International Journal of Foundations of Computer Science

Automata with Auxiliary Weights*

Peter Kostolányi and Branislav Rovan

Department of Computer Science, Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Mlynská dolina, 842 48 Bratislava, Slovakia
kostolanyi@fmph.uniba.sk
rovan@dcs.fmph.uniba.sk

Received (Day Month Year)
Revised (Day Month Year)
Accepted (Day Month Year)
Communicated by (xxxxxxxxxx)

We introduce a new generalization of weighted automata – automata with auxiliary weights. They allow to compute several quantities not directly computable over semirings by generalizing both semiring addition and semiring multiplication. Moreover, our automata retain key properties of ordinary weighted automata. We prove that automata over multi-hemirings of Droste and Kuich [7] can be modelled using our framework, and that automata with auxiliary weights can be used to describe several additional settings. Finally, we introduce rational expressions with auxiliary weights and prove that they are equivalent to our automata.

Keywords: Automaton with auxiliary weights; auxiliary weighting structure, weighted automaton, multi-hemiring.

1. Introduction

Weighted finite automata can be described as finite automata with transitions carrying weights (usually) taken from a semiring, which model some quantity related to the execution of a given transition, such as cost, reward, reliability, probability, etc. The classical theory of weighted automata has grown into a vast body of knowledge – for an overview of principal topics, we refer the reader to the handbook [6].

Several modifications of weighted automata (both on finite and infinite words) have recently been introduced, in which semirings are replaced by more general weighting structures. Investigations of this kind have been initiated by Chatterjee, Doyen, and Henzinger, who have studied weighted automata computing, e.g., the maximal average cost of executing a transition in a run on a given word [3–5]. Motivated by problems in verification, such automata describe important properties of systems, which cannot be directly modelled over semirings.

*This work has been supported by the grants VEGA 1/0967/16 and UK/258/2015.

These studies arising from practice have inspired an effort to introduce a class of weighting structures more general than semirings, in which a reasonable theory of weighted automata may be built, and which allows to express quantitative properties similar to those of Chatterjee, Doyen, and Henzinger [3–5]. Along this line, Droste and Meinecke have defined *automata over valuation monoids* [8–10, 14], and later Droste and Kuich have introduced *automata over multi-hemirings* [7]. Already the former model has managed to successfully incorporate many of the quantitative properties, which have been of principal interest for, e.g., Chatterjee, Doyen, and Henzinger [3–5]; the latter model may be viewed as a more direct approach to the former. Based on the nature of quantitative properties they are intended to capture, both these models generalize semiring *multiplication* by taking some more general function. Alternative approaches to generalizations of weighted automata include, e.g., those of Alur et al. [1] and of Cadilhac, Krebs, and Limaye [2].

One of the quantities which is not directly computable over semirings and which we are interested in is the average cost of a run on a given word, where the cost of a single run is computed by ordinary addition on real numbers. In order to capture quantitative properties like this, we cannot utilize automata over valuation monoids or automata over multi-hemirings, as we need to generalize semiring *addition* (in contrast to generalizing just semiring *multiplication*).

For this reason, we aim to come up with a framework for generalized weighted automata, which allows to generalize both semiring addition and semiring multiplication. This framework should be close enough to ordinary weighted automata – for instance, the behaviour of our automata should be computable via matrix multiplication over some suitable structure. On the other hand, it should be expressive enough to encompass some non-trivial settings. That is, we would like to build our framework based on the following requirements:

(R1) It should be possible to realize settings involving generalization of semiring addition. In particular, it should be possible to compute the average cost of a run on a given word and some similar quantitative properties.

(R2) It should be possible to generalize semiring multiplication as well. In particular, our model should extend automata over multi-hemirings.

(R3) The behaviour of our automata should be computable via matrix multiplication over some suitable underlying structure.

(R4) Transition weights in our automata should be "consistent" with the quantity to be computed.
This means that our generalized multiplication should have a property that a (trivial) "product" of one term x should equal x. Similarly, a "sum" of one term x should be x. In particular, in an automaton $\mathcal{A}$ consisting of a single transition on c leading from an initial state to a final state, the weight of this transition should equal the weight of the *word* c in the behaviour of $\mathcal{A}$. For instance, we shall not allow "addition" defined by "two times the ordinary sum of real numbers" – this clearly does not have the desired property.

To come up with a framework based on requirements (R1) to (R4) means to find a reasonable compromise between the former two and the latter two requirements. While (R1) and (R2) encourage the framework to be as general as possible, (R3) and (R4) pose restrictions that it has to satisfy. In fact, (R3) and (R4) seem to rule out some otherwise interesting quantities. However, there already exists a fully abstract model, which "maximizes" (R1) and (R2), but does not follow (R3) and (R4): *multi-weighted automata* of Droste and Perevoshchikov [11]. Practically any quantitative property can be realized in this model. Our aim is to introduce a framework that might be less general than multi-weighted automata, but follows requirements (R3) and (R4). Intuitively, these requirements ensure that the computation of behaviour remains "close to the runs in an automaton", just like for ordinary weighted automata over semirings.

Based on requirements (R1) to (R4), we introduce *automata with auxiliary weights*. The main idea is that we take an automaton $\mathcal{A}$ with weights from some set W with no structure in general. When the behaviour is to be computed, each transition of $\mathcal{A}$ is labelled by an additional auxiliary weight from some set X with no structure in general as well. However, we suppose that the set $W \times X$, possibly extended by some set E, forms a semiring or a hemiring – so that $\mathcal{A}$ corresponds to an automaton $\mathcal{A}'$ over $(W \times X) \cup E$, which is a semiring or a hemiring. The behaviour of $\mathcal{A}$ is the "first projection" of the behaviour of $\mathcal{A}'$ – that is, if the weight of some word is in E, then that word does not belong to the support of the behaviour of $\mathcal{A}$; if it is in $W \times X$, then we take the first component as the weight of the word in the behaviour of $\mathcal{A}$. Automata with auxiliary weights are in fact just like usual weighted automata. We find this to be one of their most appealing features.

Some similar ideas have already been exhibited by Filiot, Gentilini and Raskin [12], who have studied functional weighted automata over groups, in which there is no need for addition and where multiplication is realized over a group. In order to obtain group structure for some particular quantities, they have used a construction which can be described as introducing auxiliary weights. Automata with auxiliary weights may be viewed as a generalization of their approach.

We shall prove that automata with auxiliary weights extend automata over multi-hemirings of Droste and Kuich [7], so that automata over multi-hemirings can be viewed as a particular case of automata with auxiliary weights over hemirings, and even of automata with auxiliary weights over semirings. Moreover, we shall observe that several settings in which semiring *addition* is replaced by a more general function can be handled with ease by our framework.

A Kleene-like theorem has been obtained both for automata over valuation monoids [9] and for automata over multi-hemirings [7]. We show that the result for multi-hemirings generalizes to our framework – we introduce rational expressions with auxiliary weights and prove that they are equivalent to automata with auxiliary weights. The (conceptual) simplicity of the definition of our framework makes this result almost a direct consequence of results already known.

2. Preliminary Definitions

We shall denote by $\mathbb{N}$ the set of all *nonnegative* integers, while by $\mathbb{N}_+$ we shall denote the set of all *positive* integers. By $\mathbb{R}$, we shall denote the set of all real numbers, and we shall denote the set of all *nonnegative* real numbers by $\mathbb{R}_{\geq 0}$.

A *hemiring* is a quadruple $H = (H, +, \cdot, 0)$, where H is a set, $+$ and $\cdot$ are binary operations on H (addition and multiplication), and 0 is a distinguished element of H, such that $(H, +, 0)$ is a commutative monoid, $(H, \cdot)$ is a semigroup, multiplication is (both left- and right-) distributive over addition, and the identity $0 \cdot a = a \cdot 0 = 0$ holds for all a in H. A *semiring* is a quintuple $S = (S, +, \cdot, 0, 1)$, such that $(S, +, \cdot, 0)$ is a hemiring and $(S, \cdot, 1)$ is a monoid. In other words, a hemiring is a semiring (possibly) without a multiplicative identity.

