m]$.

We define the $(m,n)$-WFA

$$\mathscr{A} \coloneqq (Q, \Sigma, S, I, F, E),$$

where $Q \coloneqq \bigcup_{i=1}^{m} Q_i$ is the disjoint union of the given state sets, $F \coloneqq F_1 \times \cdots \times F_m$, $I \coloneqq (I^1, \ldots, I^n)$ for leaf weights $I^j \colon Q \longrightarrow S$ defined via

$$I^j(q) \coloneqq I_{i,j}(q), \text{ if } \exists i \in [m] \colon q \in Q_i,$$

and $E = (E^k \mid k \in \mathbb{N}_0)$, where $E^k \colon Q^k \times \Sigma^{(k)} \times Q \longrightarrow S$ is defined for any $q_1, \ldots, q_k, q \in Q$ and $\sigma \in \Sigma^{(k)}$ as

$$E^k(q_1, \ldots, q_k, \sigma, q) \coloneqq \begin{cases} E_{i,k}(q_1, \ldots, q_k, \sigma, q) & \text{, if } \exists i \in [m] \colon q_1, \ldots, q_k, q \in Q_i \\ 0 & \text{, otherwise.} \end{cases}$$

We claim that the weighted forest language $\tau \coloneqq \mathscr{L}(\mathscr{A})$ satisfies $\tau = \prod_{i=1}^{m} \left(\tau_i \circ \pi_i^m\right)$.

To prove this claim, take $\xi \in T(\Sigma)_n^m$ and define $\xi_i \coloneqq \pi_i^m \cdot \xi$. In essence, we simply repeat the proof from "$\Longrightarrow$", where equation $\star_3$ is replaced by

$$\prod_{i=1}^{m} \sum_{q_i \in F_i} \sum_{\rho_i \in R_{\mathscr{A}}(I, \xi_i, q_i)} c_{\mathscr{A}}(\rho_i, \xi_i) = \prod_{i=1}^{m} \sum_{q_i \in F_i} \sum_{\rho_i \in R_{\mathscr{A}_i}(I_i, \xi_i, q_i)} c_{\mathscr{A}_i}(\rho_i, \xi_i). \tag{4}$$

Note however that, in this case, the sets $R_{\mathscr{A}}(I, \xi_i, q_i)$ and $R_{\mathscr{A}_i}(I_i, \xi_i, q_i)$ are not isomorphic, as $\mathscr{A}$ contains vastly more states than $\mathscr{A}_i$. This is accounted for by the vanishing state transition weights and leaf weights in $\mathscr{A}$.

We now prove equation (4).

Let $i \in [m]$ and $q_i \in F_i$. We call a run $\rho \in R_{\mathscr{A}}(I, \xi_i, q_i)$ **$Q_i$-restricted** if for any $w \in \text{pos}(\xi_i)$ it holds that $\rho(w) \in Q_i$. The set of $Q_i$-restricted runs of $\mathscr{A}$ on $\xi_i$ starting in $I$ and ending in $q_i$ is denoted $R_{\mathscr{A}}^i(I, \xi_i, q_i)$.

Let $\rho \in R_{\mathscr{A}}(I, \xi_i, q_i) \setminus R_{\mathscr{A}}^i(\xi_i, I, q_i)$. It then holds that there exists $w \in \text{pos}(\xi_i)$ such that $\rho(w) \notin Q_i$. Moreover, as $\rho((1, \varepsilon)) = q_i \in Q_i$, $w$ can be chosen to be of the form $w = (1, ul)$, where

$$\text{lab}(\xi_i, (1, u)) = \sigma \in \Sigma^{(j)}$$

for some $j \geq 1$ and $l \in [j]$, such that $\rho((1, u)) \in Q_i$. Therefore

$$c_{\mathscr{A}}(\rho, \xi_i, w) = E^j(\rho((1, u1), \ldots, \rho(1, uj)), \sigma, \rho(u)) = 0$$

by definition of $E^j$, which proves

$$\sum_{\rho_i \in R_{\mathscr{A}}(I, \xi_i, q_i)} c_{\mathscr{A}}(\rho_i, \xi_i) = \sum_{\rho_i \in R_{\mathscr{A}}^i(I, \xi_i, q_i)} c_{\mathscr{A}}(\rho_i, \xi_i).$$

The last step is now to show that $Q_i$-restricted runs of $\mathscr{A}$ on $\xi_i$ correspond one-to-one to runs of $\mathscr{A}_i$ on $\xi_i$ where corresponding runs have equal weights.

By definition of runs, the sets $R_{\mathscr{A}}^i(\xi_i, I, q_i)$ and $R_{\mathscr{A}_i}(I_i, \xi_i, q_i)$ are in fact equal, so we only need to show that for any such run $\rho$ and any $w \in \text{pos}(\xi_i)$, it holds that $c_{\mathscr{A}}(\rho, \xi_i) = c_{\mathscr{A}_i}(\rho, \xi_i)$. This is indeed true, as

$$\begin{aligned} c_{\mathscr{A}}(\rho, \xi_i, w) &= E^j(\rho((1, u1), \ldots, (1, uj)), \sigma, \rho(w)) \\ &= E_{i,j}(\rho((1, u1), \ldots, (1, uj)), \sigma, \rho(w)) = c_{\mathscr{A}_i}(\rho, \xi_i, w) \end{aligned}$$

if $\mathrm{lab}(\xi_i, w) = \sigma \in \Sigma^{(j)}$ for some $j \geq 0$, and

$$c_{\mathscr{A}}(\rho, \xi_i, w) = I^k(\rho(w)) = I_{i,k}(\rho(w)) = c_{\mathscr{A}_i}(\rho, \xi_i, w)$$

if $\mathrm{lab}(\xi_i, w) = x_k$ and $k \in [n]$.

This concludes the proof, as we have shown

$$\begin{aligned}\prod_{i=1}^{m} \sum_{q_i \in F_i} \sum_{\rho_i \in R_{\mathscr{A}}(I, \xi_i, q_i)} c_{\mathscr{A}}(\rho_i, \xi_i) &= \prod_{i=1}^{m} \sum_{q_i \in F_i} \sum_{\rho_i \in R^i_{\mathscr{A}}(I, \xi_i, q_i)} c_{\mathscr{A}}(\rho_i, \xi_i) \\ &= \prod_{i=1}^{m} \sum_{q_i \in F_i} \sum_{\rho_i \in R_{\mathscr{A}_i}(I_i, \xi_i, q_i)} c_{\mathscr{A}_i}(\rho_i, \xi_i).\end{aligned}$$

□

**Remark 3.3.3.** Note that Proposition 3.3.2 can be seen as a techincal version of the equality

$$\mathrm{REC}(T(\Sigma)^m_n, S) = \mathrm{REC}(T(\Sigma)^1_n, S)^m.$$

■

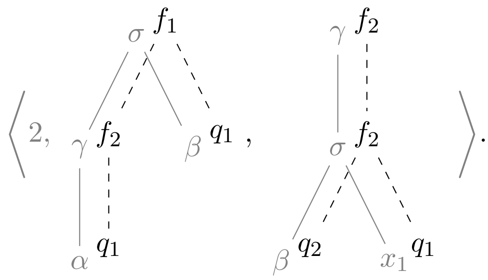
**Example 3.3.4.** Consider $\Sigma$ and $\mathscr{A}$ from Example 3.2.4. By Proposition 3.3.2, there are (1,2)-WFA $\mathscr{A}_1$ and $\mathscr{A}_2$ such that

$$\mathscr{L}(\mathscr{A}) = (\mathscr{L}(\mathscr{A}_1) \circ \pi^2_1) \times (\mathscr{L}(\mathscr{A}_2) \circ \pi^2_2).$$

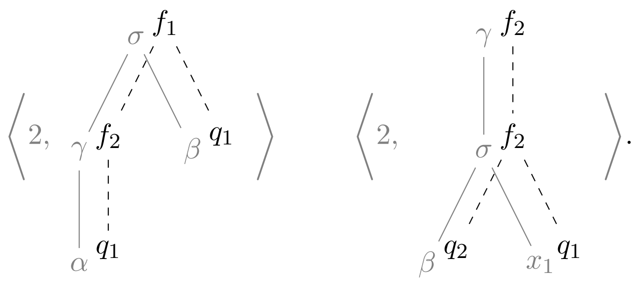
The proof of Proposition 3.3.2 gives us that

$$\mathscr{A}_1 = (Q, \Sigma, S, I, F_1, E) \qquad \text{and} \qquad \mathscr{A}_2 = (Q, \Sigma, S, I, F_2, E).$$

The following forest $\xi$ (gray) has the depicted run $\rho$ (black, dashed)



$\rho$ decomposes into the following runs $\rho_1$ (left, black) and $\rho_2$ (right, black) of $\mathscr{A}_1$ and $\mathscr{A}_2$ on the respective (1,2)-forests.



It is now a matter of multiplying transition weights to determine that $c_{\mathscr{A}}(\rho, \xi) = 0 = 0 \cdot 1 = c_{\mathscr{A}_1}(\rho_1, \xi_1) \cdot c_{\mathscr{A}_2}(\rho_2, )$.

In general, it holds that for every $\xi_1, \xi_2 \in T(\Sigma)_2^1$ such that $\mathrm{ht}(\xi_1) > 1$ and $\mathrm{ht}(\xi_2) > 1$, we have

$$\begin{aligned} \mathscr{L}(\mathscr{A}_1)(\xi_1) &= (2 \cdot 1)^{\#\,\mathrm{pos}_\alpha(\xi_1)} \cdot 0^{\#\,\mathrm{pos}_{x_2}(\xi_1)}, \text{ and} \\ \mathscr{L}(\mathscr{A}_2)(\xi_2) &\overset{\star}{=} (2 \cdot 1) \cdot (2 \cdot 1)^{\#\,\mathrm{pos}_\alpha(\xi_2)} \cdot 0^{\#\,\mathrm{pos}_{x_2}(\xi_2)}. \end{aligned}$$

Both equations are derived from the fact that $\mathscr{A}_1$ and $\mathscr{A}_2$ are in essence the WTA $\mathscr{A}$ from Example 2.3.1 extended to $\Sigma \cup X_2$. Moreover, equation $\star$ uses the fact that $\mathscr{A}_2$ has two final states $f_1$ and $f_2$, which can not simultaneously occur within a single run with non-vanishing cost. Therefore, $\mathscr{L}(\mathscr{A}_2) = (2 \cdot 1) \cdot \mathscr{L}(\mathscr{A}_1)$.

For $i \in [2]$ we have that $\mathrm{ht}(\xi_i) = 1$ implies $\mathscr{L}(\mathscr{A}_i)(\xi_i) = 0$. In total, this verifies equation (2). ■

## 3.4 Normal Forms for Weighted Forest Automata

We first introduce a “root state normal form” for WFA, isolating the root states from the remaining transition weights. For a WFA $\mathscr{A}$ in root state normal form, any run of $\mathscr{A}$ on a forest $\xi$ with non-vanishing cost can only label the roots of $\xi$ with root states.

We then introduce a respective “leaf state normal form”. In this normal form, only variables can be labeled with states that have non-vanishing leaf weight.

**Definition 3.4.1.** Let $m, n \in \mathbb{N}_0$ and $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be an $(m,n)$-WFA. We say that $\mathscr{A}$ is in **root state normal form** if $F = \{(f_1, \dots, f_m)\}$ for some distinct states $f_1, \dots, f_m \in Q$ and moreover for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q_1, \dots, q_k, q \in Q$ it holds that

$$E_k(q_1, \dots, q_k, \sigma, q) = 0, \text{ if } \exists i \in [k], j \in [m] : q_i = f_j.$$

This states that there is a single root state tuple of distinct states which only “occur” on the roots of a forest. ■

**Proposition 3.4.2.** Let $m, n \in \mathbb{N}_0$ and $\mathscr{A}$ be an $(m,n)$-WFA. There exists an $(m,n)$-WFA $\mathscr{B}$ in root state normal form such that

$$\mathscr{L}(\mathscr{A}) = \mathscr{L}(\mathscr{B}).$$

*Proof. Case 1, $m = 1$:* Let $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ and let $f$ be a fresh state symbol ($f \notin Q$). We define the (1,$n$)-WFA

$$\mathscr{B} := (Q', \Sigma, S, I', \{f\}, E'),$$

where $Q' := Q \cup \{f\}$, $I' = (I'_1, \ldots, I'_n)$ such that for any $i \in [n]$ and $q \in Q'$,

$$I'_i(q) := \begin{cases} I_i(q) & \text{, if } q \in Q \\ \sum_{e \in F} I_i(e) & \text{, if } q = f, \end{cases}$$

and $E' = (E'_k \mid k \geq 0)$ such that for any $k \geq 0$, $q_1, \ldots, q_k, q \in Q'$, and $\sigma \in \Sigma^{(k)}$,

$$E'_k(q_1, \ldots, q_k, \sigma, q) := \begin{cases} E_k(q_1, \ldots, q_k, \sigma, q) & \text{, if } q_1, \ldots, q_k, q \in Q \\ \sum_{e \in F} E_k(q_1, \ldots, q_k, \sigma, e) & \text{, if } q_1, \ldots, q_k \in Q \wedge q = f \\ 0 & \text{, otherwise.} \end{cases}$$

It immediately follows that $\mathscr{B}$ is in root state normal form.

Let $\xi \in T(\Sigma)^1_n$. We say a run $\rho \in R_{\mathscr{B}}(I', \xi, f)$ is **$f$-free** if for every $w \in \text{pos}(\xi) \setminus \{(1, \varepsilon)\}$ it holds that $\rho(w) \neq f$.

Let $\rho \in R_{\mathscr{B}}(I', \xi, f)$ be a run that is not $f$-free. There exists a position $w = (1, u) \in \text{pos}(\xi)$ and $j \in \mathbb{N}$ such that $(1, uj) \in \text{pos}(\xi)$ and $\rho((1, uj)) = f$. This implies

$$c_{\mathscr{B}}(\rho, \xi, w) = 0$$

and thus $c_{\mathscr{B}}(\rho, \xi) = 0$.

Let $\rho \in R_{\mathscr{B}}(I', \xi, f)$ be an $f$-free run and let $e \in F$. We define the run $\rho_e \in R_{\mathscr{A}}(I, \xi, e)$ by

$$\begin{aligned} \rho_e((1, \varepsilon)) &:= e && \text{, and} \\ \rho_e(w) &:= \rho(w) && \text{, for every } w \neq (1, \varepsilon). \end{aligned}$$

Using this definition, we see that for every $\rho \in R_{\mathscr{B}}(I', \xi, f)$ it holds that

$$\begin{aligned} c_{\mathscr{B}}(\rho, \xi) &= \prod_{w \in \text{pos}(\xi)} c_{\mathscr{B}}(\rho, \xi, w) \\ &= c_{\mathscr{B}}(\rho, \xi, (1, \varepsilon)) \prod_{\substack{w \in \text{pos}(\xi) \\ w \neq (1, \varepsilon)}} c_{\mathscr{B}}(\rho, \xi, w) \\ &= \Big(\sum_{e \in F} E_k(q_1, \ldots, q_k, \sigma, e)\Big) \prod_{\substack{w \in \text{pos}(\xi) \\ w \neq (1, \varepsilon)}} c_{\mathscr{B}}(\rho, \xi, w) \\ &\overset{\star}{=} \sum_{e \in F} \left( E_k(q_1, \ldots, q_k, \sigma, e) \prod_{\substack{w \in \text{pos}(\xi) \\ w \neq (1, \varepsilon)}} c_{\mathscr{A}}(\rho_e, \xi, w) \right) \\ &= \sum_{e \in F} c_{\mathscr{A}}(\rho_e, \xi). \end{aligned}$$

In equation $\star$ we have used that the costs of $\rho$ and $\rho_e$ at positions $\neq (1,\varepsilon)$ are equal by definition. Note that the map

$$R_{\mathscr{B}}(I',\xi,f) \times F \longrightarrow \bigcup_{e\in F} R_{\mathscr{A}}(I,\xi,e)$$
$$(\rho,e) \mapsto \rho_e$$

is a bijection.

This correspondence of runs implies

$$\begin{aligned}\mathscr{L}(\mathscr{B})(\xi) &= \sum_{\rho\in R_{\mathscr{B}}(I',\xi,f)} c_{\mathscr{B}}(\rho,\xi) = \sum_{\substack{\rho\in R_{\mathscr{B}}(I',\xi,f)\\ f\text{-free}}} c_{\mathscr{B}}(\rho,\xi)\\ &= \sum_{\substack{\rho\in R_{\mathscr{B}}(I',\xi,f)\\ f\text{-free}}} \sum_{e\in F} c_{\mathscr{A}}(\rho_e,\xi) = \sum_{e\in F} \sum_{\rho_e\in R_{\mathscr{A}}(I,\xi,e)} c_{\mathscr{A}}(\rho_e,\xi) = \mathscr{L}(\mathscr{A})(\xi)\end{aligned}$$

for every $\xi \in T(\Sigma)^1_n$.

*Case 2, $m > 1$:* By Proposition 3.3.2 there exist (1,$n$)-WFA $\mathscr{A}_1,\dots,\mathscr{A}_m$ such that

$$\mathscr{L}(\mathscr{A}) = \prod_{i=1}^{m} \Big(\mathscr{L}(\mathscr{A}_i) \circ \pi_i^m\Big).$$

For any $i \in [n]$, we can apply case 1 to $\mathscr{A}_i$, whence we obtain a (1,$n$)-WFA $\mathscr{B}_i$ in root state normal form such that $\mathscr{L}(\mathscr{B}_i) = \mathscr{L}(\mathscr{A}_i)$.

Without loss of generality, we can assume that the state sets of the $\mathscr{B}_i$ are pairwise disjoint. Iteratively applying Proposition 4.5.2 to the $\mathscr{B}_i$, we obtain an ($m$,$n$)-WFA $\mathscr{B}$ such that

$$\mathscr{L}(\mathscr{B}) = \prod_{i=1}^{m} \Big(\mathscr{L}(\mathscr{B}_i) \circ \pi_i^m\Big).$$

Since this implies $\mathscr{L}(\mathscr{A}) = \mathscr{L}(\mathscr{B})$, it only remains to show that $\mathscr{B}$ is in root state normal form. By construction, $\mathscr{B}$ has a single root state tuple and the components are by assumption distinct (the $\mathscr{B}_i$ have pairwise disjoint sets of states). It again follows immediately from the construction, that the state transition weights vanish, if any of the $q_i$ equals some $f_j$. This concludes the proof. □

**Example 3.4.3.** Consider $\Sigma$ and $\mathscr{A}$ from Example 3.2.4 and $\mathscr{A}_1$ and $\mathscr{A}_2$ from Example 3.3.4. Recall that $\mathscr{L}(\mathscr{A}_1)$ and $\mathscr{L}(\mathscr{A}_2)$ are the rectangular components of $\mathscr{L}(\mathscr{A})$.

As runs of $\mathscr{A}$ on a forest $\xi \in T(\Sigma)^2_2$ propagate root states through the trees in $\xi$, we find that $\mathscr{A}$ is not in root state normal form. We apply the construction from Proposition 3.4.2. In order to do this, we first consider $\mathscr{A}_1$ and $\mathscr{A}_2$, apply the construction to them and then horizontally concatenate the results. This gives a $(2,2)$-WFA $\mathscr{A}'$ in root state normal form such that $\mathscr{L}(\mathscr{A}) = \mathscr{L}(\mathscr{A}')$. As this construction is fairly intuitive, we simply provide the solutions and illustrate the runs of $\mathscr{A}'$ in comparison to the runs of $\mathscr{A}$.

