Finite-state Machines: Theory and Applications

Weighted Finite-state Automata

Thomas Hanneforth

Institut für Linguistik
Universität Potsdam

October 29, 2008

Overview

1. Semirings
2. Weighted finite-state automata
3. Semiring properties
4. Closure properties and algebra of weighted finite-state automata
5. Shortest-distance algorithms
6. Equivalence transformations

Outline

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Abstract weights

- We add *weights* to finite-state machines to enhance their capabilities.
- The notion of a weight is an abstract one: everything which fulfills the conditions of an abstract algebraic structure called a *semiring* can be used as a weight in a FSM.
- So weights can be:
  - Real numbers
  - Probabilities
  - Distances
  - Strings
  - Feature structures
  - Sets
  - Matrices
  - ...

Semirings

Monoid

Definition (Monoid)

An algebraic structure $\langle W, \diamond, \overline{1} \rangle$ is a *monoid* if:

1. $\diamond$ is a binary operation on W: $\diamond : W \times W \mapsto W$
2. $\diamond$ is associative: $\forall x, y, z \in W : x \diamond y \diamond z = (x \diamond y) \diamond z = x \diamond (y \diamond z)$
3. $\overline{1}$ is the identity element for $\diamond$: $\forall x \in W : x \diamond \overline{1} = \overline{1} \diamond x = x$

Example (The following structures are monoids:)

- $\langle \mathbb{R}, +, 0 \rangle$ (addition of real numbers)
- $\langle \mathbb{R}, \cdot, 1 \rangle$ (multiplication of real numbers)
- $\langle \mathbb{R} \cup \{\infty\}, min, \infty \rangle$ (minimum of real numbers)
- $\langle 2^M, \cup, \emptyset \rangle$ and $\langle 2^M, \cap, M \rangle$ for each set M
- $\langle \Sigma^*, \cdot, \varepsilon \rangle$ (concatenation monoid)
- $\langle 2^{\Sigma^*}, \cdot, \{\varepsilon\} \rangle$ (formal language monoid)
- $\langle \mathcal{F}, \sqcup, \perp \rangle$ (unification of feature structures, $\perp$ = empty feature structure)

Semirings

Commutative Monoid

Definition (Commutative monoid)

A monoid $\langle W, \diamond, \overline{1} \rangle$ is called *commutative* if $\forall x, y \in W : x \diamond y = y \diamond x$

Example

The following monoids are commutative:

- $\langle \mathbb{R}, +, 0 \rangle$
- $\langle \mathbb{R}, \cdot, 1 \rangle$
- $\langle \mathbb{R} \cup \{\infty\}, min, \infty \rangle$,
- $\langle \mathcal{F}, \sqcup, \bot \rangle$,
- $\langle 2^M, \cup, \emptyset \rangle$ and $\langle 2^M, \cap, M \rangle$ for each set M

The following monoids are not commutative:

- $\langle \Sigma^*, \cdot, \varepsilon \rangle$
- $\langle 2^{\Sigma^*}, \cdot, \{\varepsilon\} \rangle$

Semirings

Definition

Definition (Semiring)

An algebraic structure $\mathcal{K} = \langle W, \oplus, \otimes, \overline{0}, \overline{1} \rangle$ is a semiring if it fulfills the following conditions:

1. W is a nonempty set, called the *carrier set* of $\mathcal{K}$.
2. $\langle W, \oplus, \overline{0} \rangle$ is a commutative monoid with $\overline{0}$ as the identity element for $\oplus$.
3. $\langle W, \otimes, \overline{1} \rangle$ is a monoid with $\overline{1}$ as the identity element for $\otimes$.
4. $\otimes$ distributes over $\oplus$:
   $\forall x, y, z \in W : x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$ (*left distributivity*)
   $\forall x, y, z \in W : (y \oplus z) \otimes x = (y \otimes x) \oplus (z \otimes x)$ (*right distributivity*)
5. $\overline{0}$ is an annihilator for $\otimes$: $\forall w \in W, w \otimes \overline{0} \otimes w = \overline{0}$.

Notational convention

In the following, we will identify a semiring $\mathcal{K}$ with its carrier set.