Let Σ be an alphabet and S be a semiring. A *formal power series* over S and Σ is a mapping r from Σ^* to S. The value of r on a word w is usually denoted by (r, w) and the function r is written as a formal sum

$$r = \sum_{w \in \Sigma^*} (r, w)w.$$

The *sum of power series* r_1 and r_2 is a series $r_1 + r_2$, such that for all w in Σ^*,

$$(r_1 + r_2, w) = (r_1, w) + (r_2, w).$$

The notation in use is justified by the definition of the *(Cauchy) product of power series* r_1 and r_2, defined to be a power series $r_1 \cdot r_2$, such that for all w in Σ^*,

$$(r_1 \cdot r_2, w) = \sum_{w_1 w_2 = w} (r_1, w_1)(r_2, w_2).$$

The collection of all power series over S and Σ is denoted by $S\langle\langle \Sigma^* \rangle\rangle$, and it is a well-known fact that $S\langle\langle \Sigma^* \rangle\rangle$, equipped with the operations defined above, forms a semiring [6] called the *semiring of formal power series over S and Σ*. The *support* of a power series r is the set $\mathrm{supp}(r) = \{w \in \Sigma^* \mid (r, w) \neq 0\}$. A *polynomial* over S and Σ is a power series in $S\langle\langle \Sigma^* \rangle\rangle$ with finite support. The collection of all polynomials over S and Σ is denoted by $S\langle \Sigma^* \rangle$. Furthermore, for all finite $L \subseteq \Sigma^*$, we shall write $S\langle L \rangle$ for the collection of all power series in $S\langle\langle \Sigma^* \rangle\rangle$ with support in L, and we shall write $S\langle\langle L \rangle\rangle$ in the case L is infinite. A power series r in $S\langle\langle \Sigma^* \rangle\rangle$ is said to be *proper* if $(r, \varepsilon) = 0$, and the collection of all such series is denoted by $S\langle\langle \Sigma^+ \rangle\rangle$, which complies with the notation above.

Similar definitions related to formal power series can also be made for a hemiring H considered instead of a semiring S.

In this paper, we shall also consider formal power series over an arbitrary set U and Σ, defined by *partial* mappings from Σ^* to U. The support of such power series is its domain of definition,[a] and, similarly as for power series over semirings, the

[a] If U is a hemiring or a multi-hemiring (to be defined), we shall occasionally think of zero values as being undefined, and vice versa.

notation $U\langle\langle\Sigma^*\rangle\rangle$, $U\langle\Sigma^*\rangle$, $U\langle\langle\Sigma^+\rangle\rangle$, $U\langle L\rangle$, and $U\langle\langle L\rangle\rangle$ is used. Of course, one cannot directly define any reasonable operations on $U\langle\langle\Sigma^*\rangle\rangle$.

In order to study weighted finite automata and rational expressions over a semiring S, one has to assure that a star operation can be reasonably defined on $S\langle\langle\Sigma^*\rangle\rangle$. This can be done in a variety of ways (see, e.g., [6]). We shall use the approach taken by Droste and Meinecke [9] and Droste and Kuich [7]. That is, we shall consider *proper* power series over *arbitrary* semirings. For a proper power series r, one may define its star r^* by $(r^*, w) = \sum_{i=0}^{|w|}(r^i, w)$.

Let S be a semiring and Σ be an alphabet. A *proper finite automaton* over S and Σ is a quadruple $\mathcal{A} = (n, \alpha, A, \beta)$, where n in $\mathbb{N}$ is a *dimension* of $\mathcal{A}$, α in $(S\langle\Sigma\cup\{\varepsilon\}\rangle)^{1\times n}$ is a (row) *initial state vector*, A in $(S\langle\Sigma\rangle)^{n\times n}$ is a *transition matrix*, and β in $(S\langle\Sigma\cup\{\varepsilon\}\rangle)^{n\times 1}$ is a (column) *final state vector*. The *behaviour* of $\mathcal{A}$ is defined by $|\mathcal{A}| = \alpha A^*\beta$, via an extension of the star operation to square matrices of proper power series (see [6] for details).

Moreover, a power series r over S and Σ is *rational*, if it can be obtained from the elements of $S\langle\Sigma\cup\{\varepsilon\}\rangle$ by a finite number of applications of sum, product, and star, where the star is applied only to proper power series [6]. This process is syntactically represented by a *rational expression* over S and Σ. For the proof of the following theorem, see, e.g., [6].

Theorem 1. *Let S be a semiring and Σ be an alphabet. A series r in $S\langle\langle\Sigma^*\rangle\rangle$ is rational iff $r = |\mathcal{A}|$ for some proper finite automaton $\mathcal{A}$ over S and Σ.*

In the case of hemirings, one cannot deal with the empty word ε at all, so for a hemiring H, all considerations are made within the hemiring $H\langle\langle\Sigma^+\rangle\rangle$ with a plus operation defined by $(r^+, w) = \sum_{i=1}^{|w|}(r^i, w)$. Finite automata and rational series over hemirings have been introduced by Droste and Kuich [7]. We shall present their definitions restricted to the particular cases needed in this paper.

Let H be a hemiring and Σ be an alphabet. A *proper finite automaton* over H and Σ is a quadruple $\mathcal{A} = (n, \alpha, A, \beta)$, where n in $\mathbb{N}$ is a *dimension* of $\mathcal{A}$, α in $\mathbb{N}^{1\times n}$ is a (row) *initial state vector*, A in $(H\langle\Sigma\rangle)^{n\times n}$ is a *transition matrix*, and β in $\mathbb{N}^{n\times 1}$ is a (column) *final state vector*. The *behaviour* of $\mathcal{A}$ is defined by $|\mathcal{A}| = \alpha A^+\beta$, via an extension of the plus operation to square matrices of proper power series [7]. For n in $\mathbb{N}$ and a in H, na is defined by $na = \sum_{i=1}^{n} a$.

A power series r over H and Σ is *rational*, if it can be obtained from the elements of $S\langle\Sigma\rangle$ by a finite number of applications of sum, product, and plus. This process is syntactically represented by a *rational expression* over H and Σ. The following result has been obtained by Droste and Kuich [7].

Theorem 2 (Droste and Kuich [7]) *Let H be a hemiring and Σ be an alphabet. A series r in $H\langle\langle\Sigma^+\rangle\rangle$ is rational iff $r = |\mathcal{A}|$ for some proper finite automaton $\mathcal{A}$ over H and Σ.*

3. Automata with Auxiliary Weights

Let us now introduce our framework of automata with auxiliary weights. We shall present the fundamental definitions in this section. In the following two sections we shall first provide some examples of settings realizable in our framework (which have largely formed our motivation for introducing automata with auxiliary weights), and next we shall prove that automata with auxiliary weights extend automata over multi-hemirings.

Basically, a finite automaton with auxiliary weights is a finite automaton $\mathcal{A}$, in which transitions carry weights taken from a set W with no structure in general. When the behaviour of $\mathcal{A}$ is to be computed (or intuitively, when a run of $\mathcal{A}$ is to be executed), each transition of $\mathcal{A}$ is labelled by an additional auxiliary weight taken from some set X with no structure in general as well. However, it is assumed that the set $W \times X$, possibly extended by some set E, forms a semiring or a hemiring together with some operations $+$ and $\cdot$. Thus, after auxiliary weights are assigned, the automaton $\mathcal{A}$ becomes an automaton $\mathcal{A}'$ over $(W \times X) \cup E$, which is a semiring or a hemiring. The behaviour of the automaton $\mathcal{A}$ is then defined to be the "first projection" of the behaviour of $\mathcal{A}'$.

In this way, it is possible to compute quantities that cannot be realized by ordinary weighted automata over semirings and it is possible to generalize both semiring addition and semiring multiplication. See Section 4 for some examples and Section 5 for the proof that automata with auxiliary weights extend automata over multi-hemirings introduced by Droste and Kuich [7].

Definition 3. *A* hemiring auxiliary weighting structure (h-aws) *is an octuple* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$*, where* W*,* X*, and* E *are sets, sets* $W \times X$ *and* E *are disjoint,* $((W \times X) \cup E, +, \cdot, 0)$ *is a hemiring,* Σ *is an alphabet, and* $\upsilon \colon W \times \Sigma \to X$ *is a mapping.*

Definition 4. *A* semiring auxiliary weighting structure (s-aws) *is a nonuple* $(W, X, E, +, \cdot, 0, 1, \Sigma, \upsilon)$*, where* W*,* X*, and* E *are sets, sets* $W \times X$ *and* E *are disjoint,* $((W \times X) \cup E, +, \cdot, 0, 1)$ *is a semiring,* Σ *is an alphabet, and* $\upsilon \colon W \times \Sigma \to X$ *is a mapping.*

The reason for introducing the set E is that it often helps to establish hemiring or semiring structure – for a given set W, it is usually easier to find a hemiring or a semiring over $(W \times X) \cup E$ (for some sets X, E) than over $W \times X$ (for some X). The set E will typically contain only a few distinguished elements, such as the additive neutral element; it will mostly be impossible to obtain these elements as weights of runs in an automaton. The reader may consult Section 4 for particular examples.