Let $f$ be a fresh state symbol. The transition weights of the $(1,2)$-WFA

$$\mathscr{A}_1' = (Q \cup \{f\}, \Sigma, S, I, \{f\}, E_1') \qquad \text{and} \qquad \mathscr{A}_2' = (Q \cup \{f\}, \Sigma, S, I, \{f\}, E_2')$$

are defined as follows. Let $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $p_1, \ldots, p_k \in Q$. For every $q \in Q$ we have

$$(E_1')_k(p_1, \ldots, p_k, \sigma, q) = E_k(p_1, \ldots, p_k, \sigma, q) = (E_2')_k(p_1, \ldots, p_k, \sigma, q),$$

and the remaining case is given by

$$\begin{aligned}
&(E_1')_k(p_1, \ldots, p_k, \sigma, f) = E_k(p_1, \ldots, p_k, \sigma, f_1),\\
&(E_2')_k(p_1, \ldots, p_k, \sigma, f) = E_k(p_1, \ldots, p_k, \sigma, f_1) + E_k(p_1, \ldots, p_k, \sigma, f_2).
\end{aligned}$$

Applying the construction of Proposition 4.5.2 results in

$$\mathscr{A}' = (Q', \Sigma, S, I', F', E'),$$

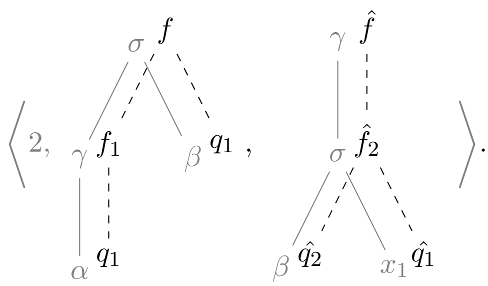
where the components are the following.

$$Q' = Q \cup \{f\} \cup \{\hat{q} \mid q \in Q\} \cup \{\hat{f}\} \qquad I' = (\mathbb{1}_{q_1} + \mathbb{1}_{\hat{q}_1}, 0) \qquad F' = \{(f, \hat{f})\}$$

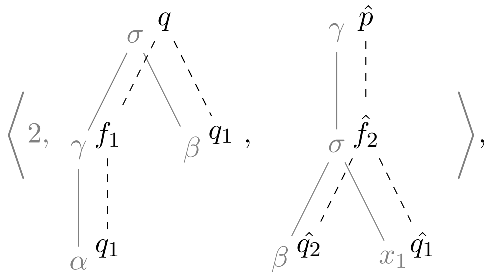
and for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $p_1, \ldots, p_k, q \in Q$ we have

$$\begin{aligned}
E_k'(p_1, \ldots, p_k, \sigma, q) &= E_k(p_1, \ldots, p_k, \sigma, q)\\
E_k'(\hat{p_1}, \ldots, \hat{p_k}, \sigma, \hat{q}) &= E_k(p_1, \ldots, p_k, \sigma, q)\\
E_k'(p_1, \ldots, p_k, \sigma, f) &= E_k(p_1, \ldots, p_k, \sigma, f_1)\\
E_k'(\hat{p_1}, \ldots, \hat{p_k}, \sigma, \hat{f}) &= E_k(p_1, \ldots, p_k, \sigma, f_1) + E_k(p_1, \ldots, p_k, \sigma, f_2).
\end{aligned}$$

The following exemplary run $\rho$ (depicted slightly above the forest) of $\mathscr{A}'$ on $\xi$ (depicted in gray)




corresponds to all runs of $\mathscr{A}$ on $\xi$ of the form



where $q \in Q$ and $\hat{p} \in \hat{Q}$. In this case, these runs of $\mathscr{A}$ on $\xi$ can only have non-vanishing cost, if $q = f_1$ and $\hat{p} = \hat{f}_2$. ■

**Definition 3.4.4.** Let $\varphi\colon T(\Sigma)_1^1 \longrightarrow S$. We call $\varphi$ **proper** if it holds that

$$\varphi(\pi_1^1) = 0.$$

■

**Definition 3.4.5.** Let $m \in \mathbb{N}_0$ and $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be an $(m,1)$-WFA. We say that $\mathscr{A}$ is in **leaf state normal form** if there exists $q^I \in Q$ such that

$$I_1 = \mathbb{1}_{q^I}$$

and for every $k \geq 0$, $q_1, \ldots, q_k \in Q$, and $\sigma \in \Sigma^{(k)}$ it holds that

$$E_k(q_1, \ldots, q_k, \sigma, q^I) = 0.$$

That is, $q^I$ is the unique leaf state for variable $x_1$ and does not occur on right hand sides of transitions in $\mathscr{A}$.

We call $\mathscr{A}$ **normalized** if $\mathscr{A}$ is in leaf state normal form and in root state normal form. ■

**Proposition 3.4.6.** Let $\varphi \in \mathrm{REC}(T(\Sigma)_1^1, S)$ proper. There exists a normalized (1,1)-WFA $\mathscr{B}$ such that

$$\mathscr{L}(\mathscr{B}) = \varphi.$$

*Proof.* Let $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be a (1,1)-WFA in root state normal form such that $\mathscr{L}(\mathscr{A}) = \varphi$. Define the (1,1)-WFA

$$\mathscr{B} \coloneqq (Q_{\mathscr{B}}, \Sigma, S, I_{\mathscr{B}}, F, E_{\mathscr{B}}),$$

where $Q_{\mathscr{B}} \coloneqq Q \dot\cup \{q^I\}$ for a fresh state symbol $q^I \notin Q$, $I_{\mathscr{B}} \coloneqq (\mathbb{1}_{q^I})$ and $E_{\mathscr{B}} \coloneqq (E_k^{\mathscr{B}} \mid k \geq 0)$. For every $k \geq 0$, $q_1, \ldots, q_k, q \in Q_{\mathscr{B}}$, and $\sigma \in \Sigma^{(k)}$ we define

$$E_k^{\mathscr{B}}(q_1, \ldots, q_k, \sigma, q) \coloneqq E_k(q_1, \ldots, q_k, \sigma, q),$$

whenever $q_1, \ldots, q_k, q \in Q$. Furthermore, if $\emptyset \neq \{i \in [k] \mid \exists j \in [n] \colon q_i = q^I\} =: K$ and $q \in Q$, then we define

$$E_k^{\mathscr{B}}(q_1, \ldots, q_k, \sigma, q) \coloneqq \sum_{\substack{q'_a \in Q \\ \forall a \in K}} \Big( \prod_{i \in K} I_1(q'_i) \Big) E_k(q'_1, \ldots, q'_k, \sigma, q).$$

Here we denoted $q'_i \coloneqq q_i$ for every $i \in [k] \setminus K$. Moreover define

$$E_k^{\mathscr{B}}(q_1, \ldots, q_k, \sigma, q) \coloneqq 0$$

in any other case.

It is clear that $\mathscr{B}$ is in leaf state normal form. Moreover, as $\mathscr{A}$ is in root state normal form (with root state $q_f$), we can show that $\mathscr{B}$ is normalized as follows. If $q_f$ occurs on the left hand side of a transition $E_k^{\mathscr{B}}(q_1, \ldots, q_k, \sigma, q)$, its weight is either directly a vanishing state transition weight or a sum over vanishing state transition weights. Therefore, $\mathscr{B}$ is in root state normal form and hence normalized.

Now we show that $\mathscr{L}(\mathscr{A}) = \mathscr{L}(\mathscr{B})$. In order to do this, we introduce the following definition.

Let $\xi \in T(\Sigma)^1_n$, $q \in Q_{\mathscr{B}}$, and $\rho \in R_{\mathscr{B}}(I_{\mathscr{B}}, \xi, q)$. We say $\rho$ is **initial** if

$$\forall w \in \mathrm{pos}(\xi) \colon (\mathrm{lab}(\xi, w) = x_1 \iff \rho(w) = q^I).$$

If $\rho$ is not initial, there either exists $w \in \mathrm{pos}(\xi)$ such that $\mathrm{lab}(\xi, w) = x_1$ and $\rho(w) \neq q^I$ or $\mathrm{lab}(\xi, w) \neq x_1$ and $\rho(w) = q^I$. In any case, $c_{\mathscr{B}}(\rho, \xi, w) = 0$ and therefore $c_{\mathscr{B}}(\rho, \xi) = 0$.

For every initial $\rho \in R_{\mathscr{B}}(I_{\mathscr{B}}, \xi, q)$, we furthermore define the set $R^\rho_{\mathscr{A}}(\xi, I, q)$ consting of all runs $\rho'$ of $\mathscr{A}$ on $\xi$ that can be obtained from $\rho$ by replacing every occurrence of $q^I$ by arbitrary states in $Q$. It follows immediately that

$$\bigcup_{\substack{\rho \in R_{\mathscr{B}}(I_{\mathscr{B}}, \xi, q_f), \\ \rho \text{ initial}}} R^\rho_{\mathscr{A}}(\xi, I, q) = R_{\mathscr{A}}(I, \xi, q),$$

which is moreover a disjoint union.

To conclude the proof, it only remains to show equation $\divideontimes$ in the following chain of equations. For any $\xi \in T(\Sigma)^1_1$ it holds that

$$\begin{aligned} \mathscr{L}(\mathscr{B})(\xi) &= \sum_{\rho \in R_{\mathscr{B}}(I_{\mathscr{B}}, \xi, q_f)} c_{\mathscr{B}}(\rho, \xi) = \sum_{\substack{\rho \in R_{\mathscr{B}}(I_{\mathscr{B}}, \xi, q_f), \\ \rho \text{ initial}}} c_{\mathscr{B}}(\rho, \xi) \\ &\overset{\divideontimes}{=} \sum_{\substack{\rho \in R_{\mathscr{B}}(I_{\mathscr{B}}, \xi, q_f), \\ \rho \text{ initial}}} \sum_{\rho' \in R^\rho_{\mathscr{A}}(\xi, I, q_f)} c_{\mathscr{A}}(\rho', \xi) \\ &= \sum_{\rho' \in R_{\mathscr{A}}(I, \xi, q_f)} c_{\mathscr{A}}(\rho', \xi) = \mathscr{L}(\mathscr{A})(\xi). \end{aligned}$$

Note that, if $\mathrm{size}(\xi) = 0$, $c_{\mathscr{B}}(\rho, \xi) = \mathbb{1}_{q^I}(q)$ (which equals $\mathscr{L}(\mathscr{A})(\xi)$, as $\varphi$ is proper) and hence in this case we do not need to prove equation $\divideontimes$.

Therefore, we (only) prove that for every $\xi \in T(\Sigma)_1^1$ with size$(\xi) \geq 1$ we have

$$\begin{aligned} &\forall q \in Q \forall \rho \in R_{\mathscr{B}}(I_{\mathscr{B}}, \xi, q) \text{ initial:} \\ &\qquad c_{\mathscr{B}}(\rho, \xi) = \sum_{\rho' \in R^{\rho}_{\mathscr{A}}(\xi, I, q)} c_{\mathscr{A}}(\rho', \xi) \end{aligned} \tag{5}$$

by induction on $\xi$.

*Case 1:* Assume that $\xi = \langle n, \alpha \rangle$ for some $\alpha \in \Sigma^{(0)}$. Let $q \in Q$ and $\rho \in R_{\mathscr{B}}(I_{\mathscr{B}}, \xi, q)$ initial. We have

$$c_{\mathscr{B}}(\rho, \xi) = E_0^{\mathscr{B}}(\alpha, q) = E_0^{\mathscr{A}}(\alpha, q) = c_{\mathscr{A}}(\rho, \xi).$$

*Case 2:* Assume that $\xi = \langle n, \sigma(t_1, \ldots, t_s) \rangle$ for some $\sigma \in \Sigma^{(s)}$, $s \geq 1$ such that equation (5) holds for $\xi_i \coloneqq \langle n, t_i \rangle \in T(\Sigma)_n^1$ for every $i \in [s]$. Let $q \in Q$ and $\rho \in R_{\mathscr{B}}(I_{\mathscr{B}}, \xi, q)$ initial. Defining $\rho_i$ as in number 2) of Remark 3.2.7 and abbreviating $q_i \coloneqq \rho((1, i))$ (for $i \in [s]$), it holds that

$$c_{\mathscr{B}}(\rho, \xi) = c_{\mathscr{B}}(\rho, \xi, (1, \varepsilon)) \prod_{i=1}^{s} c_{\mathscr{B}}(\rho_i, \xi_i) = E_s^{\mathscr{B}}(q_1, \ldots, q_s, \sigma, q) \prod_{i=1}^{s} c_{\mathscr{B}}(\rho_i, \xi_i). \tag{6}$$

Denote $K \coloneqq \{i \in [s] \mid \exists j \in [n] \colon q_i = q^I\}$. If $K = \emptyset$, we obtain by induction hypothesis

$$\begin{aligned} c_{\mathscr{B}}(\rho, \xi) &= E_s^{\mathscr{A}}(q_1, \ldots, q_s, \sigma, q) \prod_{i=1}^{s} \Big( \sum_{\rho'_i \in R^{\rho_i}_{\mathscr{A}}(\xi_i, I, q_i)} c_{\mathscr{A}}(\rho'_i, \xi_i) \Big) \\ &= \sum_{\substack{\rho'_i \in R^{\rho_i}_{\mathscr{A}}(\xi_i, I, q_i), \\ \text{for every } i \in [s]}} E_s^{\mathscr{A}}(q_1, \ldots, q_s, \sigma, q) \prod_{i=1}^{s} c_{\mathscr{A}}(\rho'_i, \xi_i) \\ &= \sum_{\rho' \in R^{\rho}_{\mathscr{A}}(\xi, I, q)} c_{\mathscr{A}}(\rho', \xi), \end{aligned}$$

where we have used the bijective correspondence

$$R^{\rho}_{\mathscr{A}}(\xi, I, q) \cong R^{\rho_1}_{\mathscr{A}}(\xi_1, I, q_1) \times \cdots \times R^{\rho_s}_{\mathscr{A}}(\xi_s, I, q_s).$$

Now let $K \neq \emptyset$ and note that $c_{\mathscr{B}}(\rho_i, \xi_i) = 1$ for every $i \in K$, as $\rho_i$ is initial. We use the same notation as in the case $K = \emptyset$ to obtain

$$
\begin{aligned}
c_{\mathscr{B}}(\rho,\xi) &\overset{\star_1}{=} E_s^{\mathscr{B}}(q_1,\dots,q_s,\sigma,q) \prod_{\substack{i\in[s]\\ i\notin K}} \Big( \sum_{\rho'_i \in R_{\mathscr{A}}^{\rho_i}(\xi_i,I,q_i)} c_{\mathscr{A}}(\rho'_i,\xi_i)\Big)\\
&\overset{\star_2}{=} \Big( \sum_{\substack{q'_i\in Q,\\ \text{for every } i\in K}} \big(\prod_{i\in K} I_1(q'_i)\big) E_s(q'_1,\dots,q'_s,\sigma,q)\Big) \prod_{\substack{i\in[s]\\ i\notin K}} \Big( \sum_{\rho'_i \in R_{\mathscr{A}}^{\rho_i}(\xi_i,I,q_i)} c_{\mathscr{A}}(\rho'_i,\xi_i)\Big)\\
&\overset{\star_3}{=} \Big( \sum_{\substack{q'_i\in Q,\\ \text{for every } i\in K}} \big(\prod_{i\in K} I_1(q'_i)\big) E_s(q'_1,\dots,q'_s,\sigma,q)\Big) \sum_{\substack{\rho'_i \in R_{\mathscr{A}}^{\rho_i}(\xi_i,I,q_i),\\ \text{for every } i\in[s]\setminus K}} \Big( \prod_{\substack{i\in[s]\\ i\notin K}} c_{\mathscr{A}}(\rho'_i,\xi_i)\Big)\\
&\overset{\star_4}{=} \sum_{\substack{q'_i\in Q,\\ \text{for every } i\in K}} \sum_{\substack{\rho'_i \in R_{\mathscr{A}}^{\rho_i}(\xi_i,I,q_i),\\ \text{for every } i\in[s]\setminus K}} \Big( \prod_{\substack{i\in[s]\\ i\notin K}} c_{\mathscr{A}}(\rho'_i,\xi_i)\Big)\Big(\prod_{i\in K} I_1(q'_i)\Big) E_s(q'_1,\dots,q'_s,\sigma,q)\\
&\overset{\star_5}{=} \sum_{\substack{q'_i\in Q\\ \text{for every } i\in K}} \sum_{\substack{\rho'_i \in R_{\mathscr{A}}^{\rho_i}(\xi_i,I,q_i)\\ \text{for every } i\in[s]\setminus K}} c_{\mathscr{A}}(\rho',\xi)\\
&\overset{\star_6}{=} \sum_{\rho' \in R_{\mathscr{A}}^{\rho}(\xi,I,q)} c_{\mathscr{A}}(\rho',\xi).
\end{aligned}
$$

Equation $\star_1$ combines equation (6) and the induction hypothesis (5). Equation $\star_2$ is the definition of $E_s^{\mathscr{B}}$. In equation $\star_3$ we used the generalized distributivity law. In equation $\star_4$ we have simply rearranged the terms. The remaining equations $\star_5$ and $\star_6$ use the following correspondence. Any family $(q'_i \in Q \mid i \in K)$ together with a family $(\rho'_i \in R_{\mathscr{A}}^{\rho_i}(\xi_i, I, q_i) \mid i \in [s] \setminus K)$ corresponds uniquely to a run $\rho'$ of $\mathscr{A}$ on $\xi$ ending in $q$ such that $\rho'((1,i)) = q'_i$ for every $i \in K$ and whose restriction to $\xi_i$ is $\rho'_i$ for every $i \in [s] \setminus K$. Equation $\star_5$ additionally assesses the cost of $\rho'$ via Remark 3.2.7. □

**Example 3.4.7.** Consider $\Sigma$ from Example 3.2.4 and the $(1,1)$-WFA $\mathscr{A} = (Q, \Sigma, S, I, F, E)$, where

$$Q = \{q, f\}, \qquad I = ((2 \cdot 1) \cdot \mathbb{1}_q), \qquad F = \{f\},$$

and $E$ is 0 except in the cases

$$E_0(\alpha, q) = E_0(\alpha, f) = E_0(\beta, q) = E_0(\beta, f) = 1 \tag{7}$$
$$E_1(q, \gamma, q) = E_1(q, \gamma, f) = 1 \tag{8}$$
$$E_2(q, q, \sigma, q) = E_2(q, q, \sigma, f) = 1. \tag{9}$$

It surely holds that for every $\xi \in T(\Sigma)_1^1$ it holds that

$$\mathscr{L}(\mathscr{A})(\xi) = \begin{cases} (2\cdot 1)^{\#pos_{x_1}(\xi)} & \text{, if size}(\xi) \geq 1 \\ 0 & \text{, otherwise.} \end{cases}$$

Hence, $\mathscr{L}(\mathscr{A})$ is proper.

Next we construct a normalized $(1,1)$-WFA $\mathscr{B}$ such that $\mathscr{L}(\mathscr{B}) = \mathscr{L}(\mathscr{A})$. Note that $\mathscr{A}$ is already in root state normal form. We apply the construction given in the proof of Proposition 3.4.6 to find that

$$\mathscr{B} = (Q \cup \{q^I\}, \Sigma, S, (\mathbb{1}_{q^I}), F, E_{\mathscr{B}}).$$

$E_{\mathscr{B}}$ is 0 except in the cases (7), (8), (9), and

$$E_{\mathscr{B},1}(q^I, \gamma, q) = E_{\mathscr{B},1}(q^I, \gamma, f) = 2 \cdot 1$$

$$E_{\mathscr{B},2}(q_1, q_2, \sigma, q) = E_{\mathscr{B},2}(q_1, q_2, \sigma, f) = \begin{cases} (2 \cdot 1) & , \text{ if } q_1 = q^I \wedge q_2 = q \\ (2 \cdot 1) & , \text{ if } q_1 = q \wedge q_2 = q^I \\ (4 \cdot 1) & , \text{ if } q_1 = q^I \wedge q_2 = q^I \end{cases}$$

■

# Chapter 4: Closure Properties of Recognizable Weighted Forest Languages

In essence, we prove that $\mathrm{REC}(T(\Sigma)^1_n, S)$ is closed under the operations that will later be called “rational operations”. These include addition, multiplication with a scalar, horizontal and vertical concatenation, and Kleene star.

In the light of Proposition 3.3.2, we might apply the respective results from [6] to obtain most of the upcoming results. The proof of closure under vertical concatenation, however, needed a new proof because of the different notion of concatenation in [6]. We tried to do the same generalization for the Kleene star, yet in this case we were not able to complete the proof without falling back to only allowing the Kleene star of recognizable weighted $(1,1)$-forest languages. In any case, we present full and detailed proofs.

In this chapter, let $\Sigma$ be a ranked alphabet and $S$ a commutative semiring.

## 4.1 The Constant Language $0$ is Recognizable

**Proposition 4.1.1.** Let $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$. It holds that $0 \in \mathrm{REC}(T(\Sigma)^m_n, S)$.

*Proof.* Let $\mathscr{A} := (Q, \Sigma, S, I, F, E)$ be an $(m,n)$-WFA, such that $F = \emptyset^m = \emptyset$. By convention, a sum over an empty index set is 0, hence

$$\mathscr{L}(\mathscr{A}) = 0,$$

which concludes the proof. □

## 4.2 Characteristic Functions of Forests are Recognizable

**Proposition 4.2.1.** Let $n \in \mathbb{N}_0$ and $\xi \in T(\Sigma)^1_n$. It holds that $\mathbb{1}_\xi \in \mathrm{REC}(T(\Sigma)^1_n, S)$.

*Proof.* Let $\xi = \langle n, t_1 \rangle$. We can apply Proposition 3.3.1 to see that the claim is proven by showing that $\mathbb{1}_{t_1} \in \mathrm{REC}(\Sigma, X_n)$, which is shown in [6, Lemmas 6.1 and 6.2].