Semirings

Common semirings

Example (Common semirings)

$\langle \mathbf{W}, \oplus, \otimes, \overline{\mathbf{0}}, \overline{\mathbf{1}} \rangle$	**Name**	
$\langle \{t, f\}, \vee, \wedge, f, t \rangle$	boolean semiring	
$\langle \mathbb{R}, +, \cdot, 0, 1 \rangle$	real semiring	
$\langle [0, 1], +, \cdot, 0, 1 \rangle$	probabilistic semiring	
$\langle \mathbb{R}_\infty, min, +, \infty, 0 \rangle$	tropical semiring	
$\langle [0, 1], max, \cdot, 0, 1 \rangle$	Viterbi semiring	
$\langle \mathbb{R}_\infty, \oplus_{log}, +, \infty, 0 \rangle$	log semiring	$x \oplus_{log} y = -ln(e^{-x} + e^{-y})$
$\langle \mathbb{R}_{-\infty}, max, +, -\infty, 0 \rangle$	arctic semiring	
$\langle [0, 1], max, min, 0, 1 \rangle$	fuzzy semiring	
$\langle 2^{\Sigma^*}, \cup, \cdot, \emptyset, \{\varepsilon\} \rangle$	concatenation semiring	2^{Σ^*} = formal languages
$\langle \Sigma^* \cup \{s_\infty\}, \wedge, \cdot, s_\infty, \varepsilon \rangle$	string semiring	$\wedge$ = *longest common prefix*
$\langle 2^M, \cup, \cap, \emptyset, M \rangle$	set semiring	M is an arbitrary set
$\langle 2^{\mathcal{F}}, \cup, \sqcup, \emptyset, \{\perp\} \rangle$	unification semiring	$\mathcal{F}$ = set of feature structures

String semiring

The string semiring is defined as $\langle \Sigma^* \cup \{s_\infty\}, \wedge, \cdot, s_\infty, \varepsilon\rangle$. That means:

- The longest common prefix operation $\wedge$ plays the role of abstract addition:
  $x \wedge y = z$ if $x = zx'$ and $y = zy'$ (z is of maximum length)
- We augment Σ^* with a special string s_∞ – the infinite string – which is the identity element of $\wedge$:
  $\forall x \in \Sigma^* \cup \{s_\infty\} : x \wedge s_\infty = s_\infty \wedge x = x$

Example

$cat \wedge car = ca$
$cat \wedge dog = \varepsilon$
$dog \wedge dogs = dog$
$cat \wedge s_\infty = cat$

Semirings

Isomorphic semirings

Definition (Isomorphic semirings)

A *semiring isomorphism* is a function f such that f and f^{-1} are *semiring homomorphisms*.

Let $\mathcal{K}_1 = \langle W_1, \oplus_1, \otimes_1, \overline{0}_1, \overline{1}_1 \rangle$ and $\mathcal{K}_2 = \langle W_2, \oplus_2, \otimes_2, \overline{0}_2, \overline{1}_2 \rangle$ be semirings.

A function $f : \mathcal{K}_1 \mapsto \mathcal{K}_2$ is a *semiring homomorphism* if:

- $f(\overline{1}_1) = \overline{1}_2$
- $f(a \oplus_1 b) = f(a) \oplus_2 f(b)$, for all a and b in $\mathcal{K}_1$
- $f(a \otimes_1 b) = f(a) \otimes_2 f(b)$ for all a and b in $\mathcal{K}_1$

Example (Isomorphic semirings)

- The tropical and Viterbi semirings are isomorphic under the $-ln()$.
- The real and log semirings are isomorphic under the $ln()$ function.

Practical importance of semiring isomorphisms

Normally, we $-log$-transform probabilities to avoid numerical underflow.