Finite automata over h-aws and s-aws are simply finite automata with weights in the set W. The sets X and E are only used to define their behaviour and hence do not appear in Definition 5.

Definition 5. *A* proper finite automaton over an h-aws $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ *is a quadruple* $\mathcal{A} = (n, \alpha, A, \beta)$*, where* n *in* $\mathbb{N}$ *is a* dimension *of* $\mathcal{A}$*,* α *in* $\mathbb{B}^{1\times n}$ *is a (row) Boolean* initial state vector*,* A *in* $(W\langle\Sigma\rangle)^{n\times n}$ *is a* transition matrix*, and* β *in* $\mathbb{B}^{n\times 1}$ *is a (column) Boolean* final state vector. Proper finite automata over s-aws *are defined in the same way.*

By *automata with auxiliary weights*, we shall mean both proper finite automata over h-aws and proper finite automata over s-aws. If needed, we may refer specifically to automata with auxiliary weights over hemirings (automata over h-aws) or to automata with auxiliary weights over semirings (automata over s-aws).

Automata with auxiliary weights do not contain auxiliary weights "physically" – they may be viewed as automata weighted by elements of the set W together with the formula for assigning auxiliary weights to transitions – this formula is given by the function υ.

The behaviour of an automaton with auxiliary weights is defined by first assigning auxiliary weights via the mapping υ, next computing the behaviour of the resulting automaton, and finally taking the "first projection". This is expressed formally by the following definition.

Definition 6. *Let* $\mathcal{A} = (n, \alpha, A, \beta)$ *be a proper finite automaton over an h-aws* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$*. Let* $\mathcal{A}' = (n, \alpha', A', \beta')$ *be a proper finite automaton over the hemiring* $((W \times X) \cup E, +, \cdot, 0)$ *and* Σ*, where* α' *is* α *with all entries* $0_\mathbb{B}$ *replaced by* $0_\mathbb{N}$ *and all entries* $1_\mathbb{B}$ *replaced by* $1_\mathbb{N}$*,* A' *is* A *with all entries* $r = \sum_{c\in\mathrm{supp}(r)}(r, c)c$ *replaced by* $r' = \sum_{c\in\mathrm{supp}(r)}[(r, c), \upsilon((r, c), c)]c + \sum_{c\in\Sigma-\mathrm{supp}(r)} 0c$*, and* β' *is* β *with all entries* $0_\mathbb{B}$ *replaced by* $0_\mathbb{N}$ *and all entries* $1_\mathbb{B}$ *replaced by* $1_\mathbb{N}$*. Let* pr_1 *denote the first projection over* $W \times X$*. The behaviour of* $\mathcal{A}$ *is defined by*

$$|\mathcal{A}| = \sum_{\substack{w\in\Sigma^+ \\ (|\mathcal{A}'|,w)\notin E}} \mathrm{pr}_1(|\mathcal{A}'|, w)w.$$

The behaviour of an automaton over an s-aws is defined in the same way, except that the elements $0_\mathbb{B}$ *and* $1_\mathbb{B}$ *of* α *and* β *are not replaced by* $0_\mathbb{N}$ *and* $1_\mathbb{N}$*, but by* 0 *and* 1 *of the s-aws.*

Remark 7. *It might seem that the presented definitions are somewhat restrictive compared to the definitions of weighted automata over semirings or hemirings, as initial and final state vectors are Boolean. However, it is well known [6] that proper automata over semirings may be transformed into a quasi-normal form (there may be a difference in constant terms) with initial and final state vectors containing only* 0 *and* 1 *as entries. Similarly, it has been proved by Droste and Kuich [7] that automata over hemirings and multi-hemirings may be transformed into a normal form with initial and final state vectors containing only* $0_\mathbb{N}$ *and* $1_\mathbb{N}$ *as entries. (We shall call such automata* 0-1 automata.*) For this reason, our requirements do not pose any serious restriction.*

Finally, let us introduce *simple* h-aws and *simple* s-aws. Automata over such structures have the same auxiliary weight x assigned to all transitions.

Definition 8. *A* simple *h-aws is an h-aws* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$*, for which a constant* x *in* X *exists, such that for all* a *in* W *and* c *in* Σ*,* $\upsilon(a, c) = x$*.* Simple *s-aws are defined in the same way.*

Remark 9. *Before going any further, let us emphasize that automata with auxiliary weights do resemble ordinary weighted automata – using the above notation, over the hemiring (or semiring)* $(W \times X) \cup E$ *– with a slight difference in the definition of their behaviour. This means that automata with auxiliary weights retain many useful properties of ordinary weighted automata. In particular, practical use of automata with auxiliary weights does not pose any serious problem compared to ordinary weighted automata.*

4. Examples

In what follows, we shall give some examples in order to provide the reader with a better intuition about the possibilities of automata with auxiliary weights. We shall begin with a well known example, which can be handled by automata over multi-hemirings as well [7]. This may be viewed as a preparation for the following section, where we shall prove that automata with auxiliary weights are stronger than automata over multi-hemirings. Examples 11 to 17, on the contrary, show that automata with auxiliary weights can be used to capture settings involving meaningful generalizations of semiring addition – as we shall see in the following section, these are not realizable over multi-hemirings.

Example 10. *Let us suppose that we are interested, for each run in an automaton, in* an average weight of a transition in the run*, and for all words* w*, in* a maximal weight of a run *on* w*. This is a well known example arising from practice, and can be realized, e.g., over multi-hemirings [7] – it is not directly realizable over classical semirings or hemirings. In what follows, we shall show that automata with auxiliary weights can be used for this purpose as well.*

This can be realized over an h-aws $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ *with* $W = \mathbb{R} \cup \{-\infty\}$*,* $X = \mathbb{N}_+$ *and* $E = \{\bot, \top\}$*, where* $+$ *is an operation on* $(W \times X) \cup E$ *defined by*

$$\begin{aligned}
(a, k) + (b, k) &= (\sup\{a, b\}, k), && a, b \in W, k \in X, \\
(a, k) + (b, m) &= \top, && a, b \in W, k, m \in X, k \neq m, \\
(a, k) + \top = \top + (a, k) &= \top, && a \in W, k \in X, \\
\top + \top &= \top, && \\
\bot + \zeta = \zeta + \bot &= \zeta, && \zeta \in (W \times X) \cup E,
\end{aligned}$$

$\cdot$ *is an operation on* $(W \times X) \cup E$ *defined by*

$$\begin{aligned}
(a,k) \cdot (b,m) &= \left(\frac{ka+mb}{k+m}, k+m\right), && a,b \in W, k,m \in X,\\
(a,k) \cdot \top = \top \cdot (a,k) &= \top, && a \in W, k \in X,\\
\top \cdot \top &= \top, &&\\
\bot \cdot \zeta = \zeta \cdot \bot &= \bot, && \zeta \in (W \times X) \cup E,
\end{aligned}$$

$0 = \bot$, *and* $\upsilon\colon W \times \Sigma \to X$ *is defined by* $\upsilon(a,c) = 1$ *for all* a *in* W *and* c *in* Σ. *It can be easily verified that* $((W \times X) \cup E, +, \cdot, 0)$ *is a hemiring, and therefore* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ *is an h-aws.*

Intuitively, the first component of an element of $W \times X$ *is the actual maximal average weight and the second component is the length of a run attaining this maximal average value. Since we consider proper automata only, all paths on the same word* w *have the same length* $|w|$. *Thus, it suffices to define addition by assigning non-*$\top$ *values only for pairs with the same values of the second component.*

For more information, one may consult, e.g., [7] – as the above construction is merely an adaptation of the construction for multi-hemirings to our framework. As we shall observe, the same adaptation is possible for each multi-hemiring (see Theorem 20).

Example 11. *Let us now show that the framework of automata with auxiliary weights can handle a setting, in which the weight of a single run is computed by* summing the transition weights in the run*, and where the resulting weight of a word* w *is* the average weight of all runs *on* w.