For the sake of completeness, we provide a direct construction. Define $Q := \mathrm{pos}(t_1)$, $F := \{\varepsilon\}$, and $I := (I_1, \ldots, I_n)$, where for $i \in [n]$ the leaf weight $I_i \colon Q \longrightarrow S$ is given for any $q \in Q$ by

$$I_i(q) := \begin{cases} 1 & \text{, if } \mathrm{lab}(t, q) = x_i \\ 0 & \text{, otherwise.} \end{cases}$$

Moreover we define the family $E = (E_k \mid k \geq 0)$ where for every $k \geq 0$ the map $E_k \colon Q^k \times \Sigma^{(k)} \times Q \longrightarrow S$ is given for every $q_1, \ldots, q_k, q \in Q$ and $\sigma \in \Sigma^{(k)}$ by

$$E_k(q_1, \ldots, q_k, \sigma, q) := \begin{cases} 1 & \text{, if } \mathrm{lab}(t, q) = \sigma \wedge \forall i \in [k] : q_i = qi \\ 0 & \text{, otherwise.} \end{cases}$$

It can be proven straightforward that the $(1,n)$-WFA

$$\mathscr{A} := (Q, \Sigma, S, I, F, E)$$

satisifes $\mathscr{L}(\mathscr{A}) = \mathbb{1}_\xi$. □

## 4.3 Closure under Scalar Multiplication

**Proposition 4.3.1.** Let $n \in \mathbb{N}_0$, $a \in S$, and $\varphi \in \mathrm{REC}(T(\Sigma)_n^1, S)$. It holds that

$$a \cdot \varphi \in \mathrm{REC}(T(\Sigma)_n^1, S).$$

*Proof.* We can apply Proposition 3.3.1 to see that this claim is proven in [6, Lemma 6.3].

We additionally provide a construction. Let $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be an (1,$n$)-WFA in root state normal form such that $\mathscr{L}(\mathscr{A}) = \varphi$. Let $f \in Q$ be the unique root state. We construct the $(1, n)$-WFA

$$\mathscr{A}' := (Q, \Sigma, S, I, F, E'),$$

where for any $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q_1, \ldots, q_k, q \in Q$ we define

$$E'_k(q_1, \ldots, q_k, \sigma, q) := \begin{cases} E_k(q_1, \ldots, q_k, \sigma, q) & \text{, if } q \neq f \\ a \cdot E_k(q_1, \ldots, q_k, \sigma, q) & \text{, if } q = f. \end{cases}$$

Let $\xi \in T(\Sigma)_n^1$ and $f \in F$. It surely holds that

$$R_{\mathscr{A}}(I, \xi, f) = R_{\mathscr{A}'}(I, \xi, f).$$

Thus we only have to show $c_{\mathscr{A}'}(\rho, \xi) = a \cdot c_{\mathscr{A}}(\rho, \xi)$ for every $\rho \in R_{\mathscr{A}}(I, \xi, f)$ in order to prove $\mathscr{L}(\mathscr{A}') = a \cdot \mathscr{L}(\mathscr{A})$.

Let $\rho \in R_{\mathscr{A}}(I, \xi, f)$ and $w \in \mathrm{pos}(\xi)$. If $f$ occurs in $\rho$ at any position other than $(1, \varepsilon)$, we have

$$c_{\mathscr{A}}(\rho, \xi) = c_{\mathscr{A}'}(\rho, \xi) = 0.$$

Therefore we can without loss of generality assume that $f$ does only occur at the root of $\xi$.

If $w \neq (1, \varepsilon)$, we have that $c_{\mathscr{A}'}(\rho, \xi, w) = c_{\mathscr{A}}(\rho, \xi, w)$ by definition of $E'$. If $w = (1, \varepsilon)$, we have that $c_{\mathscr{A}'}(\rho, \xi, w) = a \cdot c_{\mathscr{A}}(\rho, \xi, w)$. This proves $c_{\mathscr{A}'}(\rho, \xi) = a \cdot c_{\mathscr{A}}(\rho, \xi)$. □

## 4.4 Closure under Sum

**Proposition 4.4.1.** Let $n \in \mathbb{N}_0$ and $\varphi, \psi \in \mathrm{REC}(T(\Sigma)_n^1, S)$ be two weighted (1,$n$)-forest languages. It holds that

$$\varphi + \psi \in \mathrm{REC}(T(\Sigma)_n^1, S).$$

*Proof.* We can apply Proposition 3.3.1 to see that this claim is proven in [6, Lemma 6.4].

We additionally provide a construction. Let

$$\mathscr{A} = (Q, \Sigma, S, I, F, E) \qquad \text{and} \qquad \mathscr{A}' = (Q', \Sigma, S, I', F', E')$$

be two (1,$n$)-WFA such that $\mathscr{L}(\mathscr{A}) = \varphi$ and $\mathscr{L}(\mathscr{A}') = \psi$. We moreover assume that $Q \cap Q' = \emptyset$.

Define $\hat{Q} \coloneqq Q \cup Q'$, $\hat{F} \coloneqq F \cup F'$, and $\hat{I} \coloneqq (\hat{I}_1, \dots, \hat{I}_n)$, where for $i \in [n]$, the leaf weight $\hat{I}_i \colon \hat{Q} \longrightarrow S$ is given for any $q \in \hat{Q}$ by

$$\hat{I}_i(q) \coloneqq \begin{cases} I_i(q) & , \text{ if } q \in Q \\ I'_i(q) & , \text{ if } q \in Q'. \end{cases}$$

Moreover we define the family $\hat{E} = (\hat{E}_k \mid k \geq 0)$ where for every $k \geq 0$ the map $\hat{E}_k \colon \hat{Q}^k \times \Sigma^{(k)} \times \hat{Q} \longrightarrow S$ is given for every $q_1, \dots, q_k, q \in \hat{Q}$ and $\sigma \in \Sigma^{(k)}$ by

$$\hat{E}_k(q_1, \dots, q_k, \sigma, q) \coloneqq \begin{cases} E_k(q_1, \dots, q_k, \sigma, q) & , \text{ if } q_1, \dots, q_k, q \in Q \\ E'_k(q_1, \dots, q_k, \sigma, q) & , \text{ if } q_1, \dots, q_k, q \in Q' \\ 0 & , \text{ otherwise.} \end{cases}$$

We claim that the (1,$n$)-WFA $\hat{\mathscr{A}} = (\hat{Q}, \Sigma, S, \hat{I}, \hat{F}, \hat{E})$ satisfies $\mathscr{L}(\hat{\mathscr{A}}) = \varphi + \psi$.

Let $\xi \in T(\Sigma)^1_n$ and $q \in F$ and $\rho \in R_{\hat{\mathscr{A}}}(\hat{I}, \xi, q)$. It holds that

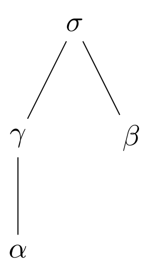
$$c_{\hat{\mathscr{A}}}(\rho, \xi) = \begin{cases} c_{\mathscr{A}}(\rho, \xi) & , \text{ if } \operatorname{im}(\rho) \subseteq Q \\ c_{\mathscr{A}'}(\rho, \xi) & , \text{ if } \operatorname{im}(\rho) \subseteq Q' \\ 0 & , \text{ otherwise,} \end{cases}$$

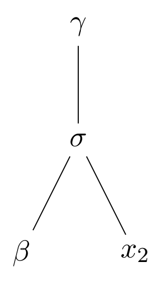
as follows directly from the definition. But this yields

$$\begin{aligned} \mathscr{L}(\hat{\mathscr{A}})(\xi) &= \sum_{q \in \hat{F}} \sum_{\rho \in R_{\hat{\mathscr{A}}}(\hat{I}, \xi, q)} c_{\hat{\mathscr{A}}}(\rho, \xi) \\ &= \sum_{q \in F} \sum_{\rho \in R_{\mathscr{A}}(I, \xi, q)} c_{\mathscr{A}}(\rho, \xi) + \sum_{q \in F'} \sum_{\rho \in R_{\mathscr{A}'}(I', \xi, q)} c_{\mathscr{A}'}(\rho, \xi) \\ &= \varphi(\xi) + \psi(\xi) \end{aligned}$$

□

**Example 4.4.2.** Consider $\Sigma$ from Example 3.2.4 and let $\xi_1$ and $\xi_2$ be the following trees.

$\xi_1 \coloneqq$ 


$\xi_2 \coloneqq$ 


Let $a_1 \coloneqq (2 \cdot 1) \in S$ and $a_2 \coloneqq (3 \cdot 1) \in S$. We construct a $(1,2)$-WFA $\mathscr{A}$ such that

$$\mathscr{L}(\mathscr{A}) = \sum_{i=1}^{2} a_i \mathbb{1}_{\xi_i}.$$

Using Propositions 4.2.1 and 4.4.1, we define

$$Q \coloneqq \mathrm{pos}(\xi_1) \dot{\cup} \mathrm{pos}(\xi_2) = \{\varepsilon, 1, 2, 11, \hat{\varepsilon}, \hat{1}, \hat{1}\hat{1}, \hat{1}\hat{2}\}.$$

Here we implemented the disjoint union via the hat symbol and projected all forest positions to tree positions. We moreover define

$$I = (0, \mathbb{1}_{\hat{1}\hat{2}}) \qquad \text{and} \qquad F = \{\varepsilon, \hat{\varepsilon}\}.$$

Note that we are constructing a $(1,2)$-WFA and hence $I$ consists of two leaf weight maps.

The weight transition map given by Propositions 4.2.1 and 4.4.1 is 0 except in the following cases. We arrange the weights in the tree structures of $\xi_1$ and $\xi_2$ to visualize the idea.

- $E_2(1, 2, \sigma, \varepsilon) = 1$
  - $E_1(11, \gamma, 1) = 1$
    - $E_0(\alpha, 11) = 1$
  - $E_0(\beta, 2) = 1$

- $E_1(\hat{1}, \gamma, \hat{\varepsilon}) = 1$
  - $E_2(\hat{1}\hat{1}, \hat{1}\hat{2}, \sigma, \hat{1}) = 1$
    - $E_0(\beta, \hat{1}\hat{1}) = 1$
    - $x_2$

We call this weight transition map $E$. The construction from Proposition 4.3.1 now forces the root values to be $a_1$ and $a_2$ respectively. Therefore we update $E$ to be

- $E_2(1, 2, \sigma, \varepsilon) = a_1$
  - $E_1(11, \gamma, 1) = 1$
    - $E_0(\alpha, 11) = 1$
  - $E_0(\beta, 2) = 1$

- $E_1(\hat{1}, \gamma, \hat{\varepsilon}) = a_2$
  - $E_2(\hat{1}\hat{1}, \hat{1}\hat{2}, \sigma, \hat{1}) = 1$
    - $E_0(\beta, \hat{1}\hat{1}) = 1$
    - $x_2$

and combine all the given parts to

$$\mathscr{A} = (Q, \Sigma, S, I, F, E).$$

This $(1,2)$-WFA now accepts the desired weighted $(1,2)$-forest language. ■

## 4.5 Closure under Horizontal Concatenation

**Definition 4.5.1.** Let $m, m', n \in \mathbb{N}_0$, $\varphi\colon T(\Sigma)_n^m \longrightarrow S$, and $\psi\colon T(\Sigma)_n^{m'} \longrightarrow S$. We define the **horizontal concatenation of $\varphi$ and $\psi$** as the map

$$\varphi \times \psi\colon T(\Sigma)_n^{m+m'} \longrightarrow S,$$

given for every $\langle n, t_1, \ldots, t_m, s_1, \ldots, s_{m'} \rangle \in T(\Sigma)_n^{m+m'}$ by

$$(\varphi \times \psi)(\langle n, t_1, \ldots, t_m, s_1, \ldots, s_{m'} \rangle) \coloneqq \varphi(\langle n, t_1, \ldots, t_m \rangle)\psi(\langle n, s_1, \ldots, s_{m'} \rangle).$$

■

**Proposition 4.5.2.** Let $m, m', n \in \mathbb{N}_0$, $\varphi \in \mathrm{REC}(T(\Sigma)^m_n, S)$, and $\psi \in \mathrm{REC}(T(\Sigma)^{m'}_n, S)$. It holds that

$$\varphi \times \psi \in \mathrm{REC}(T(\Sigma)^{m+m'}_n, S).$$

*Proof.* Let $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ be a $(m,n)$-WFA such that $\mathscr{L}(\mathscr{A}) = \varphi$ and let $\mathscr{A}' = (Q', \Sigma, S, I', F', E')$ be a $(m',n)$-WFA such that $\mathscr{L}(\mathscr{A}') = \psi$ and $Q \cap Q' = \emptyset$.

Define the $(m+m',n)$-WFA

$$\hat{\mathscr{A}} = (\hat{Q}, \Sigma, S, \hat{I}, \hat{F}, \hat{E}),$$

where the components are given as follows. $\hat{Q} \coloneqq Q \cup Q'$, $\hat{F} \coloneqq F \times F'$, and $\hat{I} = (\hat{I}_1, \ldots, \hat{I}_n)$, where for any $i \in [n]$, the leaf weight $\hat{I}_i \colon \hat{Q} \longrightarrow S$ is given for every $q \in Q$ by

$$\hat{I}_i(q) \coloneqq \begin{cases} I_i(q) & , \text{ if } q \in Q \\ I'_i(q) & , \text{ if } q \in Q'. \end{cases}$$

Moreover the family $\hat{E} = (\hat{E}_k \mid k \geq 0)$ is given for every $k \geq 0$ by the map $\hat{E}_k \colon \hat{Q}^k \times \Sigma^{(k)} \times \hat{Q} \longrightarrow S$, where for every $q_1, \ldots, q_k, q \in \hat{Q}$ and $\sigma \in \Sigma^{(k)}$ we define

$$\hat{E}_k(q_1, \ldots, q_k, \sigma, q) \coloneqq \begin{cases} E_k(q_1, \ldots, q_k, \sigma, q) & , \text{ if } q_1, \ldots, q_k, q \in Q \\ E'_k(q_1, \ldots, q_k, \sigma, q) & , \text{ if } q_1, \ldots, q_k, q \in Q' \\ 0 & , \text{ otherwise.} \end{cases}$$

Let $\xi \in T(\Sigma)^{m+m'}_n$ such that $\xi = \langle n, t_1, \ldots, t_m, s_1, \ldots, s_{m'} \rangle$. Moreover denote $\xi_1 \coloneqq \langle n, t_1, \ldots, t_m \rangle \in T(\Sigma)^m_n$ and $\xi_2 \coloneqq \langle n, s_1, \ldots, s_{m'} \rangle T(\Sigma)^{m'}_n$. Let $\hat{q} \in \hat{F}$ and let $q \in F$ and $q' \in F'$ such that $\hat{q} = (q, q')$.

We say that a run $\rho \in R_{\hat{\mathscr{A}}}(\hat{I}, \xi, \hat{q})$ is **$Q$-$Q'$-restricted**, if

$$\mathrm{im}(\rho_1) \subseteq Q \text{ and } \mathrm{im}(\rho_2) \subseteq Q',$$

where $\rho_1$ and $\rho_2$ are the restrictions of $\rho$ to $\xi_1$ and $\xi_2$, respectively. We denote the set of $Q$-$Q'$-restricted runs by $R^{Q;Q'}_{\hat{\mathscr{A}}}(\hat{I}, \xi, \hat{q})$.

Let $\rho \in R_{\hat{\mathscr{A}}}(\hat{I}, \xi, \hat{q})$ be a run that is not $Q$-$Q'$-restricted. Without loss of generality, there exists a $w = (i, u) \in \mathrm{pos}(\xi_1)$ such that $\rho_1(w) \notin Q$. As $\rho((i, \varepsilon)) \in Q$, we can in particular choose $w$ such that there exists $w' = (i, u') \in \mathrm{pos}(\xi_1)$ satisfying $\rho(w') \in Q$ and $u'l = u$ for some $l \in \mathbb{N}$. This yields

$$c_{\hat{\mathscr{A}}}(\rho, \xi, w') = 0$$

by definition of $\hat{E}$ and hence implies $c_{\hat{\mathscr{A}}}(\rho, \xi) = 0$.

We claim that for every run $\rho \in R^{Q;Q'}_{\hat{\mathscr{A}}}(\xi, \hat{I}, \hat{q})$ it holds that

$$c_{\hat{\mathscr{A}}}(\rho, \xi) = c_{\mathscr{A}}(\rho_1, \xi_1) c_{\mathscr{A}'}(\rho_2, \xi_2). \tag{10}$$

Note that this is well-defined, as $\rho_1$ is a run of $\mathscr{A}$ on $\xi_1$ (and $\rho_2$ a run of $\mathscr{A}'$ on $\xi_2$, respectively).

By Remark 3.2.7 it surely holds that

$$c_{\hat{\mathscr{A}}}(\rho,\xi) = c_{\hat{\mathscr{A}}}(\rho_1,\xi_1)c_{\hat{\mathscr{A}}}(\rho_2,\xi_2),$$

whence it suffices to show that

$$c_{\hat{\mathscr{A}}}(\rho_1,\xi_1) = c_{\mathscr{A}}(\rho_1,\xi_1) \qquad \text{and} \qquad c_{\hat{\mathscr{A}}}(\rho_2,\xi_2) = c_{\mathscr{A}'}(\rho_2,\xi_2).$$

However, this follows directly from the definition of $\hat{E}$.

We conclude that

$$\begin{aligned}
\mathscr{L}(\hat{\mathscr{A}})(\xi) &= \sum_{\hat{q}\in\hat{F}} \sum_{\rho\in R_{\hat{\mathscr{A}}}(\hat{I},\xi,\hat{q})} c_{\hat{\mathscr{A}}}(\rho,\xi) \\
&= \sum_{\hat{q}\in\hat{F}} \sum_{\rho\in R^{Q;Q'}_{\hat{\mathscr{A}}}(\xi,\hat{I},\hat{q})} c_{\hat{\mathscr{A}}}(\rho,\xi) \\
&= \sum_{\hat{q}\in\hat{F}} \sum_{\rho\in R^{Q;Q'}_{\hat{\mathscr{A}}}(\xi,\hat{I},\hat{q})} c_{\mathscr{A}}(\rho_1,\xi_1)c_{\mathscr{A}'}(\rho_2,\xi_2) \\
&\overset{\star}{=} \left(\sum_{q\in F} \sum_{\rho_1\in R_{\mathscr{A}}(I,\xi_1,q)} c_{\mathscr{A}}(\rho_1,\xi_1)\right)\left(\sum_{q'\in F'} \sum_{\rho\in R_{\mathscr{A}'}(I',\xi_2,q')} c_{\mathscr{A}'}(\rho_2,\xi_2)\right) \\
&= \mathscr{L}(\mathscr{A})(\xi_1)\mathscr{L}(\mathscr{A}')(\xi_2).
\end{aligned}$$

In equality $\star$ we have used that every $Q$-$Q'$-restricted run of $\hat{\mathscr{A}}$ on $\xi$ is uniquely defined by a run of $\mathscr{A}$ on $\xi_1$ and a run of $\mathscr{A}'$ on $\xi_2$ to apply the generalized distributivity law in $S$. This proves the claim. □

**Remark 4.5.3.** Note that the construction of $\hat{\mathscr{A}}$ is equal in Propositions 4.4.1 and 4.5.2. The reason why we cannot lift Proposition 4.4.1 to forests of arbitrary upper ranks lies in Proposition 3.3.2: recognizable weighted $(m,n)$-forest languages are rectangular, but rectangular weight maps are not closed under sum. This also gives rise to the restrictions we will have to make in order to define vertical concatenation of weighted forest languages. ■

**Example 4.5.4.** We continue Example 4.4.2 by constructing a $(2,2)$-WFA $\mathscr{A}'$ such that

$$\mathscr{L}(\mathscr{A}') = \bigtimes_{i=1}^{2} a_i\mathbb{1}_{\xi_i} = a_1\mathbb{1}_{\xi_1} \times a_2\mathbb{1}_{\xi_2}.$$

The construction for $\mathscr{A}'$ is similar to $\mathscr{A}$. In fact, the only difference is the set of root states. We obtain

$$\mathscr{A}' = (Q, \Sigma, S, I, \{(\varepsilon,\hat{\varepsilon})\}, E).$$

■

## 4.6 Closure under Vertical Concatenation

**Definition 4.6.1.** Let $m, n, p \in \mathbb{N}_0$, $\xi \in T(\Sigma)_p^m$, and denote for every $i \in [n]$ the number of occurences of $x_i$ in $\xi$ by $l_i$. Let moreover $\zeta \in T(\Sigma)_n^m$ and $\eta_j^i \in T(\Sigma)_p^1$ for every $i \in [n]$ and $j \in [l_i]$. We call the tuple

$$(\zeta, (\eta_j^i \mid i \in [n], j \in [l_i]))$$

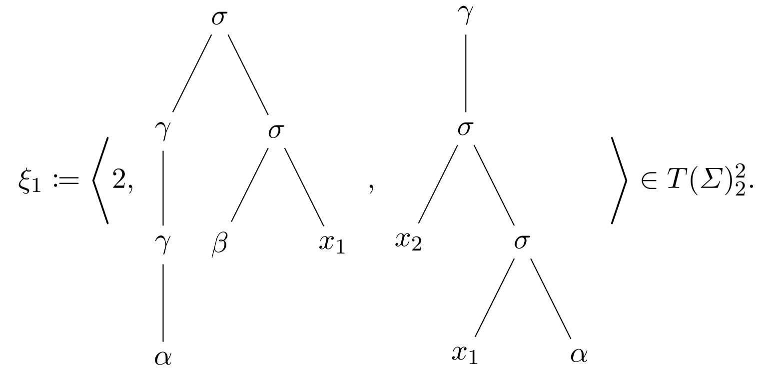
an $n$**-decomposition of** $\xi$ if

$$\xi = \zeta[\forall i \in [n] \colon x_i \leftarrow \eta_1^i, \dots, \eta_{l_i}^i].$$