This can be realized over an h-aws $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$, *where* $W = \mathbb{R}$, $X = \mathbb{N}_+$, $E = \{\bot\}$, $+$ *is an operation on* $(W \times X) \cup E$ *defined by*

$$\begin{aligned}
(a,k) + (b,m) &= \left(\frac{ka+mb}{k+m}, k+m\right), && a,b \in W, k,m \in X,\\
\zeta + \bot = \bot + \zeta &= \zeta, && \zeta \in (W \times X) \cup E,
\end{aligned}$$

$\cdot$ *is an operation on* $(W \times X) \cup E$ *defined by*

$$\begin{aligned}
(a,k) \cdot (b,m) &= (a+b, km), && a,b \in W, k,m \in X,\\
\zeta \cdot \bot = \bot \cdot \zeta &= \bot, && \zeta \in (W \times X) \cup E,
\end{aligned}$$

$0 = \bot$, *and* $\upsilon\colon W \times \Sigma \to X$ *is defined by* $\upsilon(a,c) = 1$ *for all* a *in* W *and* c *in* Σ. *It can be easily verified that* $((W \times X) \cup E, +, \cdot, 0)$ *is a hemiring, and therefore* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ *is an h-aws.*

Intuitively, the first component of elements of $W \times X$ *is the average weight of runs and the second component is the number of these runs. Indeed, if* a *is the* *average of k weights $a_1, \ldots, a_k$ and b is the average of m weights $b_1, \ldots, b_m$, then*

$$(a,k) + (b,m) = \left(\frac{a_1 + \ldots + a_k}{k}, k\right) + \left(\frac{b_1 + \ldots + b_m}{m}, m\right) = \\ = \left(\frac{a_1 + \ldots + a_k + b_1 + \ldots + b_m}{k+m}, k+m\right),$$

i.e., the average of $k+m$ weights $a_1, \ldots, a_k, b_1, \ldots, b_m$. Similarly,

$$(a,k) \cdot (b,m) = \left(\frac{a_1 + \ldots + a_k}{k}, k\right) + \left(\frac{b_1 + \ldots + b_m}{m}, m\right) = \\ = \left(\frac{ma_1 + \ldots ma_k + kb_1 + \ldots kb_m}{km}, km\right) = \\ = \left(\frac{\sum_{i=1}^{k} \sum_{j=1}^{m} (a_i + b_j)}{km}, km\right),$$

i.e., the average of km weights $a_i + b_j$ for $i = 1, \ldots, k$ and $j = 1, \ldots, m$.

Example 12. *Let $\mathcal{A} = (2, \alpha, A, \beta)$ be a proper finite automaton over the h-aws of Example 11 (with $\Sigma = \{a, b\}$), given as follows:*

$$\alpha = (1, 0), \qquad A = \begin{pmatrix} 2a & 3a + 3b \\ 4b & 3b \end{pmatrix}, \qquad \beta = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Let us now compute, using matrix multiplication, the weight $(|\mathcal{A}|, ab)$ of the word ab in the behaviour of $\mathcal{A}$.

According to Definition 6, the behaviour of $\mathcal{A}$ is defined via the behaviour of $\mathcal{A}' = (2, \alpha', A', \beta')$, which is an automaton over the hemiring $((W \times X) \cup E, +, \cdot, 0)$ given by

$$\alpha' = (1, 0), \qquad A' = \begin{pmatrix} [2,1]a & [3,1]a + [3,1]b \\ [4,1]b & [3,1]b \end{pmatrix}, \qquad \beta' = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Now,

$$(A')^2 = \begin{pmatrix} [4,1]aa + [7,1]ab + [7,1]bb & [5,1]aa + [\frac{11}{2}, 2]ab + [6,1]bb \\ [6,1]ba + [7,1]bb & [7,1]ba + [\frac{13}{2}, 2]bb \end{pmatrix},$$

so that

$$\alpha' \cdot (A')^2 \cdot \beta' = [9/2, 2]aa + [6,3]ab + [13/2, 2]bb$$

and

$$(|\mathcal{A}'|, ab) = [6, 3].$$

As $[6,3] \notin E$, we obtain

$$(|\mathcal{A}|, ab) = \mathrm{pr}_1(|\mathcal{A}'|, ab) = 6.$$

That is, the average weight of an accepting run on the word ab is 6. This result can be easily verified by inspection – there are three accepting runs on the word ab in $\mathcal{A}$, with weights 5, 6, and 7.

Example 13. *A setting similar to the one of Example 11, in which the weight of a single run is computed by* multiplying transition weights in the run*, and where the resulting weight of a word* w *is* the average weight of particular runs *on* w*, can be realized over an h-aws* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$*, which is defined in the same way as in Example 11, except that* $(a, k) \cdot (b, m) = (ab, km)$ *for* a, b *in* W *and* k, m *in* X*.*

One might think of several additional settings, which are realizable in the framework of automata with auxiliary weights, and which cannot be directly modelled using automata over semirings or hemirings. Example 14 shows that automata with auxiliary weights can handle a setting, in which the weight of a run is computed by the usual addition on real numbers, and the weight of a word w is the sum of the weights of runs on w.

Example 14. *As a slight modification of Example 11, let us consider a setting, in which the weight of a run is computed by* summing the transition weights in the run*, and the weight of a word* w *is* the sum of the weights of all runs *on* w*.*

This is handled by an h-aws $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$*, where* $W = \mathbb{R}$*,* $X = \mathbb{N}_+$*,* $E = \{\perp\}$*,* $+$ *is an operation on* $(W \times X) \cup E$ *defined by*

$$(a, k) + (b, m) = (a + b, k + m)\,, \qquad a, b \in W, k, m \in X,$$
$$\zeta + \perp = \perp + \zeta = \zeta, \qquad \zeta \in (W \times X) \cup E,$$

$\cdot$ *is an operation on* $(W \times X) \cup E$ *defined by*

$$(a, k) \cdot (b, m) = (ma + kb, km)\,, \qquad a, b \in W, k, m \in X,$$
$$\zeta \cdot \perp = \perp \cdot \zeta = \perp, \qquad \zeta \in (W \times X) \cup E,$$

$0 = \perp$*, and* $\upsilon\colon W \times \Sigma \to X$ *is defined by* $\upsilon(a, c) = 1$ *for all* a *in* W *and* c *in* Σ*.*

The intuition behind this construction is similar to that in Example 11 – auxiliary weights correspond to the number of runs, from the weights of which the sum is computed.

When the weight of a run is computed by the usual addition, one may utilize automata with auxiliary weights in order to compute the difference between the most expensive and the cheapest run on a given word, together with some other quantity, such as the average. We elaborate on this idea in the following example. It is impossible to compute *only* the above mentioned difference, as our requirement (R4) would be violated in this case (see also Remark 16).

Example 15. *Let us show that automata with auxiliary weights can handle a setting, in which the weight of a run is computed by* summing the transition weights in the run *representing their costs, and where we are interested in the difference between the weight of the most expensive and the cheapest run on a given word, together with, e.g., the average of Example 11. Our addition thus corresponds to computing* the average and the diameter of a set*.*

This cannot be modelled if only the costs of transitions are given initially. In this case, one of our underlying requirements explained above – (R4) *– would be* *violated, as in the automaton consisting of a single transition on some letter c, the weight of this transition would in general be different from the resulting weight of the word c.*

For this reason, our construction will work only if the weights of transitions are given in the form $[a, 0]$, where a is the weight itself and 0 stands for "diameter zero". We shall compute the average in the first component and the diameter in the second component.

Our h-aws is given as a tuple $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$, where $W = \mathbb{R} \times \mathbb{R}$, $X = \mathbb{R} \times \mathbb{R} \times \mathbb{N}$, $E = \{\bot\}$, $+$ is an operation on $(W \times X) \cup E$ defined by

$$([a, d], [m, M, n]) + ([a', d'], [m', M', n']) =$$
$$= \left(\left[\frac{na + n'a'}{n' + n'}, \max(M, M') - \min(m, m') \right], [\min(m, m'), \max(M, M'), n + n'] \right)$$

for $a, d, m, M, a', d', m', M'$ in $\mathbb{R}$ and $n, n' \in \mathbb{N}$ and by

$$\zeta + \bot = \bot + \zeta = \zeta$$

for $\zeta \in (W \times X) \cup E$, $\cdot$ is an operation on $(W \times X) \cup E$ defined by

$$([a, d], [m, M, n]) \cdot ([a', d'], [m', M', n']) =$$
$$= ([a + a', M + M' - m - m'], [m + m', M + M', nn'])$$

for $a, d, m, M, a', d', m', M'$ in $\mathbb{R}$ and $n, n' \in \mathbb{N}$ and by

$$\zeta \cdot \bot = \bot \cdot \zeta = \bot$$

for $\zeta \in (W \times X) \cup E$, $0 = \bot$, and $\upsilon \colon W \times \Sigma \to X$ is defined by $\upsilon([a, 0], c) = [a, a]$ for all a in $\mathbb{R}$ and c in Σ. As we assume that the weights of the transitions are of the form $[a, 0]$ the function υ may be defined arbitrarily on the rest of its inputs.

The idea behind this construction is that the auxiliary weights $[m, M, n]$ of X contain the weight m of the cheapest run, the weight M of the most expensive run, and the number n of runs used to compute the average. The diameter d is simply the difference $M - m$.

Remark 16. *It is impossible to utilize automata with auxiliary weights to compute the diameter of Example 15 without some additional quantity. The reason is that this would violate our requirement* (R4)*, to which automata with auxiliary weights comply.*

One could consider a generalization of automata with auxiliary weights, in which several auxiliary weights taken from some sets $X_1, \ldots, X_k$ are assigned, and in which the resulting weight may be a projection either on W, or on any of the k auxiliary components. In this model, it would be possible to compute only the diameter. However, we find the original model compliant to (R4) *preferable.*

Automata with auxiliary weights may be used to compute some more complex quantitative properties as well. For instance, in the following Example 17 we describe a setting, in which each transition is assigned a nonnegative real cost, while each transition labelled by some particular symbol is assigned a penalty, which is directly proportional to its cost. We show that automata with auxiliary weights may be used to compute the cost of a run, which minimizes the penalty (in case there is more than one such run, then the cost is minimized as well).

Example 17. *As a final example in this section, let us consider a setting in which each transition is assigned a cost – this is represented by some nonnegative real number. In addition, executing a transition labelled by some particular symbol c yields a penalty directly proportional to the cost of the transition by some nonnegative real constant λ.*

The overall aim is to minimize the penalty. However, what interests us about the optimal run is its cost. That is, the weight of a run should be the sum of transition weights in the run*. The weight of a word w should be* the weight of a run *on w* minimizing the penalty*. If there are two or more runs on the same word with the same penalty, one with the minimal cost is chosen.*

Let us show that this setting can be modelled using automata with auxiliary weights. Let λ be in $\mathbb{R}_{\geq 0}$. We shall work over an h-aws given as a tuple $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$, where Σ contains a special symbol c, $W = \mathbb{R}_{\geq 0}$, $X = \mathbb{R}_{\geq 0}$, $E = \{\perp\}$, $+$ is an operation on $(W \times X) \cup E$ defined by

$$(a, a') + (b, b') = \begin{cases} (a, a') & \text{if } a' < b' \\ (\min(a, b), a') & \text{if } a' = b' \\ (b, b') & \text{if } a' > b' \end{cases}, \qquad a, b \in W, a', b' \in X,$$

$$\zeta + \perp = \perp + \zeta = \zeta, \qquad \zeta \in (W \times X) \cup E,$$

the operation $\cdot$ on $(W \times X) \cup E$ is defined by

$$(a, a') \cdot (b, b') = (a + b, a' + b'), \qquad a, b \in W, a', b' \in X,$$

$$\zeta \cdot \perp = \perp \cdot \zeta = \perp, \qquad \zeta \in (W \times X) \cup E,$$

$0 = \perp$, and $\upsilon \colon W \times \Sigma \to X$ is defined by $\upsilon(a, c) = \lambda a$ for all a in W and by $\upsilon(a, d) = 0$ for all a in W and d in $\Sigma - \{c\}$.

The idea behind this construction is that the auxiliary weight holds the overall penalty – that is λ times the sum of weights corresponding to transitions labelled by the letter c.

In addition to the examples described above, there are numerous other settings, which appear to be realizable in the framework of automata with auxiliary weights. For instance, the semiring addition may be replaced by taking a geometric mean, a quadratic mean, or the sum of squares. The reader is encouraged to elaborate the details of such constructions.

5. A Relation to Automata over Multi-Hemirings

In this section, we shall prove that automata with auxiliary weights extend automata over multi-hemirings of Droste and Kuich [7] – that is, each automaton over a multi-hemiring can be viewed as an automaton with auxiliary weights. On the other hand, we shall observe that the examples listed in Section 4 (except Example 10) cannot be realized by automata over multi-hemirings. In particular, this is because automata with auxiliary weights can be used to generalize both semiring (or hemiring) addition and semiring (or hemiring) multiplication, while automata over multi-hemirings can generalize only semiring (or hemiring) multiplication. In fact, the definition of automata with auxiliary weights can be seen as an attempt to extend the framework provided by automata over multi-hemirings in order to incorporate meaningful generalizations of semiring (or hemiring) addition.

Let us first provide some definitions and results of Droste and Kuich [7] on automata over *multi-hemirings*. The idea behind this model is to take instead of semiring multiplication a family of products parameterized by pairs of natural numbers, satisfying some simple axioms. When multiplication is applied to *power series*, the coefficients are multiplied by a product indexed by lengths of respective words. As a result, multiplication of power series is parameterized by word length, and is thus more powerful than usual multiplication in semirings.

Definition 18. *A* multi-hemiring *is a quadruple* $(D, +, \circ, 0)$*, where* $(D, +, 0)$ *is a commutative monoid and where* $\circ = (\cdot_{m,n} \mid m, n \in \mathbb{N}_+)$ *is a family of products* $\cdot_{m,n} \colon D \times D \to D$*, such that (for all* a, b, c *in* D *and all* k, m, n *in* $\mathbb{N}_+$*):*

(*i*) $0 \cdot_{m,n} a = a \cdot_{m,n} 0 = 0$,
(*ii*) $(a \cdot_{k,m} b) \cdot_{k+m,n} c = a \cdot_{k,m+n} (b \cdot_{m,n} c)$,
(*iii*) $a \cdot_{m,n} (b + c) = a \cdot_{m,n} b + a \cdot_{m,n} c$ *and* $(a + b) \cdot_{m,n} c = a \cdot_{m,n} c + b \cdot_{m,n} c$.

Let D be a multi-hemiring and Σ be an alphabet. *Formal power series* over D and Σ and their addition are defined in the same way as for semirings and hemirings. Multiplication is defined for *proper* power series as follows: for all proper power series r_1 and r_2 over D and Σ and for all w in Σ^+,

$$(r_1 \cdot r_2, w) = \sum_{\substack{uv=w \\ u,v\neq\varepsilon}} (r_1, u) \cdot_{|u|,|v|} (r_2, v).$$

The plus operation is defined in the same way as for hemirings, and the collection of all proper power series over D and Σ is denoted by $D\langle\langle\Sigma^+\rangle\rangle$. The set $D\langle\langle\Sigma^+\rangle\rangle$, together with addition, multiplication and plus forms a Conway hemiring [7].[b]

Finite automata over D and Σ and *rational* power series are defined in the same way as for hemirings, and Droste and Kuich [7] have proved the following theorem:

Theorem 19 (Droste and Kuich [7]) *Let* D *be a multi-hemiring and* Σ *be an alphabet. A series* r *in* $D\langle\langle\Sigma^+\rangle\rangle$ *is rational iff* $r = |\mathcal{A}|$ *for some proper finite automaton* $\mathcal{A}$ *over* D *and* Σ.

[b]The notion of a Conway hemiring has been introduced by Droste and Kuich [7] and is a hemiring counterpart of the more classical notion of a Conway semiring [6].

Even before automata over multi-hemirings have been introduced, Droste and Meinecke have studied automata over valuation monoids [8–10, 14], an alternative approach to generalizations of weighted automata. In [7], Droste and Kuich have proved that automata over multi-hemirings are equivalent to automata over so-called regular Cauchy valuation monoids, which still encompass all reasonable cases of valuation monoids arising from practice.

Now, let us relate automata over multi-hemirings to automata with auxiliary weights. We shall first prove that proper finite automata over multi-hemirings can be seen as proper finite automata over simple h-aws. Next, we shall show that these automata can be further seen as proper finite automata over simple s-aws. Finally, we shall conclude that automata with auxiliary weights form a stronger framework than automata over multi-hemirings.