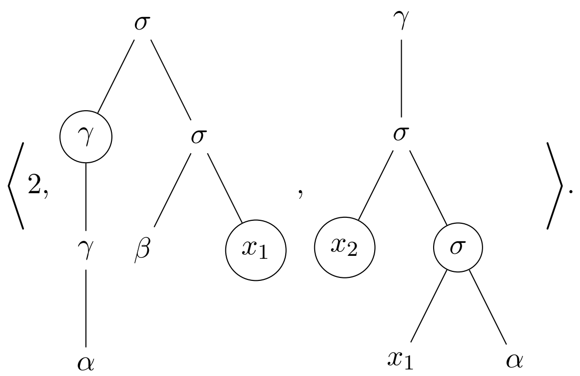
Whenever we say that $(\zeta, \eta)$ is an $n$-decomposition of $\xi$, we assume that $\eta$ is given as $(\eta_j^i \mid i \in [n], j \in [l_i])$. We denote the **set of** $n$**-decompositions of** $\xi$ by $\mathrm{Dec}_n(\xi)$. ■

**Remark 4.6.2.** Throughout this thesis, we use the following notational convention for $n$-decompositions. If we quantify $(\zeta, \eta) \in \mathrm{Dec}_n(\xi)$, then $l_i$ denotes the number of occurences of $x_i$ in $\xi$ for every $i \in [n]$. If $\eta$ is indexed by a subscript, we add this index to $l_i$. For example if we have an $I$-indexed family of $n$-decompositions $((\zeta_k, \eta_k) \mid k \in I)$, then we write $l_{k,i}$ instead of $l_i$ for every $k \in I$. ■

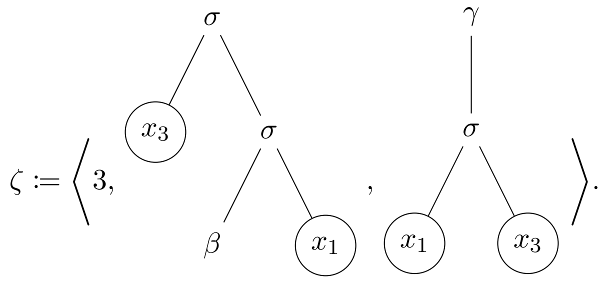
**Example 4.6.3.** Consider $\Sigma$ from Example 3.2.4. Let




A 3-decomposition of $\xi_1$ can be constructed as follows. First we fix a set of positions $U$ in $\xi_1$, given by the circles in



Next we fix a map $\vartheta\colon U \longrightarrow [3]$. By substituting $x_{\vartheta(p)}$ into $\xi_1$ at position $p$ for every $p \in U$, we obtain a new forest. In our case, we map $\varphi((1,1)) = 3$, $\varphi((1,22)) = 1$, $\varphi((2,11)) = 1$, and $\varphi((2,12)) = 3$.

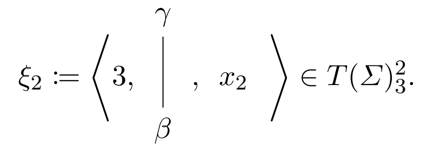



For every $p \in U$, we simply denote the subtree of $\xi_1$ at $p$ by $\eta_j^{(\vartheta(p))}$, if the $j$-th occurrence of $x_{\varphi(p)}$ in $\zeta$ is at position $p$.

$$\eta_1^1 = \langle 2, x_1\rangle, \qquad \eta_2^1 = \langle 2, x_2\rangle$$
$$\eta_1^3 = \langle 2, \gamma(\gamma(\alpha))\rangle, \qquad \eta_2^3 = \langle 2, \sigma(x_1, \alpha)\rangle.$$

In total we obtain that $(\zeta, (\eta_j^i \mid i \in \{1,3\}, j \in [2])) \in \mathrm{Dec}_3(\xi_1)$.

In preparation of upcoming examples, let




The set of 1-decompositions of $(\xi_2)$ consists of the following tuples. Note that 1-decompositions (of an arbitrary forest) are of the form $(\zeta, (\eta_1^1, \ldots, \eta_{\ell_1}^1))$. For better readability, we replace any $\eta_j^i$-forest by its single tree component.

$$\Big(\big\langle 2, \gamma(\beta),\ x_1\ \big\rangle, (\ x_2\ )\Big), \quad \Big(\big\langle 2, \gamma(\ x_1\ ),\ x_1\ \big\rangle, (\ \beta\ ,\ x_2\ )\Big), \quad \Big(\big\langle 2,\ x_1\ ,\ x_1\ \big\rangle, (\ \gamma(\beta)\ ,\ x_2\ )\Big)$$

The above arrows indicate the associations between the $\eta$-trees and the respective positions they were "cut" from. ■

**Remark 4.6.4.** Let $m, n, p \in \mathbb{N}_0$ and $\xi \in T(\Sigma)_p^m$.

1) Let $m = 1$ and $\xi = \langle p, \sigma(t_1, \ldots, t_k)\rangle$ for some $\sigma \in \Sigma^{(k)}$, $k \geq 1$. Denote $\xi' := \langle p, t_1, \ldots, t_k\rangle \in T(\Sigma)_p^k$. The following facts are derived from the definition of $n$-decompositions.

For every $(\zeta, \eta) \in \mathrm{Dec}_n(\xi)$ such that $\mathrm{size}(\zeta) \geq 1$, we know that

$$\zeta = \langle n, \sigma(\zeta_1, \ldots, \zeta_k)\rangle$$

for some $\zeta_1, \ldots, \zeta_k \in T_\Sigma(X_n)$.Denoting $\zeta' := \langle n, \zeta_1, \ldots, \zeta_s\rangle \in T(\Sigma)_n^s$, it holds that $(\zeta', \eta) \in \mathrm{Dec}_n(\xi')$. We obtain a map

$$\iota\colon \{(\zeta, \eta) \in \mathrm{Dec}_n(\xi) \mid \mathrm{size}(\zeta) \geq 1\} \longrightarrow \mathrm{Dec}_n(\xi'), \qquad \iota(\zeta, \eta) = (\zeta', \eta).$$

Vice versa, for every $(\zeta', \eta) \in \mathrm{Dec}_n(\xi')$ it holds that $(\sigma(\zeta'), \eta) \in \mathrm{Dec}_n(\xi)$. We obtain a map

$$\kappa\colon \mathrm{Dec}_n(\xi') \longrightarrow \{(\zeta, \eta) \in \mathrm{Dec}_n(\xi) \mid \mathrm{size}(\zeta) \geq 1\}, \qquad \kappa(\zeta', \eta) = (\sigma(\zeta'), \eta).$$

As $\sigma(\langle n, \zeta_1, \ldots, \zeta_s\rangle) = \zeta$ and $\langle n, \sigma(\zeta')_1, \ldots, \sigma(\zeta')_s\rangle = \zeta'$), we find that $\kappa^{-1} = \iota$ and hence

$$\{(\zeta, \eta) \in \mathrm{Dec}_n(\xi) \mid \mathrm{size}(\zeta) \geq 1\} \cong \mathrm{Dec}_n(\xi').$$

2) Let $m > 1$ and $\xi = \langle n, t_1, \ldots, t_m\rangle \in T(\Sigma)^m_n$. Denote $\xi_i \coloneqq \langle n, t_i\rangle$ for any $i \in [m]$.

Let $(\zeta, \eta) \in \mathrm{Dec}_n(\xi)$ such that $\zeta = \langle n, \zeta_1, \ldots, \zeta_m\rangle$. For every $k \in [m]$ denote by $\eta_k$ the subfamily of $\eta$ consisting of the $\eta^i_j$ that correspond to variables in $\zeta_k$. Up to a shift in the indices in $\eta_k$, we obtain $(\zeta_k, \eta_k) \in \mathrm{Dec}_n(\xi_k)$.

Vice versa, let $(\zeta_k, \eta_k) \in \mathrm{Dec}_n(\xi_k)$such that $\zeta_k = \langle n, t_k\rangle \in T(\Sigma)^1_n$ for some $t_k \in T_\Sigma(X_n)$ for any $k \in [m]$. Up to a shift in the indices of the $\eta_k$, we obtain

$$(\langle n, t_1, \ldots, t_m\rangle, \bigcup_{k=1}^{m} \eta_k) \in \mathrm{Dec}_n(\xi)$$

and ultimately the bijection

$$\mathrm{Dec}_n(\xi) \cong \mathrm{Dec}_n(\xi_1) \times \cdots \times \mathrm{Dec}_n(\xi_m).$$

■

**Definition 4.6.5.** Let $m, n, p \in \mathbb{N}_0$. For every weighted $(m,n)$-forest language $\varphi$ and every rectangular weighted $(n,p)$-forest language $\psi$, we define the **vertical concatenation of $\varphi$ and $\psi$** as the weighted $(m,p)$-forest language $\varphi \cdot \psi$ as follows. Let $\xi \in T(\Sigma)^m_p$ and denote for every $i \in [n]$ the number of occurences of $x_i$ in $\xi$ by $l_i$. Then we define

$$(\varphi \cdot \psi)(\xi) \coloneqq \sum_{(\zeta,\eta)\in \mathrm{Dec}_n(\xi)} \varphi(\zeta) \cdot \prod_{i=1}^{n} \prod_{j=1}^{l_i} \psi_i(\eta^i_j),$$

where the $\psi_1, \ldots, \psi_n$ are the rectangular components of $\psi$. ■

**Example 4.6.6.** Consider $\Sigma$ from Example 3.2.4 and recall that $S$ is an arbitrary commutative semiring. Let $\varphi\colon T(\Sigma)^2_1 \longrightarrow S$ and $\psi\colon T(\Sigma)^1_3 \longrightarrow S$ be defined as follows. For every $\xi \in T(\Sigma)^2_1$ and $\xi' \in T(\Sigma)^1_3$ let

$$\begin{aligned} \varphi(\xi) &\coloneqq (2 \cdot 1)^{\#\,\mathrm{pos}_\alpha(\xi)} \cdot (3 \cdot 1)^{\#\,\mathrm{pos}_{x_1}(\xi)}, \\ \psi(\xi') &\coloneqq (5 \cdot 1)^{\#\,\mathrm{pos}_\beta(\xi')} \qquad . \end{aligned}$$

Consider the forest $\xi_2 \in T(\Sigma)^2_3$ from Example 4.6.3 and recall the already determined set of 1-decompositions of $\xi_2$. We calculate

$$\begin{aligned} (\varphi \cdot \psi)(\xi_2) = \; & \varphi\big(\langle 2, \gamma(\beta), x_1\rangle\big) \cdot \psi\big(\langle 3, x_2\rangle\big) + \\ & \varphi\big(\langle 2, \gamma(x_1), x_1\rangle\big) \cdot \psi\big(\langle 3, \beta\rangle\big) \cdot \psi\big(\langle 3, x_2\rangle\big) + \\ & \varphi\big(\langle 2, x_1, x_1\rangle\big) \cdot \psi\big(\langle 3, \gamma(\beta)\rangle\big) \cdot \psi\big(\langle 3, x_2\rangle\big) \\ = \; & (3 \cdot 1)^1 \cdot 1 + (3 \cdot 1)^2 \cdot (5 \cdot 1)^1 + (3 \cdot 1)^2 \cdot (5 \cdot 1)^1 \\ = \; & 93 \cdot 1 \end{aligned}$$

■

Next we show that $\mathrm{REC}(T(\Sigma), S)$ is closed under concatenation. The intuition for the construction is as follows. Given an $(m,n)$-WFA $\mathscr{A}$ and an $(n,p)$-WFA $\mathscr{B}$, we construct an $(m,p)$-WFA that nondeterministically chooses a decomposition of the input forest and then processes the upper half like $\mathscr{A}$ and the lower half like $\mathscr{B}$. At the switching positions, the automaton takes all possible root weights of $\mathscr{B}$ into account.

**Proposition 4.6.7.** Let $m, n, p \in \mathbb{N}_0$, $\varphi \in \mathrm{REC}(T(\Sigma)^m_n, S)$, and $\psi \in \mathrm{REC}(T(\Sigma)^n_p, S)$. It holds that

$$\varphi \cdot \psi \in \mathrm{REC}(T(\Sigma)^m_p, S).$$

*Proof.* First note that the claim is well-defined, as by Proposition 3.3.2, recognizable weighted forest languages are rectangular.

There exists an $(m,n)$-WFA $\mathscr{A} = (Q_\varphi, \Sigma, S, I_\varphi, F_\varphi, E_\varphi)$ such that $\mathscr{L}(\mathscr{A}) = \varphi$ and an $(n,p)$-WFA $\mathscr{B} = (Q_\psi, \Sigma, S, I_\psi, F_\psi, E_\psi)$ in root state normal form such that $\mathscr{L}(\mathscr{B}) = \psi$. Moreover, let $Q_\varphi \cap Q_\psi = \emptyset$ and $F_\psi = \{f_\psi\}$ for some $f_\psi = (f_{\psi,1}, \ldots, f_{\psi,n}) \in Q_\psi^n$.

We construct the $(m,p)$-WFA

$$\mathscr{C} \coloneqq (Q, \Sigma, S, I, F, E),$$

where $Q \coloneqq Q_\varphi \cup Q_\psi$, $F \coloneqq F_\varphi$, and $I = (I_1, \ldots, I_p)$ such that for every $i \in [p]$ and $q \in Q$,

$$I_i(q) \coloneqq \begin{cases} I_{\psi,i}(q) & \text{, if } q \in Q_\psi \\ \sum_{j=1}^{n} I_{\psi,i}(f_{\psi,j}) I_{\varphi,j}(q) & \text{, if } q \in Q_\varphi. \end{cases}$$

Moreover $E = (E_k \mid k \geq 0)$, where for every $k \geq 0$, $q_1, \ldots, q_k, q \in Q$, and $\sigma \in \Sigma^{(k)}$,

$$E_k(q_1, \ldots, q_k, \sigma, q) \coloneqq$$
$$\begin{cases} E_{\psi,k}(q_1, \ldots, q_k, \sigma, q) & \text{, if } q_1, \ldots, q_k, q \in Q_\psi \\ E_{\varphi,k}(q_1, \ldots, q_k, \sigma, q) & \text{, if } q_1, \ldots, q_k, q \in Q_\varphi \wedge k \geq 1 \\ \sum_{j=1}^{n} \big(E_{\psi,k}(q_1, \ldots, q_k, \sigma, f_{\psi,j}) \cdot I_{\varphi,j}(q)\big) & \text{, if } q_1, \ldots, q_k \in Q_\psi, q \in Q_\varphi \wedge k \geq 1 \\ E_{\varphi,0}(\sigma, q) + \sum_{j=1}^{n} \big(E_{\psi,0}(\sigma, f_{\psi,j}) \cdot I_{\varphi,j}(q)\big) & \text{, if } q \in Q_\varphi \wedge k = 0 \\ 0 & \text{, otherwise.} \end{cases}$$

In order to show that $\mathscr{L}(\mathscr{C}) = \mathscr{L}(\mathscr{A}) \cdot \mathscr{L}(\mathscr{B})$, we have to prove equation $\diamond$ in the following chain of equations. Let $\xi \in T(\Sigma)^m_p$. It holds that

$$
\begin{aligned}
\mathscr{L}(\mathscr{C})(\xi) &= \sum_{f \in F} E^{\mathscr{C}}_{p,m}(I,\xi,f) \\
&\overset{\diamond}{=} \sum_{f \in F} \sum_{(\zeta,\eta) \in \mathrm{Dec}_n(\xi)} \Big( E^{\mathscr{A}}_{n,m}(I_\varphi,\zeta,f) \prod_{i=1}^{n} \prod_{j=1}^{l_i} E^{\mathscr{B}}_{p,1}(I_\psi,\eta^i_j,f_{\psi,i}) \Big) \\
&= \sum_{(\zeta,\eta) \in \mathrm{Dec}_n(\xi)} \Big( \big( \sum_{f \in F} E^{\mathscr{A}}_{n,m}(I_\varphi,\zeta,f) \big) \prod_{i=1}^{n} \prod_{j=1}^{l_i} E^{\mathscr{B}}_{p,1}(I_\psi,\eta^i_j,f_{\psi,i}) \Big) \\
&= \sum_{(\zeta,\eta) \in \mathrm{Dec}_n(\xi)} \Big( \mathscr{L}(\mathscr{A})(\zeta) \prod_{i=1}^{n} \prod_{j=1}^{l_i} \mathscr{L}(\mathscr{B})_i(\eta^i_j) \Big) \\
&= (\mathscr{L}(\mathscr{A}) \cdot \mathscr{L}(\mathscr{B}))(\xi).
\end{aligned}
$$

Note that equation $\diamond$ formalizes the aforementioned motivation of the construction of $\mathscr{C}$.

To prove equation $\diamond$, we show the following more general statement. For every $\xi \in T(\Sigma)_p$ (in the following, $l$ denotes the upper rank of $\xi$), we have

$$
\forall q \in Q^l_\varphi \colon E_{p,l}(I,\xi,q) = \sum_{(\zeta,\eta) \in \mathrm{Dec}_n(\xi)} E^{\mathscr{A}}_{n,l}(I_\varphi,\zeta,q) \prod_{i=1}^{n} \prod_{j=1}^{l_i} E^{\mathscr{B}}_{p,1}(I_\psi,\eta^i_j,f_{\psi,i}), \tag{11}
$$

Note that this is only shown for states of $\mathscr{A}$, as the term $E^{\mathscr{A}}_{n,l}(I_\varphi,\zeta,q)$ would not be well-defined in any other case. This goes hand in hand with the idea of this proof. Decomposing $\xi$ means allowing the state behaviour of $\mathscr{B}$ on a bottom part of $\xi$ but forcing transitions to $\mathscr{A}$ eventually, namely at the root positions of $\eta$. This happens, even if $\zeta$ does not contain symbols from $\Sigma$.

We prove equation (11) by induction on $\xi$. Note that the important bit of this proof is to correctly determine $\mathrm{Dec}_n(\xi)$ for the different cases.

*Case 1:* Assume that $l = 1$ and $\xi = \langle p, x_i \rangle$ for some $i \in [p]$.

Let $(\zeta,\eta) \in \mathrm{Dec}_n(\xi)$. As $\mathrm{size}(\xi) \geq \mathrm{size}(\zeta)$, we find that $\mathrm{size}(\zeta) = 0$ (analogously, $\mathrm{size}(\eta^k_j) = 0$ for every $k \in [n]$ and $j \in [l_k]$). Therefore, there exists $k \in [n]$ such that $\zeta = \langle n, x_k \rangle$. Thus, the familiy $\eta$ contains but a single tree, $\eta^k_1$, which in turn satisfies $\eta^k_1 = \langle p, x_i \rangle$. We have thus determined

$$
\mathrm{Dec}_n(\xi) = \{ (\pi^n_k, \pi^k_i) \mid k \in [n] \}.
$$

Therefore we have for any $q \in Q^l$

$$
E_{p,l}(I,\xi,q) = I_i(q) = \sum_{k=1}^{n} I_{\psi,i}(f_{\psi,k}) I_{\varphi,k}(q) = \sum_{k=1}^{n} E^{\mathscr{B}}_{p,1}(I_\psi,\pi^k_i,f_{\psi,k}) E^{\mathscr{A}}_{n,l}(I_\varphi,\pi^n_k,q).
$$

It is clear that the sum on the right hand side of this equation runs over all $n$-decompositions of $\xi$ and therefore we obtain equation (11).

*Case 2:* Assume that $l = 1$ and $\xi = \langle p, \alpha \rangle$ for some $\alpha \in \Sigma^{(0)}$.