Let us first describe the transition from multi-hemirings to simple h-aws. In a multi-hemiring, the product depends on the lengths of a particular pair of words. The main idea in what follows is to make these word lengths our auxiliary weights, which essentially transforms a multi-hemiring into a hemiring. The formal construction is as follows:

Theorem 20. *Let* $(D, +, \circ, 0)$ *be a multi-hemiring, let* Σ *be an alphabet, and let* $\mathcal{A} = (n, \alpha, A, \beta)$ *be a proper finite 0-1 automaton over* D *and* Σ*. Then* $\mathcal{A}$ *can be viewed as a proper finite automaton* $\mathcal{B}$ *over a simple h-aws* $(D, X, E, \oplus, \otimes, \mathbf{0}, \mathbf{1}, \Sigma, \upsilon)$*, such that* $|\mathcal{B}| = |\mathcal{A}|$*.*

Proof. Let $X = \mathbb{N}_+$ and $E = \{\bot, \top\}$. Let $\oplus$ be an operation on $(D \times X) \cup E$ defined by

$$\begin{aligned} (a,k) \oplus (b,k) &= (a+b,k), && a,b \in D, k \in X, \\ (a,k) \oplus (b,m) &= \top, && a,b \in D, k,m \in X, k \neq m, \\ (a,k) \oplus \top = \top \oplus (a,k) &= \top, && a \in D, k \in X, \\ \top \oplus \top &= \top, && \\ \bot \oplus \zeta = \zeta \oplus \bot &= \zeta, && \zeta \in (D \times X) \cup E. \end{aligned}$$

Furthermore, let $\otimes$ be an operation on $(D \times X) \cup E$ defined by

$$\begin{aligned} (a,k) \otimes (b,m) &= (a \cdot_{k,m} b, k+m), && a,b \in D, k,m \in X, \\ (a,k) \otimes \top = \top \otimes (a,k) &= \top, && a \in D, k \in X, \\ \top \otimes \top &= \top, && \\ \bot \otimes \zeta = \zeta \otimes \bot &= \bot, && \zeta \in (D \times X) \cup E. \end{aligned}$$

One could easily convince himself that the set $(D \times X) \cup E$, equipped with the operations $\oplus$ and $\otimes$, forms a hemiring with an additive identity given by $\mathbf{0} = \bot$. Thus, if $\upsilon\colon D \times \Sigma \to X$ is defined by $\upsilon(a, c) = 1$ for all a in D and c in Σ, then $(D, X, E, \oplus, \otimes, \mathbf{0}, \Sigma, \upsilon)$ is a simple h-aws.

Let $\mathcal{B}$ be $\mathcal{A}$ seen as an automaton over this h-aws (this means that we naturally identify $0_{\mathbb{N}}$ with $0_{\mathbb{B}}$ and $1_{\mathbb{N}}$ with $1_{\mathbb{B}}$; furthermore, as noted in Section 2, we shall think of zero weights of $\mathcal{A}$ as being undefined in $\mathcal{B}$, as D is no longer seen as a multi-hemiring, but merely as a set). By Definition 6, $|\mathcal{B}|$ is the "first projection" of $|\mathcal{B}'|$, where $\mathcal{B}'$ is an automaton over the hemiring $(D \times X) \cup E$ with a transition matrix A', which is A with each nonzero weight a in D replaced by $(a, 1)$ in $D \times X$ (and with zero weights canonically replaced by $\perp$).

One could prove by induction that for all i in $\mathbb{N}_+$, the i-th power of A' is exactly the i-th power of A, with each nonzero weight a in D replaced by (a, i) in $W \times X$ (and with zero weights replaced by $\perp$). Thus, if

$$|\mathcal{A}| = \alpha A^+ \beta = \sum_{w \in \operatorname{supp}|\mathcal{A}|} (|\mathcal{A}|, w) w + \sum_{w \notin \operatorname{supp}|\mathcal{A}|} 0w,$$

then

$$|\mathcal{B}'| = \alpha (A')^+ \beta = \sum_{w \in \operatorname{supp}|\mathcal{A}|} ((|\mathcal{A}|, w), |w|) w + \sum_{w \notin \operatorname{supp}|\mathcal{A}|} \perp w,$$

and by taking the "first projection", the equality $|\mathcal{B}| = |\mathcal{A}|$ follows (after identifying undefined values of $|\mathcal{B}|$ with zero values of $|\mathcal{A}|$). The theorem is proved. □

Remark 21. *The formal meaning of "can be viewed" in the statement of Theorem 20 has already been clarified in its proof: the automaton $\mathcal{B}$ in fact* is *the automaton $\mathcal{A}$ interpreted as an automaton over an h-aws. This involves, as a minor formal detail, switching from natural numbers* 0 *and* 1 *to respective Boolean values, and from zero weights in D to undefined values. Similarly, the equivalence of automata $\mathcal{A}$ and $\mathcal{B}$ holds only after identifying undefined values of $|\mathcal{B}|$ with zero weights of $|\mathcal{A}|$.*

Let H be a hemiring. The *Dorroh extension* of H by $\mathbb{N}$ (see, e.g., Golan [13]) is a semiring $H \times \mathbb{N}$ with addition defined by $(a, k) + (b, m) = (a + b, k + m)$ and with multiplication defined by $(a, k) \cdot (b, m) = (ma + kb + ab, km)$. This construction has been applied extensively by Droste and Kuich while developing the theory of weighted automata over hemirings [7], and will also be crucial for the proof of the following theorem.

In what follows, we shall say that a proper finite automaton $\mathcal{A} = (n, \alpha, A, \beta)$ over an h-aws is *normal*, if for $\alpha = (\alpha_1, \ldots, \alpha_n)$ and for $\beta = (\beta_1, \ldots, \beta_n)^T$, the identity $\alpha_i \beta_i = 0$ holds for all i. In other words, $\mathcal{A}$ is normal iff it has no state, which is at the same time initial and final. It is a trivial exercise to prove that this indeed is a normal form.

Theorem 22. *Let $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ be an h-aws (a simple h-aws), and let $\mathcal{A} = (n, \alpha, A, \beta)$ be a normal proper finite automaton over the h-aws (the simple h-aws) $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$. Then $\mathcal{A}$ can be viewed as a proper finite automaton $\mathcal{B}$ over an s-aws (a simple s-aws) $(W, Y, F, \oplus, \otimes, \mathbf{0}, \mathbf{1}, \Sigma, \tau)$, such that $|\mathcal{B}| = |\mathcal{A}|$.*

Proof. As $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ is an h-aws, $((W \times X) \cup E, +, \cdot, 0)$ is a hemiring. By the Dorroh extension, we obtain a semiring $((W \times X \times \mathbb{N}) \cup (E \times \mathbb{N}), \oplus, \otimes, \mathbf{0}, \mathbf{1})$, where, for all ζ, η in $(W \times X) \cup E$ and k, m in $\mathbb{N}$, $\oplus$ is given by

$$(\zeta, k) \oplus (\eta, m) = (\zeta + \eta, k + m),$$

$\otimes$ is given by

$$(\zeta, k) \otimes (\eta, m) = (m\zeta + k\eta + \zeta \cdot \eta, km),$$

$\mathbf{0} = (0, 0_{\mathbb{N}})$, and $\mathbf{1} = (0, 1_{\mathbb{N}})$. Set $Y = X \times \mathbb{N}$ and $F = E \times \mathbb{N}$, and define $\tau \colon W \times \Sigma \to Y$ by $\tau(a, c) = (\upsilon(a, c), 0_{\mathbb{N}})$ for all a in W and c in Σ. Then $(W, Y, F, \oplus, \otimes, \mathbf{0}, \mathbf{1}, \Sigma, \tau)$ is an s-aws. Let $\mathcal{B}$ be $\mathcal{A}$ seen as an automaton over $(W, Y, F, \oplus, \otimes, \mathbf{0}, \mathbf{1}, \Sigma, \tau)$. The behaviour $|\mathcal{A}|$ is the "first projection" of $|\mathcal{A}'|$ and $|\mathcal{B}|$ is the "first projection" of $|\mathcal{B}'|$, where $\mathcal{A}'$ and $\mathcal{B}'$ are given by Definition 6.