Let $(\zeta, \eta) \in \mathrm{Dec}_n(\xi)$. As $\mathrm{size}_\Sigma(\xi) \geq \mathrm{size}_\Sigma(\zeta)$, we find that $\mathrm{size}(\zeta) \leq 1$ (analogously, $\mathrm{size}(\eta_j^k) \leq 1$ for every $k \in [n]$ and $j \in [l_k]$). If $\mathrm{size}(\zeta) = 0$, then there exists a $k \in [n]$ such that $\zeta = \langle n, x_k \rangle$ and $\eta$ contains but the single tree $\eta_1^k = \langle p, \alpha \rangle = \xi$. If $\mathrm{size}(\zeta) = 1$, then $\zeta = \langle n, \alpha \rangle$ and $\eta$ is an empty family. We have thus determined

$$\mathrm{Dec}_n(\xi) = \{(\pi_k^n, \langle p, \alpha \rangle) \mid k \in [n]\} \cup \{(\langle n, \alpha \rangle, ())\}.$$

Therefore we have for any $q \in Q^l$

$$\begin{aligned} E_{p,l}(I, \xi, q) = E_0(\alpha, q) &= E_{\varphi,0}(\alpha, q) + \sum_{k=1}^{n} \big(I_{\varphi,k}(q) \cdot E_{\psi,0}(\alpha, f_{\psi,k})\big) \\ &= E_{n,l}^{\mathscr{A}}(I_\varphi, \langle n, \alpha \rangle, q) + \sum_{k=1}^{n} E_{n,l}^{\mathscr{A}}(I_\varphi, \pi_k^n, q) \cdot E_{p,1}^{\mathscr{B}}(I_\psi, \langle p, \alpha \rangle, f_{\psi,k}) \end{aligned}$$

Using the above description of $\mathrm{Dec}_n(\xi)$ for this case, we obtain the desired equation (11).

*Case 3:* Assume that $l = 1$ and $\xi = \langle p, \sigma(t_1, \ldots, t_s) \rangle$ for some $\sigma \in \Sigma^{(s)}$, $s \geq 1$ such that equation (11) holds for $\xi' := \langle p, t_1, \ldots, t_s \rangle \in T(\Sigma)_p^s$.

Let $(\zeta, \eta) \in \mathrm{Dec}_n(\xi)$. If $\mathrm{size}(\zeta) = 0$, it holds that $\zeta = \langle n, x_k \rangle$ for some $k \in [n]$ and $\eta$ contains but the single tree $\eta_1^k = \xi$. By Remark 4.6.4, the case $\mathrm{size}(\zeta) \geq 1$ covers exactly the $n$-decompositions $(\zeta', \eta)$ of $\xi'$. Therefore we have determined

$$Dec_n(\xi) = \{(\sigma(\zeta'), \eta) \mid (\zeta', \eta) \in \mathrm{Dec}_n(\xi')\} \cup \{(\pi_k^n, \xi) \mid k \in [n]\}.$$

Now, let $q \in Q^l$. It holds that

$$\begin{aligned} E_{p,l}(I, \xi, q) &= \sum_{q' \in Q^s} E_{p,s}(I, \xi', q') E_s(q', \sigma, q) \\ &= \underbrace{\sum_{q' \in Q_\varphi^s} E_{p,s}(I, \xi', q') E_s(q', \sigma, q)}_{=:\Sigma_1} + \underbrace{\sum_{q' \in Q_\psi^s} E_{p,s}(I, \xi', q') E_s(q', \sigma, q)}_{=:\Sigma_2}. \end{aligned}$$

We moreover know that

$$\begin{aligned} \Sigma_2 &= \sum_{q' \in Q_\psi^s} E_{p,s}(I, \xi', q') \sum_{k=1}^{n} \big(E_{\psi,s}(q', \sigma, f_{\psi,k}) \cdot I_{\varphi,k}(q)\big) \\ &\doteq \sum_{k=1}^{n} I_{\varphi,k}(q) \sum_{q' \in Q_\psi^s} E_{p,s}^{\mathscr{B}}(I, \xi', q') E_{\psi,s}(q', \sigma, f_{\psi,k}) \\ &= \sum_{k=1}^{n} I_{\varphi,k}(q) E_{p,1}^{\mathscr{B}}(I, \xi, f_{\psi,k}) \\ &= \sum_{k=1}^{n} E_{n,l}^{\mathscr{A}}(I_\varphi, \pi_k^n, q) E_{p,1}^{\mathscr{B}}(I, \xi, f_{\psi,k}). \end{aligned}$$

In equation $\bullet$, we first rearranged the terms and then applied the easy fact that $E_{p,s}(I,\xi',q') = E^{\mathscr{B}}_{p,s}(I,\xi',q')$, as $q'$ consists of states in $Q_\psi$ and state transition weights can only transition from $\mathscr{B}$ to $\mathscr{A}$ but never back. We have thus seen that $\Sigma_2$ amounts to the summands in the right hand side of equation (11) where the decomposition of $\xi$ is of the form $(\pi^n_k, \xi)$.

Furthermore, using the induction hypothesis we know that

$$\begin{aligned}
\Sigma_1 &= \sum_{q'\in Q^s_\varphi} \Big( \sum_{(\zeta',\eta)\in \mathrm{Dec}_n(\xi')} E^{\mathscr{A}}_{n,s}(I_\varphi,\zeta',q') \prod_{i=1}^{n}\prod_{j=1}^{l_i} E^{\mathscr{B}}_{p,1}(I_\psi,\eta^i_j,f_{\psi,i})\Big) E_{\varphi,s}(q',\sigma,q) \\
&= \sum_{(\zeta',\eta)\in \mathrm{Dec}_n(\xi')} \Big( \sum_{q'\in Q^s_\varphi} E^{\mathscr{A}}_{n,s}(I_\varphi,\zeta',q') E_{\varphi,s}(q',\sigma,q)\Big) \prod_{i=1}^{n}\prod_{j=1}^{l_i} E^{\mathscr{B}}_{p,1}(I_\psi,\eta^i_j,f_{\psi,i}) \\
&= \sum_{(\zeta',\eta)\in \mathrm{Dec}_n(\xi')} E^{\mathscr{A}}_{n,1}(I_\varphi,\sigma(\zeta'),q) \prod_{i=1}^{n}\prod_{j=1}^{l_i} E^{\mathscr{B}}_{p,1}(I_\psi,\eta^i_j,f_{\psi,i}),
\end{aligned}$$

which amounts to the summands in the right hand side of equation (11) where the decomposition of $\xi$ is of the form $(\sigma(\zeta'),\eta)$.

*Case 4:* Assume that $l > 1$, $\xi = \langle n, t_1, \dots, t_l\rangle$, and $q = (q_1,\dots,q_l) \in Q^l_\varphi$ such that equation (11) holds for $\xi_i \coloneqq \langle n, t_i\rangle$ for any $i \in [l]$.

In Remark 4.6.4 we have seen that

$$\mathrm{Dec}_n(\xi) \cong \mathrm{Dec}_n(\xi_1) \times \cdots \times \mathrm{Dec}_n(\xi_l).$$

Therefore we obtain that

$$\begin{aligned}
E_{p,l}(I,\xi,q) &= \prod_{k=1}^{l} E_{p,1}(I,\xi_k,q_k) \\
&= \prod_{k=1}^{l} \Big( \sum_{(\zeta_k,\eta_k)\in \mathrm{Dec}_n(\xi_k)} E^{\mathscr{A}}_{n,1}(I_\varphi,\zeta_k,q_k) \prod_{i=1}^{n}\prod_{j=1}^{l_{k,i}} E^{\mathscr{B}}_{p,1}(I_\psi,(\eta_k)^i_j,f_{\psi,i})\Big) \\
&\overset{\dagger}{=} \sum_{(\zeta,\eta)\in \mathrm{Dec}_n(\xi)} \prod_{k=1}^{l} \Big( E^{\mathscr{A}}_{n,1}(I_\varphi,\zeta_k,q_k) \prod_{i=1}^{n}\prod_{j=1}^{l_{k,i}} E^{\mathscr{B}}_{p,1}(I_\psi,(\eta_k)^i_j,f_{\psi,i})\Big) \\
&\overset{\ddagger}{=} \sum_{(\zeta,\eta)\in \mathrm{Dec}_n(\xi)} E^{\mathscr{A}}_{n,l}(I_\varphi,\zeta,q) \prod_{i=1}^{n}\prod_{j=1}^{l_i} E^{\mathscr{B}}_{p,1}(I_\psi,\eta^i_j,f_{\psi,i}),
\end{aligned}$$

where equation $\dagger$ uses the aforementioned bijection that maps $(\zeta,\eta) \mapsto ((\zeta_1,\eta_1),\dots,(\zeta_l,\eta_l))$. Equation $\ddagger$ uses the fact that

$$\prod_{k=1}^{l} E^{\mathscr{A}}_{n,1}(I_\varphi,\zeta_k,q_k) = E^{\mathscr{A}}_{n,l}(I_\varphi,\zeta,q)$$

and moreover that for any $i \in [n]$,

$$\prod_{k=1}^{l} \prod_{j=1}^{l_{k,i}} E^{\mathscr{B}}_{p,1}(I_\psi, (\eta_k)^i_j, f_{\psi,i}) = \prod_{j=1}^{l_i} E^{\mathscr{B}}_{p,1}(I_\psi, \eta^i_j, f_{\psi,i}).$$

This concludes the final case and hence proves the proposition. □

**Example 4.6.8.** Consider $\Sigma$, $\varphi$, and $\psi$ from Example 4.6.6. We first show that $\varphi$ and $\psi$ are recognizable and then apply the construction from Proposition 4.6.7 to see that $\varphi \cdot \psi$ is recognizable as well.

We define the automata

$$\mathscr{A} \coloneqq (Q_\varphi, \Sigma, S, I_\varphi, F_\varphi, E_\varphi) \qquad \text{and} \qquad \mathscr{B} \coloneqq (Q_\psi, \Sigma, S, I_\psi, F_\psi, E_\psi),$$

where

$$\begin{aligned} Q_\varphi &\coloneqq \{q_1, f_1, f_2\} & Q_\psi &\coloneqq \{q_1, f_1\} \\ F_\varphi &\coloneqq \{(f_1, f_2)\} & F_\psi &\coloneqq \{f_1\} \\ I_\varphi &\coloneqq ((3 \cdot 1) \cdot \mathbb{1}_{Q_\varphi}) & I_\psi &\coloneqq (\mathbb{1}_{Q_\psi}, \mathbb{1}_{Q_\psi}, \mathbb{1}_{Q_\psi}), \end{aligned}$$

and $E_\varphi$ and $E_\psi$ are defined as 0 except in the following cases.

$$\begin{aligned} E_{\varphi,0}(\alpha, q_\varphi) &\coloneqq (2 \cdot 1) & E_{\psi,0}(\alpha, q_\psi) &\coloneqq 1 \\ E_{\varphi,0}(\beta, q_\varphi) &\coloneqq 1 & E_{\psi,0}(\beta, q_\psi) &\coloneqq (5 \cdot 1) \\ E_{\varphi,1}(q_1, \gamma, q_\varphi) &\coloneqq 1 & E_{\psi,1}(q_1, \gamma, q_\psi) &\coloneqq 1 \\ E_{\varphi,2}(q_1, q_1, \sigma, q_\varphi) &\coloneqq 1 & E_{\psi,2}(q_1, q_1, \sigma, q_\psi) &\coloneqq 1, \end{aligned}$$

where $q_\varphi \in Q_\varphi$ and $q_\psi \in Q_\psi$.

It follows immediately, that a run $\rho$ of $\mathscr{A}$ (or $\mathscr{B}$) on a tree $\xi \in T(\Sigma)^2_1$ (or in $T(\Sigma)^1_3$, respectively) can only have non-vanishing cost if it labels root positions in $\xi$ with the unique root states and every other position with $q_1$. Therefore, we obtain

$$\mathscr{L}(\mathscr{A}) = \varphi \qquad \text{and} \qquad \mathscr{L}(\mathscr{B}) = \psi.$$

The $(2,3)$-WFA $\mathscr{C} = (Q', \Sigma, S, I', F', E')$ that satisfies

$$\mathscr{L}(\mathscr{C}) = \mathscr{L}(\mathscr{A}) \cdot \mathscr{L}(\mathscr{B})$$

is given by the proof of Proposition 4.6.7. We introduce a copy $\tilde{Q_\psi}$ of the set $Q_\psi$ and determine the components of $\mathscr{C}$ as

$$Q' = Q_\varphi \cup \tilde{Q_\psi} = \{q_1, f_1, f_2, \tilde{q_1}, \tilde{f_1}\},$$

$$I' = (I'_1, I'_2, I'_3), \quad I'_1 = I'_2 = I'_3 = (3 \cdot 1) \cdot \mathbb{1}_{Q_\varphi} + \mathbb{1}_{\tilde{Q_\psi}},$$

$$F' = F_\varphi = \{(f_1, f_2)\},$$

and the transition weights are defined as 0 except in the following cases.

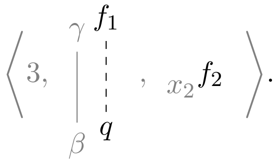
$$E'_0(\alpha, q_\varphi) = (2 \cdot 1) + 1 \cdot (3 \cdot 1) = 5 \cdot 1 \qquad E'_0(\alpha, q_\psi) = 1$$
$$E'_0(\beta, q_\varphi) = 1 + (5 \cdot 1) \cdot (3 \cdot 1) = 16 \cdot 1 \qquad E'_0(\beta, q_\psi) = 5 \cdot 1$$
$$E'_1(q_1, \gamma, q_\varphi) = E'_1(\tilde{q}_1, \gamma, q_\psi) = 1 \qquad E'_1(\tilde{q}_1, \gamma, q_\varphi) = 1 \cdot (3 \cdot 1) = 3 \cdot 1$$
$$E'_2(q_1, q_1, \sigma, q_\varphi) = E'_2(\tilde{q}_1, \tilde{q}_1, \sigma, q_\psi) = 1 \qquad E'_2(\tilde{q}_1, \tilde{q}_1, \sigma, q_\varphi) = 1 \cdot (3 \cdot 1) = (3 \cdot 1),$$

where $q_\varphi \in Q_\varphi$ and $q_\psi \in Q_\psi$.

Consider the forest $\xi_2 \in T(\Sigma)^2_3$ from Example 4.6.3 and recall that

$$(\mathscr{L}(\mathscr{A}) \cdot \mathscr{L}(\mathscr{B}))(\xi_2) = 93 \cdot 1.$$

We calculate $\mathscr{L}(\mathscr{C})(\xi_2)$. Let $\rho \in R_\mathscr{C}(I', \xi_2, (f_1, f_2))$. Therefore, $\rho$ is already uniquely determined by its value $q \in Q'$ at position $(1, 1) \in \mathrm{pos}(\xi_2)$.




If $q$ is either $f_1$, $f_2$, or $\tilde{f}_1$, the cost of $\rho$ is 0. Hence, we only have to consider the cases $q = q_1$ and $q = \tilde{q}_1$. Denote the respective runs by $\rho_{q_1}$ and $\rho_{\tilde{q}_1}$. We obtain

$$\mathscr{L}(\mathscr{C})(\xi_2) = c_\mathscr{C}(\rho_{q_1}, \xi_2) + c_\mathscr{C}(\rho_{\tilde{q}_1}, \xi_2) = (5 \cdot 1) \cdot (3 \cdot 1) \cdot (3 \cdot 1) + (16 \cdot 1) \cdot 1 \cdot (3 \cdot 1) = 93 \cdot 1,$$

as expected. ■

## 4.7 Closure under Kleene Star

**Definition 4.7.1.** Let $n \in \mathbb{N}_0$ and $\varphi \colon T(\Sigma)^1_1 \longrightarrow S$. We define inductively for $k \in \mathbb{N}_0$

$$\varphi^0 = \mathbb{1}_{\pi^1_1},$$
$$\varphi^{k+1} = \varphi \cdot \varphi^k + \mathbb{1}_{\pi^1_1},$$

and call $\varphi^k$ the **$k$-th vertical power of** $\varphi$. ■

**Example 4.7.2.** We continue Example 3.4.7 and denote $\varphi \coloneqq \mathscr{L}(\mathscr{A})$. The tree

$$\xi \coloneqq \gamma(\gamma(\alpha)) \in T(\Sigma)^1_1$$

has exactly the following 1-decompositions.

$$\mathrm{Dec}_1(\xi) = \Big\{ \big(\xi, ()\big), \big(\gamma(\gamma(x_1)), (\alpha)\big), \big(\gamma(x_1), (\gamma(\alpha))\big), \big(x_1, (\xi)\big) \Big\}.$$

It moreover holds that

$$\begin{aligned}\mathrm{Dec}_1(\alpha) &= \{(\alpha, ()), (x_1, (\alpha))\}, \text{ and}\\ \mathrm{Dec}_1(\gamma(\alpha)) &= \{(\gamma(\alpha), ()), (\gamma(x_1), (\alpha)), (x_1, (\gamma(\alpha)))\}.\end{aligned}$$

We calculate

$$\begin{aligned}\varphi^0(\xi) &= 0,\\ \varphi^1(\xi) &= 1,\\ \varphi^2(\xi) &= \sum_{(\zeta,\eta)\in \mathrm{Dec}_1(\xi)} \varphi(\zeta)\cdot \prod_{j=1}^{l_1} \varphi(\eta^1_j)\\ &= 1 + (2\cdot 1)\cdot 1 + (2\cdot 1)\cdot 1 + 0\cdot 1 = 5\cdot 1,\\ \varphi^3(\xi) &= \sum_{(\zeta,\eta)\in \mathrm{Dec}_1(\xi)} \varphi(\zeta)\cdot \prod_{j=1}^{l_1} \varphi^2(\eta^1_j)\\ &= 1 + (2\cdot 1)\cdot \varphi^2(\alpha) + (2\cdot 1)\cdot \varphi^2(\gamma(\alpha))\\ &= 1 + (2\cdot 1)\cdot (\varphi(\alpha) + \varphi(x_1)\cdot\varphi(\alpha))\\ &\quad + (2\cdot 1)\cdot(\varphi(\gamma(\alpha)) + \varphi(\gamma(x_1))\cdot\varphi(\alpha) + \varphi(x_1)\cdot\varphi(\gamma(\alpha)))\\ &= 1 + (2\cdot 1)\cdot(1+0) + (2\cdot 1)\cdot(1 + (2\cdot 1)\cdot 1 + 0\cdot 1) = 9\cdot 1.\end{aligned}$$

Similarly, one can proceed to calculate $\varphi^n(\xi)$ for $n \geq 4$. Lemma 4.7.3 shows that this is not necessary and $\varphi^n(\xi) = 9\cdot 1$ for every $n \geq 4$. ■

**Lemma 4.7.3.** Let $n \in \mathbb{N}_0$, $\varphi\colon T(\Sigma)^1_1 \longrightarrow S$ proper, and $\xi \in T(\Sigma)^1_1$. It holds that

$$\varphi^{l+1}(\xi) = \varphi^l(\xi)$$

for every $l \geq \mathrm{ht}(\xi) + 1$.

*Proof.* We proceed by induction on the height of $\xi$.

*Case* ht $= 0$*:* Assume that $\xi = \langle 1, x_1\rangle$. It surely holds that $\varphi^0(\xi) = 1$. Assume that $\varphi^l(\xi) = 1$ for some $l \in \mathbb{N}_0$. Since

$$\mathrm{Dec}_1(\xi) = \{(\pi^1_1, \pi^1_1)\},$$

we can apply properness of $\varphi$ to see that

$$\begin{aligned}\varphi^{l+1}(\xi) &= (\varphi\cdot\varphi^l)(\xi) + 1\\ &= \sum_{(\zeta,\eta)\in \mathrm{Dec}_1(\xi)} \underbrace{\varphi(\zeta)}_{=0} \prod_{i=1}^{1}\prod_{j=1}^{l_i} \varphi^l(\eta^i_j) + 1\\ &= 1.\end{aligned}$$

In particular, $\varphi^{l+1}(\xi) = \varphi^l(\xi)$ for every $l \geq 0$.

*Case* ht $= 1$*:* Assume that $\xi = \langle 1, \alpha \rangle$ for some $\alpha \in \Sigma^{(0)}$. Let $l \geq 1$. It holds that $\mathbb{1}_{\pi}(\xi) = 0$ and

$$\mathrm{Dec}_1(\xi) = \{(\pi_1^1, \langle 1, \alpha \rangle)\} \cup \{(\langle 1, \alpha \rangle, ())\}.$$

Therefore

$$\begin{aligned} \varphi^{l+1}(\xi) &= (\varphi \cdot \varphi^l)(\xi) + 0 \\ &= \sum_{(\zeta,\eta) \in \mathrm{Dec}_1(\xi)} \varphi(\zeta) \prod_{i=1}^{1} \prod_{j=1}^{l_i} \varphi^l(\eta_j^i) \\ &= \varphi(\xi), \end{aligned}$$

where we again apply properness of $\varphi$ in the last equation to remove the summand where $\zeta = \pi_1^1$. This shows in particular $\varphi^{l+1}(\xi) = \varphi^l(\xi)$.

*Case* ht $\geq 1$*:* Assume that $\xi = \langle 1, \sigma(t_1, \ldots, t_s) \rangle$ for some $\sigma \in \Sigma^{(s)}$, $s \geq 1$ such that the claim holds for any tree $\xi'$ with $\mathrm{ht}(\xi') < \mathrm{ht}(\xi)$. Let $l \geq \mathrm{ht}(\xi) + 1$. It holds that $\mathbb{1}_{\pi_1^1}(\xi) = 0$ and

$$\mathrm{Dec}_1(\xi) = \{(\sigma(\zeta'), \eta) \mid (\zeta', \eta) \in \mathrm{Dec}_1(\xi')\} \cup \{(\pi_1^1, \xi)\}.$$

Therefore

$$\begin{aligned} \varphi^{l+1}(\xi) &= (\varphi \cdot \varphi^l)(\xi) + 0 \\ &= \sum_{(\zeta,\eta) \in \mathrm{Dec}_1(\xi)} \varphi(\zeta) \prod_{i=1}^{1} \prod_{j=1}^{l_i} \varphi^l(\eta_j^i) \overset{\star_1}{=} \sum_{\substack{(\zeta,\eta) \in \mathrm{Dec}_1(\xi) \\ \zeta = \sigma(\zeta')}} \varphi(\zeta) \prod_{i=1}^{1} \prod_{j=1}^{l_i} \varphi^l(\eta_j^i) \\ &\overset{\star_2}{=} \sum_{\substack{(\zeta,\eta) \in \mathrm{Dec}_1(\xi) \\ \zeta = \sigma(\zeta')}} \varphi(\zeta) \prod_{i=1}^{1} \prod_{j=1}^{l_i} \varphi^{l-1}(\eta_j^i) \overset{\star_3}{=} \sum_{(\zeta,\eta) \in \mathrm{Dec}_1(\xi)} \varphi(\zeta) \prod_{i=1}^{1} \prod_{j=1}^{l_i} \varphi^{l-1}(\eta_j^i) \\ &= (\varphi \cdot \varphi^{l-1})(\xi) + 0 = \varphi^l(\xi). \end{aligned}$$

In equations $\star_1$ and $\star_3$ we use that the summand where $\zeta = \pi_1^1$ vanishes. In equation $\star_2$ we use that $\mathrm{ht}(\eta_j^i) < \mathrm{ht}(\xi)$, as $\mathrm{ht}(\zeta) > 0$. □

**Definition 4.7.4.** Let $n \in \mathbb{N}_0$, $\varphi\colon T(\Sigma)_1^1 \longrightarrow S$ proper. We define the **Kleene star of** $\varphi$, denoted $\varphi^*\colon T(\Sigma)_1^1 \longrightarrow S$, by

$$\varphi^*(\xi) \coloneqq \varphi^{\mathrm{ht}(\xi)+1}(\xi),$$

for every $\xi \in T(\Sigma)_1^1$. ■

**Lemma 4.7.5.** Let $n \in \mathbb{N}_0$, $\varphi\colon T(\Sigma)_1^1 \longrightarrow S$ proper. It holds that

$$\varphi^* = \varphi \cdot \varphi^* + \mathbb{1}_{\pi_1^1}.$$

*Proof.* Let $\xi \in T(\Sigma)_1^1$ and $h \coloneqq \mathrm{ht}(\xi)$. We immediately obtain

$$\varphi^*(\xi) = \varphi^{h+1}(\xi) = \varphi^{h+2}(\xi) = (\varphi \cdot \varphi^{h+1} + \mathbb{1}_{\pi_1^1})(\xi),$$

as $\varphi$ is proper. We moreover obtain

$$(\varphi \cdot \varphi^{h+1})(\xi) = (\varphi \cdot \varphi^*)(\xi),$$

which proves the claim. □

**Proposition 4.7.6.** Let $\varphi \in \mathrm{REC}(T(\Sigma)_1^1, S)$ proper. It holds that

$$\varphi^* \in \mathrm{REC}(T(\Sigma)_1^1, S).$$

*Proof.* By Proposition 3.4.6, there exists a normalized (1,1)-WFA $\mathscr{A} = (Q, \Sigma, S, I, F, E)$ such that $\mathscr{L}(\mathscr{A}) = \varphi$. Let $q_f$ and $q^I$ be the unique root and leaf state of $\mathscr{A}$, respectively.

We define the (1,1)-WFA

$$\mathscr{A}^* \coloneqq (Q^*, \Sigma, S, I, F^*, E^*),$$

where $Q^* \coloneqq Q \setminus \{q_f\}$, $F^* \coloneqq \{q^I\}$, and for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q, q_1, \ldots, q_k \in Q$

$$E_k^*(q_1, \ldots, q_k, \sigma, q) \coloneqq \begin{cases} E_k(q_1, \ldots, q_k, \sigma, q) & , \text{ if } q \neq q^I \\ E_k(q_1, \ldots, q_k, \sigma, q_f) & , \text{ if } q = q^I. \end{cases}$$

We first prove a technical tool. For any $\xi \in T(\Sigma)_1^1$ it holds that

$$\forall q \in Q^* \setminus \{q^I\} \colon E_{1,1}^{\mathscr{A}^*}(I, \xi, q) = \sum_{(\zeta,\eta) \in \mathrm{Dec}_1(\xi)} E_{1,1}^{\mathscr{A}}(I, \zeta, q) \prod_{i=1}^{1} \prod_{j=1}^{l_i} \mathscr{L}(\mathscr{A}^*)(\eta_j^i), \tag{12}$$

by induction on $\xi$.

*Case 1:* Assume that $\xi = \langle 1, x_1 \rangle$ and let $q \in Q^* \setminus \{q^I\}$. It holds that

$$E_{1,1}^{\mathscr{A}^*}(I, \xi, q) = I_1(q) = 0 = E_{1,1}^{\mathscr{A}}(I, \pi_1^1, q) \mathscr{L}(\mathscr{A}^*)(\pi_1^1),$$

which concludes this case.

*Case 2:* Assume that $\xi = \langle 1, \alpha \rangle$ for some $\alpha \in \Sigma^{(0)}$ and let $q \in Q^* \setminus \{q^I\}$. It holds that

$$E_{1,1}^{\mathscr{A}^*}(I, \xi, q) = E_0^*(\alpha, q) = E_0(\alpha, q) + I_1(q) \mathscr{L}(\mathscr{A}^*)(\alpha),$$

which concludes this case.

*Case 3:* Assume that $\xi = \langle 1, \sigma(t_1, \ldots, t_s) \rangle$ for some $\sigma \in \Sigma^{(s)}$, $s \geq 1$ such that the claim holds for $\xi_i \coloneqq \langle 1, t_i \rangle \in T(\Sigma)_1^1$ for every $i \in [s]$. Let $q \in Q^* \setminus \{q^I\}$. It holds that

$$E_{1,1}^{\mathscr{A}^*}(I, \xi, q) = \sum_{q_1, \ldots, q_s \in Q^*} E_s^*(q_1, \ldots, q_s, \sigma, q) \prod_{k=1}^{s} E_{1,1}^{\mathscr{A}^*}(I, \xi_k, q_k), \tag{13}$$

and by induction hypothesis

$$E^{\mathscr{A}^*}_{1,1}(I,\xi_k,q_k) = \sum_{(\zeta_k,\eta_k)\in\mathrm{Dec}_1(\xi_k)} E^{\mathscr{A}^*}_{1,1}(I,\zeta_k,q_k)\prod_{i=1}^{1}\prod_{j=1}^{l_{k,i}}\mathscr{L}(\mathscr{A}^*)((\eta_k)^i_j), \tag{14}$$

for any $k\in[s]$. Combining equations (13) and (14) and using the fact that

$$\{(\zeta,\eta)\in\mathrm{Dec}_1(\xi)\mid \mathrm{size}(\zeta)\geq 1\}\cong\mathrm{Dec}_1(\xi_1)\times\cdots\times\mathrm{Dec}_1(\xi_s),$$

we arrive at

$$\begin{aligned}
E^{\mathscr{A}^*}_{1,1}(I,\xi,q) &= \sum_{q_1,\dots,q_s\in Q^*} E^*_s(q_1,\dots,q_s,\sigma,q)\cdot \\
&\qquad\cdot\prod_{k=1}^{s}\Bigg(\sum_{(\zeta_k,\eta_k)\in\mathrm{Dec}_1(\xi_k)} E^{\mathscr{A}^*}_{1,1}(I,\zeta_k,q_k)\prod_{i=1}^{1}\prod_{j=1}^{l_{k,i}}\mathscr{L}(\mathscr{A}^*)((\eta_k)^i_j)\Bigg) \\
&= \sum_{\substack{(\zeta,\eta)\in\mathrm{Dec}_1(\xi)\\ \mathrm{size}(\zeta)\geq 1}}\ \sum_{q_1,\dots,q_s\in Q^*} E^*_s(q_1,\dots,q_s,\sigma,q)\cdot \\
&\qquad\cdot\prod_{k=1}^{s}\Bigg(E^{\mathscr{A}^*}_{1,1}(I,\zeta_k,q)\prod_{i=1}^{1}\prod_{j=1}^{l_{k,i}}\mathscr{L}(\mathscr{A}^*)((\eta_k)^i_j)\Bigg) \\
&= \sum_{\substack{(\zeta,\eta)\in\mathrm{Dec}_1(\xi)\\ \mathrm{size}(\zeta)\geq 1}}\Bigg(\sum_{q_1,\dots,q_s\in Q^*} E^*_s(q_1,\dots,q_s,\sigma,q)\prod_{k=1}^{s}E^{\mathscr{A}^*}_{1,1}(I,\zeta_k,q)\Bigg)\cdot \\
&\qquad\cdot\prod_{i=1}^{1}\prod_{j=1}^{l_i}\mathscr{L}(\mathscr{A}^*)(\eta^i_j) \\
&= \sum_{\substack{(\zeta,\eta)\in\mathrm{Dec}_1(\xi)\\ \mathrm{size}(\zeta)\geq 1}} E^{\mathscr{A}^*}_{1,1}(I,\zeta,q)\prod_{i=1}^{1}\prod_{j=1}^{l_i}\mathscr{L}(\mathscr{A}^*)(\eta^i_j).
\end{aligned}$$

This concludes the proof of equation (12), as summands on the right hand side with $\mathrm{size}(\zeta)=0$ vanish.

Now we use equation (12) to prove that

$$\mathscr{L}(\mathscr{A})^*(\xi)=\mathscr{L}(\mathscr{A}^*)(\xi)$$

by induction on the height of $\xi$.

*Case 1:* Assume that $\xi=\langle 1,x_1\rangle$. It holds that

$$\mathscr{L}(\mathscr{A})^*(\xi)=1=I_1(q^I)=\mathscr{L}(\mathscr{A}^*)(\xi),$$

which concludes this case.

*Case 2:* Assume that $\xi = \langle 1, \alpha \rangle$ for some $\alpha \in \Sigma^{(0)}$. By case 2 in the proof of Lemma 4.7.3, we have $\mathscr{L}(\mathscr{A})^*(\xi) = \mathscr{L}(\mathscr{A})(\xi)$, whence

$$\mathscr{L}(\mathscr{A})^*(\xi) = E_0(\alpha, q_f) = E_0^*(\alpha, q^I) = E_{1,1}^{\mathscr{A}^*}(I, \alpha, q^I) = \mathscr{L}(\mathscr{A}^*)(\xi),$$

which concludes this case.

*Case 3:* Assume that $\xi = \langle 1, \sigma(t_1, \ldots, t_s) \rangle$ for some $\sigma \in \Sigma^{(s)}$, $s \geq 1$ such that the claim holds for every $\xi'$ such that $\operatorname{ht} \xi' < \operatorname{ht} \xi$. Using the equality from Lemma 4.7.5, the fact that $\mathscr{L}(\mathscr{A})$ is proper, and the induction hypothesis, we find that

$$\begin{aligned}
\mathscr{L}(\mathscr{A})^*(\xi) &= \sum_{(\zeta,\eta) \in \mathrm{Dec}_1(\xi)} \mathscr{L}(\mathscr{A})(\zeta) \prod_{j=1}^{l_1} \mathscr{L}(\mathscr{A})^*(\eta_j^i) \\
&= \sum_{\substack{(\zeta,\eta) \in \mathrm{Dec}_1(\xi) \\ \mathrm{size}(\zeta) \geq 1}} \mathscr{L}(\mathscr{A})(\zeta) \prod_{j=1}^{l_1} \mathscr{L}(\mathscr{A})^*(\eta_j^i) \\
&= \sum_{\substack{(\zeta,\eta) \in \mathrm{Dec}_1(\xi) \\ \mathrm{size}(\zeta) \geq 1}} \mathscr{L}(\mathscr{A})(\zeta) \prod_{j=1}^{l_1} \mathscr{L}(\mathscr{A}^*)(\eta_j^i).
\end{aligned}$$

Moreover we have in each summand

$$\mathscr{L}(\mathscr{A})(\zeta) = E_{1,1}^{\mathscr{A}}(I, \zeta, q_f) = \sum_{p_1, \ldots, p_s \in Q} E_s(p_1, \ldots, p_s, \sigma, q_f) \prod_{k=1}^{s} E_{1,1}^{\mathscr{A}}(I, \zeta_k, p_k).$$

Reordering the occurring terms, we obtain

$$\begin{aligned}
\mathscr{L}(\mathscr{A})^*(\xi) &= \sum_{\substack{(\zeta,\eta) \in \mathrm{Dec}_1(\xi) \\ \mathrm{size}(\zeta) \geq 1}} \Big( \sum_{p_1, \ldots, p_s \in Q} E_s(p_1, \ldots, p_s, \sigma, q_f) \prod_{k=1}^{s} E_{1,1}^{\mathscr{A}}(I, \zeta_k, p_k) \Big) \prod_{j=1}^{l_1} \mathscr{L}(\mathscr{A}^*)(\eta_j^i) \\
&= \sum_{p_1, \ldots, p_s \in Q} E_s(p_1, \ldots, p_s, \sigma, q_f) \Big( \sum_{\substack{(\zeta,\eta) \in \mathrm{Dec}_1(\xi) \\ \mathrm{size}(\zeta) \geq 1}} \prod_{k=1}^{s} E_{1,1}^{\mathscr{A}}(I, \zeta_k, p_k) \prod_{j=1}^{l_1} \mathscr{L}(\mathscr{A}^*)(\eta_j^i) \Big).
\end{aligned}$$

We moreover know that it suffices to sum over $p_1, \ldots, p_s \in Q^*$, as $\mathscr{A}$ is normalized. Therefore, using the definitions of $E$ and $E^*$, we see that

$$\begin{aligned}
\mathscr{L}(\mathscr{A})^*(\xi) &= \sum_{\substack{p_1, \ldots, p_s \\ \in Q^*}} E_s^*(p_1, \ldots, p_s, \sigma, q^I) \Big( \sum_{\substack{(\zeta,\eta) \in \mathrm{Dec}_1(\xi) \\ \mathrm{size}(\zeta) \geq 1}} \prod_{k=1}^{s} E_{1,1}^{\mathscr{A}}(I, \zeta_k, p_k) \prod_{j=1}^{l_1} \mathscr{L}(\mathscr{A}^*)(\eta_j^i) \Big) \\
&= \sum_{\substack{p_1, \ldots, p_s \\ \in Q^*}} E_s^*(p_1, \ldots, p_s, \sigma, q^I) \prod_{k=1}^{s} \Big( \sum_{\substack{(\zeta_k, \eta_k) \\ \in \mathrm{Dec}_1(\xi_k)}} E_{1,1}^{\mathscr{A}}(I, \zeta_k, p_k) \prod_{j=1}^{l_{k,1}} \mathscr{L}(\mathscr{A}^*)((\eta_k)_j^i) \Big).
\end{aligned}$$

For every $k \in [s]$ we know that if $p_k = q^I$, we have

$$\sum_{\substack{(\zeta_k,\eta_k)\\ \in \mathrm{Dec}_1(\xi_k)}} E^{\mathscr{A}}_{1,1}(I,\zeta_k,p_k) \prod_{j=1}^{l_{k,1}} \mathscr{L}(\mathscr{A}^*)((\eta_k)^i_j)$$

$$= E^{\mathscr{A}}_{1,1}(I,\pi^1_1,q^I)\mathscr{L}(\mathscr{A}^*)(\xi_k) + 0$$

$$= I_1(q^I)\mathscr{L}(\mathscr{A}^*)(\xi_k) = \mathscr{L}(\mathscr{A}^*)(\xi_k) = E^{\mathscr{A}^*}_{1,1}(I,\xi_k,q^I).$$

Here we used that if $\mathrm{size}(\zeta_k) \geq 1$, then $E^{\mathscr{A}}_{1,1}(I,\zeta_k,p_k) = 0$, as $p_k = q^I$ can not occur above terminal symbols.

If $p_k \neq q^I$, equation (12) yields

$$\sum_{\substack{(\zeta_k,\eta_k)\\ \in \mathrm{Dec}_1(\xi_k)}} E^{\mathscr{A}}_{1,1}(I,\zeta_k,p_k) \prod_{j=1}^{l_{k,1}} \mathscr{L}(\mathscr{A}^*)((\eta_k)^i_j) = E^{\mathscr{A}^*}_{1,1}(I,\xi_k,p_k).$$

We ultimately obtain

$$\mathscr{L}(\mathscr{A})^*(\xi) = \sum_{\substack{p_1,\ldots,p_s\\ \in Q^*}} E^*_s(p_1,\ldots,p_s,\sigma,q^I) \prod_{k=1}^{s} E^{\mathscr{A}^*}_{1,1}(I,\xi_k,p_k) = \mathscr{L}(\mathscr{A}^*)(\xi),$$

which concludes the proof. □

**Example 4.7.7.** We continue example 4.7.2. The automaton $\mathscr{A}^*$ satisfying $\mathscr{L}(\mathscr{A}^*) = \mathscr{L}(\mathscr{A})^*$ is constructed in the proof of Proposition 4.7.6 from $\mathscr{B}$ as follows (recall that $\mathscr{B}$ is a normalized $(1,1)$-WFA such that $\mathscr{L}(\mathscr{B}) = \mathscr{L}(\mathscr{A})$).

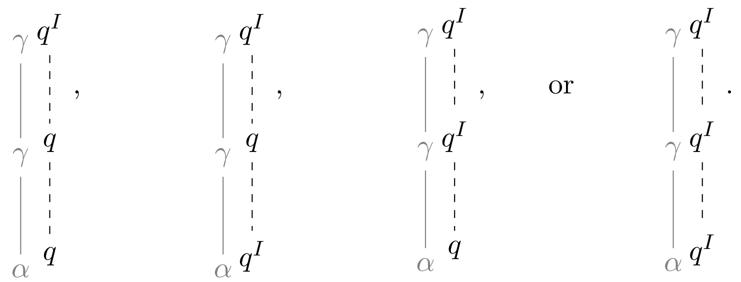
$$\mathscr{A}^* = (Q^*, \Sigma, S, (\mathbb{1}_{q^I}, F, E^*),$$

where $Q^* = (Q \cup \{q^I\}) \setminus \{f\} = \{q, q^I\}$ and $E^*$ is 0 except in the cases

$$\begin{aligned} E^*_0(\alpha,q') &= 1, & E^*_0(\beta,q') &= 1,\\ E^*_1(q,\gamma,q') &= 1, & E^*_1(q^I,\gamma,q') &= 2\cdot 1,\\ E^*_2(q,q,\sigma,q') &= 1, & E^*_2(q^I,q,\sigma,q') &= 2\cdot 1,\\ E^*_2(q,q^I,\sigma,q') &= 2\cdot 1, & E^*_2(q^I,q^I,\sigma,q') &= 4\cdot 1, \end{aligned}$$

for every $q' \in Q^*$.

Recall that $\xi = \gamma(\gamma(\alpha))$. We calculate the value of $\mathscr{L}(\mathscr{A}^*)(\xi)$. A run $\rho$ of $\mathscr{A}^*$ on $\xi$ ending in $q^I$ is either



By the definition of $E^*$, $\rho$ has cost $2^k \cdot 1$ if $\rho$ maps $k+1$ positions of $\xi$ to $q^I$. Therefore we obtain

$$\mathscr{L}(\mathscr{A}^*)(\xi) = 1 + 2 \cdot +2 \cdot 1 + 4 \cdot 1 = 9 \cdot 1.$$

Moreover we obtain the following general formula by the same argumentation as for $\xi$. Let $n \in \mathbb{N}$ and $\xi_n \coloneqq \gamma^n(\alpha)$. It holds that

$$\mathscr{L}(\mathscr{A}^*)(\xi_n) = \sum_{J \subseteq [n]} \Big( 2^{\#J} \cdot 1 \Big) = \Big( \sum_{J \subseteq [n]} 2^{\#J} \Big) \cdot 1 = \Big( \sum_{k=0}^{n} \binom{n}{k} \cdot 2^k \Big) \cdot 1 \overset{\star}{=} 3^n \cdot 1,$$

where in equation $\star$ we have used the binomial theorem $(2+1)^n = \sum_{k=0}^{n} \binom{n}{k} 2^k 1^{n-k}$. ■

# Chapter 5: Rec = Rat

We now introduce rational expressions and their corresponding weighted forest languages. Ultimately, we prove a Kleene-like theorem for weighted forest languages.