Let $\mathcal{A}'$ and $\mathcal{B}'$ have transition matrices A' and B', respectively. One could verify by induction that for all i in $\mathbb{N}_+$, the i-th power of B' is the i-th power of A' with each weight ζ in $(W \times X) \cup E$ replaced by $(\zeta, 0_{\mathbb{N}})$ in $(W \times X \times \mathbb{N}) \cup (E \times \mathbb{N})$. By normality of $\mathcal{A}$, one need not consider the 0-th power of B' at all. Thus, by taking the "first projection", one finally obtains $|\mathcal{B}| = |\mathcal{A}|$.

Now, if the h-aws $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ is simple, then it follows by Definition 8 that a constant x in X exists, such that for $\upsilon(a, c) = x$ for all a in W and c in Σ. Then, for all a in W and c in Σ, $\tau(a, c) = (\upsilon(a, c), 0_{\mathbb{N}}) = (x, 0_{\mathbb{N}})$, and the s-aws $(W, Y, F, \oplus, \otimes, \mathbf{0}, \mathbf{1}, \Sigma, \tau)$ is simple as well. □

As observed by Droste and Kuich [7], each proper finite automaton over a multi-hemiring D can be transformed into an equivalent normal proper finite 0-1 automaton over D, where "normality" is defined in the same way as for automata over h-aws. Thus the following corollary might be of some interest.

Corollary 23. *Let* $(D, +, \circ, 0)$ *be a multi-hemiring, let* Σ *be an alphabet, and let* $\mathcal{A} = (n, \alpha, A, \beta)$ *be a normal proper finite 0-1 automaton over* D *and* Σ. *Then* $\mathcal{A}$ *can be viewed as a proper finite automaton* $\mathcal{B}$ *over a simple s-aws* $(D, X, E, \oplus, \otimes, \mathbf{0}, \mathbf{1}, \Sigma, \upsilon)$, *such that* $|\mathcal{B}| = |\mathcal{A}|$.

Proof. Follows directly by Theorem 20, Theorem 22, and by an obvious fact that the transformation of Theorem 20 preserves normality. □

Finally, one might easily prove that the examples mentioned in Section 4 (except Example 10) cannot be realized over multi-hemirings. For instance, in the setting of Example 11, one might argue that the average operation is not associative, and therefore cannot be used as addition in a multi-hemiring. Given this, automata with auxiliary weights not only extend automata over multi-hemirings, but this extension is also non-trivial. Let us emphasize once again that automata with auxiliary weights may be viewed as an attempt to incorporate generalizations of semiring addition into the framework of automata over multi-hemirings.

6. Rational Expressions with Auxiliary Weights

Kleene-like theorem proving equivalence of automata and suitably modified weighted rational expressions [15] has been proved both for automata over multi-hemirings of Droste and Kuich [7] and for automata over valuation monoids of Droste and Meinecke [9].

In what follows, we shall prove a similar result for automata with auxiliary weights. We shall do so for automata with auxiliary weights over hemirings only, but the same result holds for automata with auxiliary weights over semirings as well.

We shall introduce rational expressions with auxiliary weights (over an h-aws), a generalization of weighted rational expressions, and prove that they are equivalent to finite automata with auxiliary weights (over an h-aws). The conceptual simplicity of our definitions makes this theorem almost a direct consequence of results already known.

Definition 24. *Let* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ *be an h-aws. A* nonempty rational expression *over* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ *is:*

(*i*) *Either one of the atomic formulae ac for a in W and c in* Σ*.*
(*ii*) *Or one of the formulae* $(\mathcal{E} + \mathcal{F})$*,* $(\mathcal{E} \cdot \mathcal{F})$*, and* $(\mathcal{E}^+)$*, where* $\mathcal{E}$ *and* $\mathcal{F}$ *are nonempty rational expressions.*

Definition 25. *A* rational expression *over an h-aws* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ *is either a nonempty rational expression over* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$*, or the empty formula.*

The empty rational expression denotes the power series with all values undefined. Power series denoted by nonempty expressions are defined as follows.

Definition 26. *Let* $\mathcal{E}$ *be a rational expression over an h-aws* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$*, which is given by a nonempty formula. Let* $\mathcal{E}'$ *be a rational expression over the hemiring* $((W \times X) \cup E, +, \cdot, 0)$*, which is obtained from* $\mathcal{E}$ *by replacing atomic expressions ac by* $[a, \upsilon(a, c)]c$*. The* power series denoted by $\mathcal{E}$ *is given by*

$$|\mathcal{E}| = \sum_{\substack{w \in \Sigma^+ \\ (|\mathcal{E}'|, w) \notin E}} \mathrm{pr}_1(|\mathcal{E}'|, w) w.$$

Theorem 27. *Let* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ *be an h-aws and r be in* $W\langle\langle \Sigma^+ \rangle\rangle$*. The series r is a behaviour of a proper finite automaton over* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$ *iff r is denoted by a rational expression over* $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$*.*

Proof. First let $r = |\mathcal{A}|$ for some proper finite automaton $\mathcal{A} = (n, \alpha, A, \beta)$ over the h-aws $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$. Then $|\mathcal{A}|$ is defined in terms of $|\mathcal{A}'|$, where $\mathcal{A}'$ is an automaton over the hemiring $((W \times X) \cup E, +, \cdot, 0)$, such that its transition matrix A' has atoms ac of A replaced by $[a, \upsilon(a, c)]c$ (see Definition 6). One can clearly construct a rational expression $\mathcal{E}'$ over the hemiring $((W \times X) \cup E, +, \cdot, 0)$, such that $|\mathcal{E}'| = |\mathcal{A}'|$ and the set of atoms in $\mathcal{E}'$ is the set of atoms in A'. In particular, for all atoms xc of $\mathcal{E}'$, $x = [a, \upsilon(a, c)]$ for some a in W. As a result, $\mathcal{A}$ is equivalent to some rational expression $\mathcal{E}$ over the h-aws $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$.

Conversely, let $r = |\mathcal{E}|$ for some rational expression $\mathcal{E}$ over $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$. The claim is trivial for $\mathcal{E}$ empty. If $\mathcal{E}$ is nonempty, then $|\mathcal{E}|$ is defined in terms of $|\mathcal{E}'|$, where $\mathcal{E}'$ is a rational expression over the hemiring $((W \times X) \cup E, +, \cdot, 0)$ with atoms ac of $\mathcal{E}$ replaced by $[a, \upsilon(a, c)]c$. A 0-1-automaton over the hemiring $((W \times X) \cup E, +, \cdot, 0)$, equivalent to $\mathcal{E}'$, can be constructed by structural induction [7]. However, the construction of [7] might not work in our case, as it involves – in the definition of automata for product and Kleene plus – addition of atoms in the transition matrix of the automaton. This may spoil the desired property that for all atoms xc in the transition matrix, $x = [a, \upsilon(a, c)]$ for some a in W.

In what follows, we shall show that this property can be achieved if the construction of [7] is carefully modified, and that for each rational expression over the hemiring $((W \times X) \cup E, +, \cdot, 0)$, it is possible to construct a 0-1-automaton over $((W \times X) \cup E, +, \cdot, 0)$ with the same set of atomic expressions.

Clearly, for each atomic expression ac (with a in $(W \times X) \cup E$ and with c in Σ), a 0-1-automaton $\mathcal{B}_{ac} = (n_{ac}, \alpha_{ac}, B_{ac}, \beta_{ac})$ can be constructed, such that $|\mathcal{B}_{ac}| = ac$. It suffices to take $n_{ac} = 2$ and

$$\alpha_{ac} = (1, 0), \qquad B_{ac} = \begin{pmatrix} 0 & ac \\ 0 & 0 \end{pmatrix}, \qquad \beta_{ac} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Now, let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two expressions over the hemiring $((W \times X) \cup E, +, \cdot, 0)$. By the induction hypothesis, we may suppose that $|\mathcal{F}_1| = |\mathcal{B}_1|$ and $|\mathcal{F}_2| = |\mathcal{B}_2|$ for some 0-1-automata $\mathcal{B}_1 = (n_1, \alpha_1, B_1, \beta_1)$ and $\mathcal{B}_2 = (n_2, \alpha_2, B_2, \beta_2)$, such that the set of atomic expressions in B_1 (resp. B_2) is the same as the set of atomic expressions in $\mathcal{F}_1$ (resp. $\mathcal{F}_2$). Now, the expression $\mathcal{F}_1 + \mathcal{F}_2$ is clearly equivalent to the automaton $\mathcal{B}_1 + \mathcal{B}_2 = (n_1 + n_2, \alpha_{1,+,2}, B_{1,+,2}, \beta_{1,+,2})$, such that

$$\alpha_{1,+,2} = (\alpha_1, \alpha_2), \qquad B_{1,+,2} = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}, \qquad \beta_{1,+,2} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}.$$