In this chapter, let $\Sigma$ be a ranked alphabet and $S$ a commutative semiring.

## 5.1 Rational Languages are Recognizable

**Definition 5.1.1.** We define the **set of weighted rational forest expressions** over $\Sigma$ and $S$, denoted $\mathrm{rat}(T(\Sigma), S)$, and for every such expression $r$ the **weighted forest language generated by** $r$, denoted $\langle\!| r |\!\rangle$, as the smallest biranked set $\mathscr{R} = \bigcup_{m,n\in\mathbb{N}_0} \mathscr{R}_n^m$ and the rectangular weighted forest language inductively by

1) $0 \in \mathscr{R}_n^m$ and $\langle\!| 0 |\!\rangle = 0$
   for every $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$,
2) $\xi \in \mathscr{R}_n^1$ and $\langle\!| \xi |\!\rangle = \mathbb{1}_\xi$
   for every $\xi \in T(\Sigma)_n^1$ for some $n \in \mathbb{N}_0$,
3) $a \cdot r_1 \in \mathscr{R}_n^1$ and $\langle\!| a \cdot r_1 |\!\rangle = a \cdot \langle\!| r_1 |\!\rangle$
   for every $a \in S$, $r_1 \in \mathscr{R}_n^1$, and $n \in \mathbb{N}_0$,
4) $r_1 + r_2 \in \mathscr{R}_n^1$ and $\langle\!| r_1 + r_2 |\!\rangle = \langle\!| r_1 |\!\rangle + \langle\!| r_2 |\!\rangle$
   for every $r_1, r_2 \in \mathscr{R}_n^1$ and $n \in \mathbb{N}_0$,
5) $r_1 \times r_2 \in \mathscr{R}_n^{m+m'}$ and $\langle\!| r_1 \times r_2 |\!\rangle = \langle\!| r_1 |\!\rangle \times \langle\!| r_2 |\!\rangle$
   for every $r_1 \in \mathscr{R}_n^m$,$r_2 \in \mathscr{R}_n^{m'}$, and $m, m', n \in \mathbb{N}_0$,
6) $r_1 \cdot r_2 \in \mathscr{R}_p^m$ and $\langle\!| r_1 \cdot r_2 |\!\rangle = \langle\!| r_1 |\!\rangle \cdot \langle\!| r_2 |\!\rangle$
   for every $r_1 \in \mathscr{R}_n^m$,$r_2 \in \mathscr{R}_p^n$ and $m, n, p \in \mathbb{N}_0$,
7) $r_1^* \in \mathscr{R}_1^1$ and $\langle\!| r_1^* |\!\rangle = \langle\!| r_1 |\!\rangle^*$
   for every $r_1 \in \mathscr{R}_1^1$ such that $\langle\!| r_1 |\!\rangle$ is proper.

We denote by $\mathrm{RAT}(T(\Sigma, S)$ the set of weighted forest languages generated by weighted rational forest expressions over $\Sigma$ and $S$. Moreover we denote $\mathrm{RAT}(T(\Sigma)_n^m, S) :=$ $\mathrm{RAT}(T(\Sigma), S)_n^m$ for the sake of consistency with the definition of recognizable weighted forest languages. ■

**Theorem 5.1.2.** For every $m, n \in \mathbb{N}_0$ it holds that

$$\mathrm{RAT}(T(\Sigma)_n^m, S) \subseteq \mathrm{REC}(T(\Sigma)_n^m, S).$$

*Proof.* We prove the claim by structural induction on rational expressions. Let $r \in \mathrm{rat}(T(\Sigma)^m_n, S)$. We list the corresponding propositions.

*Case 1* ($r = 0$): Proposition 4.1.1.
*Case 2* ($r = \xi$): Proposition 4.2.1.
*Case 3* ($r = a \cdot r_1$): Proposition 4.3.1.
*Case 4* ($r = r_1 + r_2$): Proposition 4.4.1.
*Case 5* ($r = r_1 \times r_2$): Proposition 4.5.2.
*Case 6* ($r = r_1 \cdot r_2$): Proposition 4.6.7.
*Case 7* ($r = r_1^*$): Proposition 4.7.6. □

## 5.2 Recognizable Languages are Rational

We show how the weighted language defined by a WFA $\mathscr{A}$ decomposes through rational operations into weighted languages with finite support. We then use the fact that weighted languages with finite support are rational in order to prove that $\mathscr{L}(\mathscr{A})$ is rational as well. We heavily use the ideas from the proof given in [6].

First, we prove the rationality of $\mathscr{L}(\mathscr{A})$ for $(1,0)$-WFA and then use closure under horizontal concatenation in order to prove the rationality for aribrary WFA.

We now outline the proof in the case of a $(1,0)$-WFA $\mathscr{A}$. Let $\xi$ be a forest. We want to understand the possible runs of $\mathscr{A}$ on $\xi$. Given a run $\rho$ and some state $p$ of $\mathscr{A}$, we can mark the topmost positions of $\xi$ where $\rho$ takes on the value $p$. This cut decomposes $\xi$ into a top part and some bottom parts. Moreover, this cut decomposes $\rho$ into a respective top part and some bottom parts. The top part of $\rho$, however, does not use $p$ anymore. Iteratively using this trick results in runs over the empty set. That is, runs that only assign a state at the root of a tree. We can easily find rational expressions that generate the same weighted languages as runs over the empty set. The described "cutting" of $\xi$ and $\rho$ then amounts to a vertical concatenation operation and a Kleene star.

In this subchapter, let $\mathscr{A} = (Q, \Sigma, S, \emptyset, F, E)$ be a (1,0)-WFA and moreover assume that $Q = \{q_1, \dots, q_n\}$ and $\#Q = n$. Furthermore, given a state $q \in Q$, we denote by $\mathrm{ind}(q)$ the index $i \in [n]$ such that $q = q_i$.

**Definition 5.2.1.** We define the (1,$n$)-WFA $\mathscr{A}' := (Q, \Sigma, S, I, F, E)$ with the leaf weight $I = (\mathbb{1}_{q_1}, \dots, \mathbb{1}_{q_n})$.

Let $P \subseteq Q$ and $q \in Q$. We define the map

$$S_{\mathscr{A}}(P, q) \colon T(\Sigma)^1_n \longrightarrow S,$$

where for every $\xi \in T(\Sigma)$

$$S_{\mathscr{A}}(P,q)(\xi) := \begin{cases} 0 & \text{, if } \exists i \in [n] \colon \xi = \pi^n_i \\ \displaystyle\sum_{\rho \in R^P_{\mathscr{A}'}(I,\xi,q)} c_{\mathscr{A}'}(\rho, \xi) & \text{, otherwise} \end{cases}$$

■

**Definition 5.2.2.** Let $\varphi\colon T(\Sigma)^1_n \longrightarrow S$ be a weighted forest language and $q \in Q$. We denote by $\varphi^q$ the weighted forest language $\psi\colon T(\Sigma \cup X_n)^1_1 \longrightarrow S$ that is obtained from $\varphi$ by viewing each variable except $x_{\mathrm{ind}(q)}$ as a terminal symbol.

Analogously, given $\xi \in T(\Sigma)^1_n$, we denote by $\xi^q$ the tree $\zeta \in T(\Sigma \cup X_n)^1_1$ obtained from $\xi$ by viewing each variable except $x_{\mathrm{ind}(q)}$ as a terminal symbol. ■

**Remark 5.2.3.** First note that $S_{\mathscr{A}}(P, q)$ maps empty forests to 0. We refer to this property as properness, which is in line with the original definition of properness.

Next we remark a decomposition correspondence for runs on a tree. Let $P \subseteq Q$, $p, q \in Q$ such that $p \notin P$, and $\xi \in T(\Sigma)^1_n$.

Let $\rho \in R^{P\cup\{p\}}_{\mathscr{A}'}(I, \xi, q)$. If $q \neq p$, define the set $\mathrm{cut}(\rho, \xi, p) \subseteq \mathrm{pos}(\xi)$ inductively as follows. Let

$$\mathrm{cut}(\rho, \pi^n_i, p) = \begin{cases} \{\varepsilon\} & , i = \mathrm{ind}(p) \\ \emptyset & , i \neq \mathrm{ind}(p) \end{cases}$$

and for any $k \geq 0$, $\sigma \in \Sigma^{(k)}$ and $t_1, \dots, t_k \in T_\Sigma(X_1)$

$$\mathrm{cut}(\rho, \langle 1, \sigma(t_1, \dots, t_k)\rangle, p) = \begin{cases} \{\varepsilon\} & , \rho(\varepsilon) = p \\ \bigcup_{i=1}^k \Big(i \cdot \mathrm{cut}(\rho_i, \langle 1, t_i\rangle, p)\Big) & , \rho(\varepsilon) \neq p, \end{cases}$$

where $\rho_i$ is the restriction of $\rho$ to $t_i$ for every $i \in [k]$. We illustrate this definition in Example 5.2.4.

If $q = p$, we define $\mathrm{cut}(\rho, \xi, p) \subseteq \mathrm{pos}(\xi)$ analogously, where we ignore the root of $\xi$, for example by replacing $\rho(\varepsilon)$ by some $\star \notin Q$.

In any case, denote $\{w_1, \dots, w_l\} = \mathrm{cut}(\rho, \xi, p)$ where $\#\,\mathrm{cut}(\rho, \xi, p) = l$. Cutting $\xi$ along $\mathrm{cut}(\rho, \xi, p)$, yields a 1-decomposition of $\xi^p$. More precisely, let $\zeta \in T(\Sigma)^1_n$ be the tree obtained from $\xi$ by replacing (for every $i \in [l]$) the subtree at position $w_i$ by the variable $x_{\mathrm{ind}(p)}$. Moreover, denote $\eta^{\mathrm{ind}(p)}_i = \xi|_{w_i}$. This yields an $n$-decomposition $(\zeta, \eta) \in \mathrm{Dec}_n(\xi)$ (where all the other $\eta^i_j$ are empty trees). In fact, $(\zeta, \eta)$ is a 1-decomposition of $\xi^p$. The restriction of $\rho$ to $\zeta$ is denoted by $\rho_0$ (where we define $\rho_0(w_i) \coloneqq q_{j(i)}$). The restriction of $\rho$ to $\eta^1_i$ is denoted $\rho_i$ for every $i \in [k]$. Ultimately, we have obtained the tuple

$$((\zeta, \eta), \rho_0, \rho_1 \dots, \rho_k).$$

Conversely, such a tuple $((\zeta, \eta), \rho_0, \rho_1 \dots, \rho_k)$, where the roots of the $\rho_i$ match the values $\rho_0(w_i)$, uniquely determines $\rho$, as the $\rho_i$ cover the entirety of $\xi$.

We have thus seen that

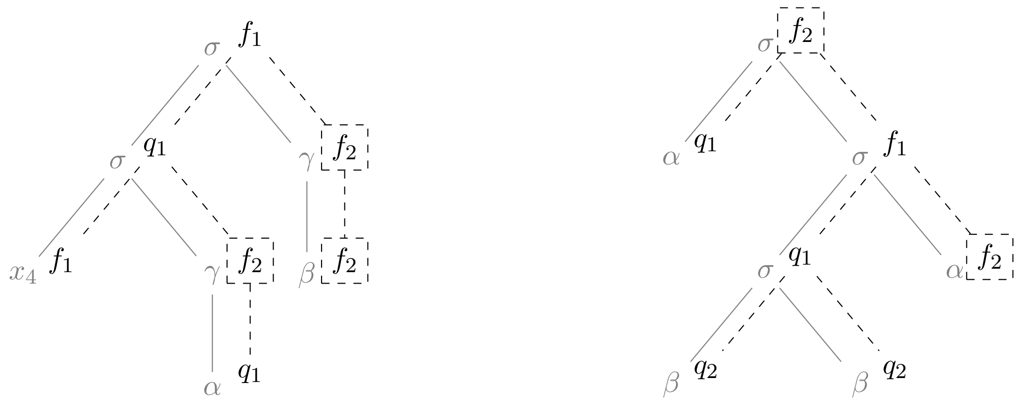
$$\begin{aligned} R^{P\cup\{p\}}_{\mathscr{A}}(I, \xi, q) \cong \{((\zeta, \eta), \rho_0, \rho_1 \dots, \rho_k) \mid & (\zeta, \eta) \in \mathrm{Dec}_1(\xi^p), \mathrm{size}(\zeta) \geq 1, \rho_0 \in R^P_{\mathscr{A}}(I, \zeta, q), \\ & \rho_0 \text{ assigns } p \text{ to the } x_{\mathrm{ind}(p)}\text{-positions}, \\ & \forall i \in [l_{\mathrm{ind}(p)}]\colon \rho_i \in R^{P\cup\{p\}}_{\mathscr{A}}(I, \eta^{\mathrm{ind}(p)}_i, p)\} \end{aligned}$$

Moreover, one sees immediately that the costs satisfy

$$c_{\mathscr{A}'}(\rho, \xi) = c_{\mathscr{A}'}(\rho_0, \zeta) \prod_{i=1}^{l_{\mathrm{ind}(p)}} c_{\mathscr{A}'}(\rho_i, \eta_i^{\mathrm{ind}(p)}).$$

■

**Example 5.2.4.** Let $\Sigma$ and $\mathscr{A}$ be as in Example 3.2.4. Consider the following trees $\xi_1$ and $\xi_2$ (from left to right, in gray) together with their respective runs $\rho_1$ and $\rho_2$ (slightly above the trees, dashed). In order to point out the positions at which these runs take on the state $f_2$, we draw dashed boxes around these positions.



It holds that $\mathrm{cut}(\rho_1, \xi_1, f_2) = \{11, 12, 2\}$ and $\mathrm{cut}(\rho_2, \xi_2, f_2) = \{22\}$. Note that $\rho_1(11) \neq f_2$, but $\xi_1(11) = x_4$ and $\mathrm{ind}(f_2) = 4$, hence $11 \in \mathrm{cut}(\rho_1, \xi_1, f_2)$. Note moreover that $\rho_2(\varepsilon)$, yet by definition we do not cut directly at the root position. ■

**Lemma 5.2.5.** Let $P \subseteq Q$ and $p, q \in Q$ such that $p \notin P$. Then,

$$S_{\mathscr{A}}(P \cup \{p\}, q) = S_{\mathscr{A}}(P, q)^p \cdot (S_{\mathscr{A}}(P, p)^p)^*,$$

where we interpret the right hand side as a map of type $T(\Sigma)^1_n \longrightarrow S$ (instead of $T(\Sigma \cup X_n)^1_1 \longrightarrow S$).

*Proof.* Let $\xi \in T(\Sigma)^1_1$. We show the desired equation by structural induction on $\xi$.

*Case 1:* Assume that $\xi = \langle n, x_i \rangle$. By properness, it holds that

$$S_{\mathscr{A}}(P \cup \{p\}, q)(\xi) = 0 = (S_{\mathscr{A}}(P, q)^p \cdot (S_{\mathscr{A}}(P, p)^p)^*)(\xi)$$

*Case 2:* Assume that $\xi = \langle 1, \alpha \rangle$ for some $\alpha \in \Sigma^{(0)}$. Note that there is exactly one run of $\mathscr{A}'$ on $\xi$ (in both cases, using $P$, and using $P \cup \{p\}$) ending in $q$. Therefore,

$$\begin{aligned} S_{\mathscr{A}}(P \cup \{p\}, q)(\xi) &= \sum_{\rho \in R^{P \cup \{p\}}_{\mathscr{A}'}(I, \xi, q)} c_{\mathscr{A}'}(\rho, \xi) \\ &= \sum_{\rho \in R^{P}_{\mathscr{A}'}(I, \xi, q)} c_{\mathscr{A}'}(\rho, \xi) = S_{\mathscr{A}}(P, q)(\xi). \end{aligned}$$

As $S_{\mathscr{A}}(P,q)$ is proper, we moreover obtain

$$\begin{aligned}
S_{\mathscr{A}}(P,q)(\xi) &= \sum_{(\zeta,\eta)\in\mathrm{Dec}_1(\xi^p)} S_{\mathscr{A}}(P,q)^p(\zeta) \prod_{j=1}^{l_{\mathrm{ind}(p)}} (S_{\mathscr{A}}(P,p)^p)^*(\eta_j^1)\\
&= (S_{\mathscr{A}}(P,q)^p \cdot (S_{\mathscr{A}}(P,p)^p)^*)(\xi)
\end{aligned}$$

*Case 3:* Assume that $\xi = \langle 1, \sigma(t_1,\dots,t_s)\rangle$ for some $\sigma \in \Sigma^{(s)}$, $s \geq 1$ such that the claim holds for all proper subtrees of $\xi$ (note that this is well-defined). It holds that

$$\begin{aligned}
S_{\mathscr{A}}(P\cup\{p\},q)(\xi) &= \sum_{\rho\in R_{\mathscr{A}'}^{P\cup\{p\}}(I,\xi,q)} c_{\mathscr{A}'}(\rho,\xi)\\
&\overset{\Phi}{=} \sum_{\substack{(\zeta,\eta)\in\mathrm{Dec}_1(\xi^p)\\ \mathrm{size}(\zeta)\geq 1}} \sum_{\substack{\rho_0\in R_{\mathscr{A}'}^{P}(I,\zeta,q),\\ \rho_j\in R_{\mathscr{A}'}^{P\cup\{p\}}(I,\eta_j^{\mathrm{ind}(p)},p)\\ \text{for every } j\in[l_{\mathrm{ind}(p)}]}} c_{\mathscr{A}'}(\rho_0,\zeta) \prod_{j=1}^{l_{\mathrm{ind}(p)}} c_{\mathscr{A}'}(\rho_j,\eta_j^{\mathrm{ind}(p)})\\
&= \sum_{\substack{(\zeta,\eta)\in\mathrm{Dec}_1(\xi^p)\\ \mathrm{size}(\zeta)\geq 1}} \Bigg( \sum_{\rho_0\in R_{\mathscr{A}'}^{P}(I,\zeta,q)} c_{\mathscr{A}'}(\rho_0,\zeta)\Bigg)\cdot\\
&\qquad \cdot \prod_{j=1}^{l_{\mathrm{ind}(p)}} \Bigg( \sum_{\rho_j\in R_{\mathscr{A}'}^{P\cup\{p\}}(I,\eta_j^{\mathrm{ind}(p)},p)} c_{\mathscr{A}'}(\rho_j,\eta_j^{\mathrm{ind}(p)})\Bigg)\\
&\overset{\curlyvee}{=} \sum_{\substack{(\zeta,\eta)\in\mathrm{Dec}_1(\xi^p)\\ \mathrm{size}(\zeta)\geq 1}} S_{\mathscr{A}}(P,q)^p(\zeta) \prod_{j=1}^{l_{\mathrm{ind}(p)}} (S_{\mathscr{A}}(P\cup\{p\},p)^p + \mathbb{1}_{\pi^n_{\mathrm{ind}(p)}})(\eta_j^{\mathrm{ind}(p)})\\
&\overset{\mathrm{IH}}{=} \sum_{\substack{(\zeta,\eta)\in\mathrm{Dec}_1(\xi^p)\\ \mathrm{size}(\zeta)\geq 1}} S_{\mathscr{A}}(P,q)^p(\zeta)\cdot\\
&\qquad \cdot \prod_{j=1}^{l_{\mathrm{ind}(p)}} (S_{\mathscr{A}}(P,p)^p \cdot (S_{\mathscr{A}}(P,p)^p)^* + \mathbb{1}_{\pi^n_{\mathrm{ind}(p)}})(\eta_j^{\mathrm{ind}(p)})\\
&\overset{\pm}{=} \sum_{\substack{(\zeta,\eta)\in\mathrm{Dec}_1(\xi^p)\\ \mathrm{size}(\zeta)\geq 1}} S_{\mathscr{A}}(P,q)^p(\zeta) \prod_{j=1}^{l_{\mathrm{ind}(p)}} (S_{\mathscr{A}}(P,p)^p)^*(\eta_j^{\mathrm{ind}(p)})\\
&= (S_{\mathscr{A}}(P,q)^p \cdot (S_{\mathscr{A}}(P,p)^p)^*)(\xi).
\end{aligned}$$

Equation $\Phi$ applies Remark 5.2.3 and the fact that if $\rho_0$ does not map the $x_i$-positions of $\zeta$ to $p$, its cost vanishes.

Equation $\curlyvee$ replaces the sums for $\zeta$ and $\eta_j^{\mathrm{ind}(p)}$ by the respective weighted languages

$S_{\mathscr{A}}$. As it is possible that $\eta_j^{\mathrm{ind}(p)} = \pi^n_{\mathrm{ind}(p)}$ for some $j \in [l_{\mathrm{ind}(p)}]$, in which case

$$\sum_{\rho_j \in R^{P\cup\{p\}}_{\mathscr{A}'}(I,\eta_j^{\mathrm{ind}(p)},q)} c_{\mathscr{A}'}(\rho_j, \eta_j^{\mathrm{ind}(p)}) = 1 \neq 0 = S_{\mathscr{A}}(P \cup \{p\}, p)(\eta_j^{\mathrm{ind}(p)}),$$

equation $\curlyvee$ moreover adds $\mathbb{1}_{\pi^n_{\mathrm{ind}(p)}}$.

In equation $\dotplus$ we used Lemma 4.7.5. $\square$

**Lemma 5.2.6.** Let $q \in Q$. It holds that

$$S_{\mathscr{A}}(\emptyset, q) \in \mathrm{RAT}(T(\Sigma)^1_n, S).$$

*Proof.* Let $\xi \in T(\Sigma)^1_n$. If $\mathrm{size}(\xi) = 0$, we have $S_{\mathscr{A}}(\emptyset, q) = 0$ by definition.

If $\mathrm{size}(\xi) = 1$, there exists $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $i_1, \ldots, i_k \in [n]$ such that we have $\xi = \langle n, \sigma(x_{i_1}, \ldots, x_{i_k})\rangle$. We denote $a_\sigma^{i_1,\ldots,i_k} := S_{\mathscr{A}}(\emptyset, q)(\xi)$.

If $\mathrm{size}(\xi) \geq 2$, there is no run on $\xi$ using $\emptyset$, whence again $S_{\mathscr{A}}(\emptyset, q) = 0$.

Alltogether we obtain

$$S_{\mathscr{A}}(\emptyset, q) = \sum_{\sigma \in \Sigma} \sum_{i_1,\ldots,i_k \in [n]} a_\sigma^{i_1,\ldots,i_k} \mathbb{1}_{\sigma(x_{i_1},\ldots,x_{i_k})},$$

which is a finite sum. Cases 2 and 3 in the definition of $\mathrm{RAT}(T(\Sigma)^1_1, S)$ yield the claim. $\square$

**Proposition 5.2.7.** Let $P \subseteq Q$ and $q \in Q$. It holds that

$$S_{\mathscr{A}}(P, q) \in \mathrm{RAT}(T(\Sigma)^1_n, S).$$

*Proof.* The proof is by induction on $\#P$. The induction base $\#P = 0$ is Lemma 5.2.6 and the induction step $P \rightsquigarrow P \cup \{p\}$ is Lemma 5.2.5.

Note that the switching between $T(\Sigma)^1_n$ and $T(\Sigma \cup X_n)^1_1$ does not affect the rationality. In fact, we could have done the proof of Lemma 5.2.5 without this trick (by inflating the occurring weighted tree languages by characteristic functions of empty trees), yet it was notationally more convenient to switch between the interpretations of $X_n$ as terminals and as variables. $\square$

**Proposition 5.2.8.** It holds that

$$\mathscr{L}(\mathscr{A}) = \sum_{q \in F} \Big( S_{\mathscr{A}}(Q, q) \cdot 0 \Big). \tag{15}$$

In particular, $\mathscr{L}(\mathscr{A}) \in \mathrm{RAT}(T(\Sigma)^1_0, S)$.

*Proof.* Let $\xi \in T(\Sigma)^1_0$. It holds that

$$(S_{\mathscr{A}}(Q, q) \cdot 0)(\xi) = \sum_{(\zeta,\eta) \in \mathrm{Dec}_1(\xi)} S_{\mathscr{A}}(Q, q)(\zeta) \prod_{j=1}^{l_1} 0(\eta_j^1) = S_{\mathscr{A}}(Q, q)(\xi).$$

By assumption $\xi \neq x_1$, whence $c_{\mathscr{A}'}(\rho, \xi) = c_{\mathscr{A}}(\rho, \xi)$ for every $\rho \in R_{\mathscr{A}'}(I, \xi, q)$. Moreover it surely holds that $R_{\mathscr{A}'}(I, \xi, q) = R_{\mathscr{A}}(I, \xi, q)$. We conclude that

$$S_{\mathscr{A}}(Q, q)(\xi) = \sum_{\rho \in R_{\mathscr{A}'}(I,\xi,q)} c_{\mathscr{A}'}(\rho, \xi) = \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} c_{\mathscr{A}}(\rho, \xi),$$

and hence

$$\sum_{q \in F} \Big( S_{\mathscr{A}}(Q, q) \cdot 0 \Big)(\xi) = \sum_{q \in F} \sum_{\rho \in R_{\mathscr{A}}(I,\xi,q)} c_{\mathscr{A}}(\rho, \xi) = \mathscr{L}(\mathscr{A})(\xi).$$

□

**Example 5.2.9.** Consider $\Sigma$ from Example 3.2.4 and let $\mathscr{A} = (\{q_1, q_2\}, \Sigma, S, (), \{q_1, q_2\}, E)$ be the $(1, 0)$-WFA, where $E$ vanishes except in the following cases.

$$\begin{aligned} E_2(q_1, q_1, \sigma, q_1) &= 3 \cdot 1, & E_2(q_2, q_2, \sigma, q_2) &= 1, \\ E_1(q_1, \sigma, q_1) &= 1, & E_1(q_2, \sigma, q_2) &= 5 \cdot 1, \\ E_0(\sigma, q_1) &= 1, & E_0(\sigma, q_2) &= 1, \\ E_0(\sigma, q_1) &= 2 \cdot 1, & E_0(\sigma, q_2) &= 1. \end{aligned}$$

One can easily check that

$$\mathscr{L}(\mathscr{A}) = (2 \cdot 1)^{\#\operatorname{pos}_\alpha(\xi)} \cdot (3 \cdot 1)^{\#\operatorname{pos}_\sigma(\xi)} + (5 \cdot 1)^{\#\operatorname{pos}_\gamma(\xi)}.$$

We use Lemmas 5.2.6 and 5.2.5 and Proposition 5.2.8 to find some $r \in \operatorname{rat}(T(\Sigma), S)$ such that $\langle\!| r |\!\rangle = \mathscr{L}(\mathscr{A})$.

The definition of $S_{\mathscr{A}}$ immediately yields

$$\begin{aligned} S_{\mathscr{A}}(\emptyset, q_1) &= \langle\!| 3 \cdot \sigma(x_1, x_1) + \phantom{5 \cdot} \gamma(x_1) + \beta + 2 \cdot \alpha |\!\rangle, \text{ and} \\ S_{\mathscr{A}}(\emptyset, q_2) &= \langle\!| \phantom{3 \cdot} \sigma(x_1, x_1) + 5 \cdot \gamma(x_1) + \beta + \phantom{2 \cdot} \alpha |\!\rangle. \end{aligned}$$

The following formulas follow from Lemma 5.2.5.

$$\begin{aligned} S_{\mathscr{A}}(\{q_1, q_2\}, q_1) &= S_{\mathscr{A}}(\{q_2\}, q_1)^{q_1} \cdot (S_{\mathscr{A}}(\{q_2\}, q_1)^{q_1})^* = (S_{\mathscr{A}}(\{q_2\}, q_1)^{q_1})^*, \\ S_{\mathscr{A}}(\{q_1, q_2\}, q_2) &= S_{\mathscr{A}}(\{q_1\}, q_2)^{q_2} \cdot (S_{\mathscr{A}}(\{q_1\}, q_2)^{q_2})^* = (S_{\mathscr{A}}(\{q_1\}, q_2)^{q_2})^*, \end{aligned}$$

$$\begin{aligned} S_{\mathscr{A}}(\{q_2\}, q_1) &= S_{\mathscr{A}}(\emptyset, q_1)^{q_2} \cdot (S_{\mathscr{A}}(\emptyset, q_2)^{q_2})^*, \\ S_{\mathscr{A}}(\{q_1\}, q_2) &= S_{\mathscr{A}}(\emptyset, q_2)^{q_1} \cdot (S_{\mathscr{A}}(\emptyset, q_1)^{q_1})^*. \end{aligned}$$

Using Proposition 5.2.8, we find that

$$\begin{aligned} \mathscr{L}(\mathscr{A}) &= S_{\mathscr{A}}(\{q_1, q_2\}, q_1) \cdot 0 + S_{\mathscr{A}}(\{q_1, q_2\}, q_2) \cdot 0 \\ &= (S_{\mathscr{A}}(\{q_2\}, q_1)^{q_1})^* \cdot 0 + (S_{\mathscr{A}}(\{q_1\}, q_2)^{q_2})^* \cdot 0 \\ &= ((S_{\mathscr{A}}(\emptyset, q_1)^{q_2} \cdot (S_{\mathscr{A}}(\emptyset, q_2)^{q_2})^*)^{q_1})^* \cdot 0 + \\ &\qquad + ((S_{\mathscr{A}}(\emptyset, q_2)^{q_1} \cdot (S_{\mathscr{A}}(\emptyset, q_1)^{q_1})^*)^{q_2})^* \cdot 0. \end{aligned}$$

Plugging the rational expressions for $S_{\mathscr{A}}(\emptyset, q_1)$ and $S_{\mathscr{A}}(\emptyset, q_2)$ into the above equation results in $r$.

One can now check that in fact $\langle\!\langle r \rangle\!\rangle = \mathscr{L}(\mathscr{A})$. We iteratively evalutate the subexpressions of $r$ using the definitions of the corresponding operations. The detailed calculations tend to become lengthy, hence we only provide the solutions.

$$\begin{aligned}(S_{\mathscr{A}}(\emptyset, q_1)^{q_1})^* &= (2 \cdot 1)^{\#\,\mathrm{pos}_\alpha(\xi)} \cdot (3 \cdot 1)^{\#\,\mathrm{pos}_\sigma(\xi)} \\ (S_{\mathscr{A}}(\emptyset, q_2)^{q_2})^* &= (5 \cdot 1)^{\#\,\mathrm{pos}_\gamma(\xi)}\end{aligned}$$

$$S_{\mathscr{A}}(\emptyset, q_1)^{q_2} \cdot (S_{\mathscr{A}}(\emptyset, q_2)^{q_2})^* = \langle\!\langle 3 \cdot \sigma(x_1, x_1) + \gamma(x_1) + \beta + 2 \cdot \alpha \rangle\!\rangle \tag{16}$$

$$S_{\mathscr{A}}(\emptyset, q_2)^{q_1} \cdot (S_{\mathscr{A}}(\emptyset, q_1)^{q_1})^* = \langle\!\langle \sigma(x_1, x_1) + 5 \cdot \gamma(x_1) + \beta + \alpha \rangle\!\rangle \tag{17}$$

$$\begin{aligned}((S_{\mathscr{A}}(\emptyset, q_1)^{q_2} \cdot (S_{\mathscr{A}}(\emptyset, q_2)^{q_2})^*)^{q_1})^* \cdot 0 &= (2 \cdot 1)^{\#\,\mathrm{pos}_\alpha(\xi)} \cdot (3 \cdot 1)^{\#\,\mathrm{pos}_\sigma(\xi)} \\ ((S_{\mathscr{A}}(\emptyset, q_2)^{q_1} \cdot (S_{\mathscr{A}}(\emptyset, q_1)^{q_1})^*)^{q_2})^* \cdot 0 &= (5 \cdot 1)^{\#\,\mathrm{pos}_\gamma(\xi)}\end{aligned}$$

and hence $\langle\!\langle r \rangle\!\rangle = \mathscr{L}(\mathscr{A})$, as claimed.

Note how the subformulas in equations 16 and 17 degenerated into $S_{\mathscr{A}}(\emptyset, q_1)$ and $S_{\mathscr{A}}(\emptyset, q_2)$, respectively. This is due to the fact that $q_1$ and $q_2$ can not be mixed in $\mathscr{A}$ without getting vanishing costs. This goes hand in hand with the intuition for $\mathscr{A}$ as an automaton generating the sum of two weighted languages. Both these summand languages are generated by a single state ($q_1$ and $q_2$, respectively) and their state behaviours can not mix (by construction of an automaton for the sum of two weighted languages). ■

**Theorem 5.2.10.** Let $m \in \mathbb{N}_0$. It holds that

$$\mathrm{REC}(T(\Sigma)_0^m, S) \subseteq \mathrm{RAT}(T(\Sigma)_0^m, S).$$

*Proof.* Let $\varphi \in \mathrm{REC}(T(\Sigma)_0^m, S)$ with rectangular components $\varphi_1, \ldots, \varphi_m \in \mathrm{REC}(T(\Sigma)_0^1, S)$. For every $i \in [m]$ there exists a (1,0)-WFA

$$\mathscr{A}_i = (Q_i, \Sigma, S, \emptyset, F_i, E_i),$$

such that $\mathscr{L}(\mathscr{A}_i) = \varphi_i$. By Proposition 5.2.8 there are $r_1, \ldots, r_m \in \mathrm{RAT}(T(\Sigma)_0^1, S)$ such that

$$\mathscr{L}(\mathscr{A}_i) = r_i$$

for every $i \in [m]$. Therefore,

$$\begin{aligned}\varphi = \varphi_1 \times \cdots \times \varphi_m \qquad &= \mathscr{L}(\mathscr{A}_1) \times \cdots \times \mathscr{L}(\mathscr{A}_m) \\ = \langle\!\langle r_1 \rangle\!\rangle \times \cdots \times \langle\!\langle r_m \rangle\!\rangle \qquad &= \langle\!\langle r_1 \times \cdots \times r_m \rangle\!\rangle,\end{aligned}$$

which proves the claim. □

**Corollary 5.2.11.** Let $m, n \in \mathbb{N}_0$. It holds that

$$\mathrm{REC}(T(\Sigma)_n^m, S) \subseteq \mathrm{RAT}(T(\Sigma \cup X_n)_0^m, S).$$

*Proof.* After identification of $\mathrm{REC}(T(\Sigma)_n^m, S)$ and $\mathrm{REC}(T(\Sigma \cup X_n)_0^m, S)$, this follows immediately from Theorem 5.2.10. $\square$

We can thus conclude this thesis with our Kleene theorem

**Theorem 5.2.12.** Let $m, n \in \mathbb{N}_0$. It holds that

$$\mathrm{REC}(T(\Sigma)_n^m, S) = \mathrm{RAT}(T(\Sigma \cup X_n)_0^m, S)$$

and in particular

$$\mathrm{REC}(T(\Sigma)_0^m, S) = \mathrm{RAT}(T(\Sigma)_0^m, S).$$

*Proof.* Theorems 5.2.10 and 5.1.2. $\square$

# Chapter 6: Conclusion

## 6.1 Prospectus

We have combined the Kleene results from [6] and [18] to arrive at a Kleene result for weighted forest languages.

After providing insight into the mathematical tools, we introduced forests and weighted forest automata. A first major result was the rectangularity of recognizable weighted forest languages. This opened up a connection between weighted forest languages and weighted tree languages. We moreover proved two normal forms that would become very helpful throughout the remaining thesis.

We defined a Kleene star operation on weighted forest languages and saw that a property called *properness* became relevant in order to arrive at a well-defined operation. As an alternative to allowing only proper weighted languages inside a Kleene star, we could have also restricted our case to complete semirings. However, properness is not a restriction on the assumptions of our Kleene result. It is simply a fact that in our automata analysis (the generation of a rational expression accepting the same language as an automaton) the terms occurring inside Kleene stars are proper weighted languages. Hence the rational operations need not include a Kleene star for non-proper weighted languages.

The two directions in the proof of our Kleene result were expression synthesis (the construction of an automaton generating the same language as a rational expression) and automata analysis.

In our synthesis, we could have simply cited many of the results from the tree case [6]. However, it was important for us to demonstrate the formal nature of forests. Hence we explicitly provided all necessary constructions, proofs, and examples. Moreover the proof for vertical concatenation was done in the general case of forests, instead of using the rectangularity property.

The proof of closure under Kleene star is done by utilizing the rectangularity property and then executing the construction given in [6] in the tree case. Writing this thesis, we also made an attempt to prove closure of (arbitrary) recognizable weighted forest languages under Kleene star, yet were not able to find an understandable proof that was linked to some intuition.

In our analysis, we redid the proof presented in [6] and lifted the result to forest languages through the rectangularity property.

## 6.2 Future Work

In this section we want to give a very brief outlook at what one could do next with respect to the theory of weighted forest languages.

First of all, it seems worthwhile to do the automaton analysis from this thesis without using the rectangularity property. In [18], Straßburger provides a proof for the unweighted case. However, he uses a Kleene star operation for forests of arbitrary horizontal size (in contrast to forests of horizontal size 1, as we do). We believe that using his idea of proof in the weighted case might shorten the proof of the analysis without losing its intuition.

Using the rectangularity property, one could also dive into other famous theorems and lift them to the forest case. Some very prominent theorems are *Büchi's Theorem* (which links recognizable languages to languages generated by monadic first order logic), *Nivat's Theorem* (which provides a decomposition formula for languages generated by transducers into homomorphisms and regular languages), and *pumping lemmas* (which provide criteria to decide that a language is not recognizable).

One could also introduce different automata models than the one presented in this thesis. The rectangularity of (our) recognizable languages makes our theory less appealing, as proofs tend to break down into the tree case very easily. Other automata models might process entire forests at once, instead of single symbols. Because of the partial monoid structure in magmoids (the vertical concatenation), this might be similar to the string case and bring up interesting connections between forest languages and string languages. However, these ideas are very speculative and to our knowledge, no efforts have been made in these directions.

## References

[1] A. Arnold and M. Dauchet. Théorie des magmoïdes. I. *RAIRO Inform. Théor.*, 12(3):235–257, v, 1978.

[2] A. Arnold and M. Dauchet. Théorie des magmoïdes. II. *RAIRO Inform. Théor.*, 13(2):135–154, 1979.

[3] M. F. Atiyah and I. G. Macdonald. *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969.

[4] J. Berstel and C. Reutenauer. Recognizable formal power series on trees. *Theoret. Comput. Sci.*, 18(2):115–148, 1982.

[5] F. Dörband. Context-free grammars over the free magmoid. Dresden, Technische Universität, 2016.

[6] M. Droste, C. Pech, and H. Vogler. A Kleene theorem for weighted tree automata. *Theory Comput. Syst.*, 38(1):1–38, 2005.

[7] Manfred Droste, Werner Kuich, and Heiko Vogler, editors. *Handbook of weighted automata*. Monographs in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2009.

[8] Samuel Eilenberg. *Automata, languages, and machines. Vol. A*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York, 1974. Pure and Applied Mathematics, Vol. 58.

[9] Joost Engelfriet. Tree automata and tree grammars. *ArXiv*, abs/1510.02036, 2015.

[10] Ferenc Gécseg and Magnus Steinby. *Tree automata*. Akadémiai Kiadó (Publishing House of the Hungarian Academy of Sciences), Budapest, 1984.

[11] J. A. Goguen, J. W. Thatcher, E. G. Wagner, and J. B. Wright. Initial algebra semantics and continuous algebras. *J. Assoc. Comput. Mach.*, 24(1):68–95, 1977.

[12] Jonathan S. Golan. *Semirings and their applications*. Kluwer Academic Publishers, Dordrecht, 1999. Updated and expanded version of ıt The theory of semirings, with applications to mathematics and theoretical computer science [Longman Sci. Tech., Harlow, 1992; MR1163371 (93b:16085)].

[13] Udo Hebisch and Hanns Joachim Weinert. *Halbringe*. Teubner Studienbücher Mathematik. [Teubner Mathematical Textbooks]. B. G. Teubner, Stuttgart, 1993. Algebraische Theorie und Anwendungen in der Informatik. [Algebraic theory and applications in computer science].

[14] S. C. Kleene. Representation of events in nerve nets and finite automata. In *Automata studies*, Annals of mathematics studies, no. 34, pages 3–41. Princeton University Press, Princeton, N. J., 1956.

[15] M. P. Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.

[16] Helmut Seidl. Finite tree automata with cost functions. In *CAAP '92 (Rennes, 1992)*, volume 581 of *Lecture Notes in Comput. Sci.*, pages 279–299. Springer, Berlin, 1992.

[17] Helmut Seidl. Finite tree automata with cost functions. *Theoretical Computer Science*, 126(1):113–142, April 1994.

[18] L. Straßburger. A Kleene theorem for forest languages. In *Language and automata theory and applications*, volume 5457 of *Lecture Notes in Comput. Sci.*, pages 715–727. Springer, Berlin, 2009.

[19] J. W. Thatcher and J. B. Wright. Generalized finite automata theory with an application to a decision problem of second-order logic. *Math. Systems Theory*, 2:57–81, 1968.