Furthermore, as might easily be observed, the expression $\mathcal{F}_1 \cdot \mathcal{F}_2$ is equivalent to the automaton $\mathcal{B}_1 \cdot \mathcal{B}_2 = (n_1 + n_1 \cdot n_2 + n_2, \alpha_{1,\cdot,2}, B_{1,\cdot,2}, \beta_{1,\cdot,2})$, such that

$$\alpha_{1,\cdot,2} = (\alpha_1, 0, 0), \qquad B_{1,\cdot,2} = \begin{pmatrix} B_1 & C_1 & 0 \\ 0 & 0 & C_2 \\ 0 & 0 & B_2 \end{pmatrix}, \qquad \beta_{1,\cdot,2} = \begin{pmatrix} 0 \\ 0 \\ \beta_2 \end{pmatrix}.$$

Here, C_1 is an $n_1 \times (n_1 \cdot n_2)$ matrix, such that its horizontal dimension is indexed by pairs (j, k) with $1 \leq j \leq n_1$ and $1 \leq k \leq n_2$, and its $[i, (j, k)]$-th entry is defined by

$$C_1[i, (j, k)] = \begin{cases} B_1[i, j] & \text{if } \beta_1[j] = 1 \\ 0 & \text{otherwise} \end{cases}$$

for all $1 \leq i, j \leq n_1$ and $1 \leq k \leq n_2$. Similarly, C_2 is an $(n_1 \cdot n_2) \times n_2$ matrix, such that its vertical dimension is indexed by pairs (i,j) with $1 \leq i \leq n_1$ and $1 \leq j \leq n_2$, and its $[(i,j),k]$-th entry is defined by

$$C_2[(i,j),k] = \begin{cases} B_2[j,k] & \text{if } \alpha_2[j] = 1 \\ 0 & \text{otherwise} \end{cases}$$

for all $1 \leq i, j \leq n_1$ and $1 \leq k \leq n_2$. The set of atomic expressions in $B_{1,\cdot,2}$ is clearly the same as the set of atomic expressions in $\mathcal{F}_1 \cdot \mathcal{F}_2$.

Finally, if $\mathcal{F}$ is a rational expression over the hemiring $((W \times X) \cup E, +, \cdot, 0)$, it is equivalent by the induction hypothesis to some 0-1 automaton $\mathcal{B}_3 = (n_3, \alpha_3, B_3, \beta_3)$. A rational expression $\mathcal{F}^+$ is equivalent to the automaton $\mathcal{B}_+ = (n_3 + n_3^2, \alpha_+, B_+, \beta_+)$, such that

$$\alpha_+ = (\alpha_3, 0), \qquad B_+ = \begin{pmatrix} B_3 & C \\ D & 0 \end{pmatrix}, \qquad \beta_{1+2} = \begin{pmatrix} \beta_3 \\ 0 \end{pmatrix}.$$

Here, C is an $n_3 \times n_3^2$ matrix, such that its horizontal dimension is indexed by pairs (j,k) with $1 \leq j, k \leq n_3$, and its $[i,(j,k)]$-th entry is defined by

$$C[i,(j,k)] = \begin{cases} B_3[i,j] & \text{if } \beta_3[j] = 1 \\ 0 & \text{otherwise} \end{cases}$$

for all $1 \leq i, j, k \leq n_3$. Similarly, D is an $n_3^2 \times n_3$ matrix, such that its vertical dimension is indexed by pairs (i,j) with $1 \leq i, j \leq n_3$, and its $[(i,j),k]$-th entry is defined by

$$D[(i,j),k] = \begin{cases} B_3[j,k] & \text{if } \alpha_3[j] = 1 \\ 0 & \text{otherwise} \end{cases}$$

for all $1 \leq i, j, k \leq n_3$. Obviously, the set of atomic expressions in B_+ is the same as the set of atomic expressions in $\mathcal{F}^+$.

As a result, $\mathcal{E}$ is equivalent to some proper finite automaton $\mathcal{A}$ over the h-aws $(W, X, E, +, \cdot, 0, \Sigma, \upsilon)$. □

Remark 28. *The same result can be obtained for rational expressions and automata over s-aws as well.*

Example 29. *Suppose we have k types of blocks $B_1, \ldots, B_k$ of (not necessarily distinct) positive integer heights $h_1, \ldots, h_k$. Blocks of different types have different colours. For all positive integers n, let F_n denote the minimum number of blocks (in total through all types) needed in order to build all possible towers of height n.*

Let us take the h-aws of Example 14 – that is, both addition and multiplication in our structure correspond to ordinary addition on real numbers. Then the rational expression

$$\mathcal{E} = \left(a^{h_1} + a^{h_2} + \ldots + a^{h_k}\right)^+$$

denotes the power series

$$|\mathcal{E}| = \sum_{n \in \mathbb{N}_+} F_n a^n.$$

As demonstrated by the above example, rational expressions with auxiliary weights can be used as a convenient means for modelling, which is not limited to a direct "translation" of an automaton to the corresponding rational expression.

References

[1] R. Alur et al., Regular Functions and Cost Register Automata, *Logic in Computer Science, 28th Annual ACM/IEEE Symposium, LICS 2013* (IEEE Computer Society, 2013), pp. 13–22.

[2] M. Cadilhac, A. Krebs and N. Limaye, Value Automata with Filters, *Seventh Workshop on Non-Classical Models of Automata and Applications, NCMA 2015, Short Papers*, eds. R. Freund et al. (ÖCG, Vienna, 2015), pp. 13–21.

[3] K. Chatterjee, L. Doyen and T. A. Henzinger, Alternating Weighted Automata, *Fundamentals of Computation Theory, 17th Int. Symposium, FCT 2009*, eds. M. Kutyłowski, W. Charatonik and M. Gębala (Springer, Heidelberg, 2009), pp. 3–13.

[4] K. Chatterjee, L. Doyen and T. A. Henzinger, Expressiveness and Closure Properties for Quantitative Languages, *Proc. 24th Annual IEEE Symposium on Logic in Computer Science, LICS 2009* (IEEE Comp. Soc. Press, 2009), pp. 199–208.

[5] K. Chatterjee, L. Doyen and T. A. Henzinger, Quantitative Languages, *ACM Trans. Comput. Log.* **11**(4) (2010) Article 23.

[6] M. Droste, W. Kuich and H. Vogler (eds.), *Handbook of Weighted Automata* (Springer, Heidelberg, 2009).

[7] M. Droste and W. Kuich, Weighted Finite Automata over Hemirings, *Theor. Comput. Sci.* **485** (2013) 38–48.

[8] M. Droste and I. Meinecke, Describing Average- and Longtime-Behavior by Weighted MSO Logics, *Mathematical Foundations of Computer Science 2010, 35th Int. Symposium, MFCS 2010*, eds. P. Hliněný and A. Kučera (Springer, Heidelberg, 2010), pp. 537–548.

[9] M. Droste and I. Meinecke, Weighted Automata and Regular Expressions over Valuation Monoids, *Int. J. Found. Comput. Sci.* **22**(8) (2011) 1829–1844.

[10] M. Droste and I. Meinecke, Weighted Automata and Weighted MSO Logics for Average and Long-Time Behaviors, *Inf. Comput.* **220** (2012) 44–59.

[11] M. Droste and V. Perevoshchikov, Multi-weighted Automata and MSO Logic, *8th Int. Computer Science Symposium in Russia, CSR 2013*, eds. A. A. Bulatov and A. M. Shur (Springer, Heidelberg, 2013), pp. 418–430.

[12] E. Filiot, R. Gentilini and J.-F. Raskin, Quantitative Languages Defined by Functional Automata, *Concurrency Theory, 23rd Int. Conference, CONCUR 2012*, eds. M. Koutny and I. Ulidowski (Springer, Heidelberg, 2012), pp. 132–146.

[13] J. S. Golan, *Semirings and their Applications* (Kluwer Academic Publishers, Dordrecht, 1999).

[14] I. Meinecke, Valuations of Weighted Automata: Doing It in a Rational Way, *Algebraic Foundations in Computer Science*, eds. W. Kuich and G. Rahonis (Springer, Heidelberg, 2011), pp. 309–346.

[15] J. Sakarovitch, *Elements of Automata Theory* (Cambridge University Press, Cambridge, 2009).