Outline

Weighted Finite-state Automata

Weighted finite-state acceptors

Definition (Weighted finite-state acceptor)

A *weighted finite-state acceptor* (WFSA) $A = \langle \Sigma, Q, q_0, F, E, \lambda, \rho \rangle$ over a semiring $\mathcal{K}$ is a 7-tuple with:

- Σ, the finite input alphabet
- Q, the finite set of states
- $q_0 \in Q$, the start state
- $F \subseteq Q$ the set of final states
- $E \subseteq Q \times (\Sigma \cup \{\varepsilon\}) \times \mathcal{K} \times Q$, the set of transitions
- $\lambda \in K$, the *initial weight* and
- $\rho : F \mapsto K$, the *final weight function*

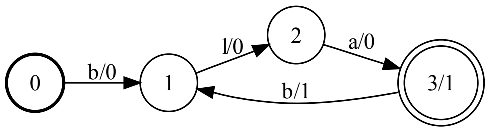
Notation

If t is a transition $\in E$, we denote with $w[t]$ the weight of t and $l[t]$ the label $\in \Sigma \cup \{\varepsilon\}$ of t.

Weighted Finite-state Automata

Weighted finite-state acceptors

Example (WFSA accepting/generating politicians' speeches)

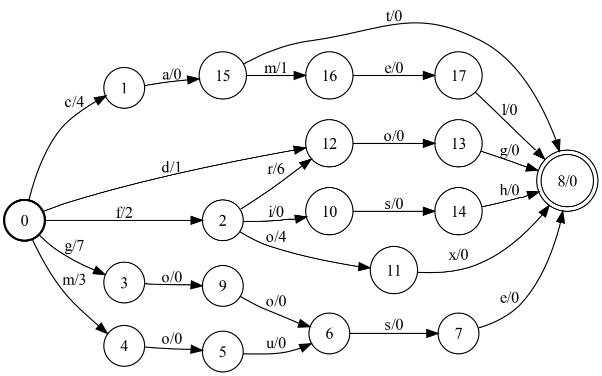


- Under the assumption that we **add** weights along a path (and in the end **add** the weight of the reached final state), this WFSA defines the following weighted language: $L(A_{blabla}) = \{bla \rightarrow 1, blabla \rightarrow 2, blablabla \rightarrow 3, \ldots, bla^n \rightarrow n\}$.
- This means: A_{blabla} can measure the length of these speeches.
- This in turn shows: Weighted finite-state automata can **count**!

Weighted Finite-state Automata

Weighted finite-state acceptors

Example (WFSA mapping words to word indices)



- Again, if we **add** weights along a path we obtain for each accepted word a unique number, for example: *fox* $\mapsto 2 + 4 + 0 + 0 = 6$.
- This technique is also called *perfect hashing*.

Weighted Finite-state Automata

Weighted finite-state transducers

Definition (Weighted finite-state transducer)

A *weighted finite-state transducer* (WFST) $A = \langle \Sigma, \Delta, Q, q_0, F, E, \lambda, \rho \rangle$ over a semiring $\mathcal{K}$ is a 8-tuple with

- Σ, the finite input alphabet
- Δ, the finite output alphabet
- Q, the finite set of states
- $q_0 \in Q$, the start state
- $F \subseteq Q$ the set of final states
- $E \subseteq Q \times (\Sigma \cup \{\varepsilon\}) \times (\Delta \cup \{\varepsilon\}) \times \mathcal{K} \times Q$, the set of transitions
- $\lambda \in K$, the *initial weight* and
- $\rho : F \mapsto K$, the *final weight function* mapping final states to elements in $\mathcal{K}$

Example

Weighted Finite-state Automata

Weight of a string - the Master formula

Definition (Master formula)

Let $A = \langle \Sigma, Q, q_0, F, E, \lambda, \rho \rangle$ be a WFSA and $\mathcal{K} = \langle W, \oplus, \otimes, \overline{0}, \overline{1} \rangle$ the semiring associated with A. Let $\pi = t_1 t_2 \ldots t_k$ be a path in A, that is, a sequence of adjacent transitions.
Let $\omega[\pi]$ be the abstract multiplication of the weights along π:
$\omega[\pi] = w[t_1] \otimes w[t_2] \otimes \ldots \otimes w[t_k]$.
Let $x \in \Sigma^*$ be a string, such that $x = l[t_1] \cdot l[t_2] \cdot \ldots \cdot l[t_k]$.
Let $\Pi(P, x, P')$ be the set of paths starting at states in P, ending in states in P' and thereby deriving the string x.
The weight $[\![A]\!](x)$ assigned to a given string x is defined as follows:

$$[\![A]\!](x) = \bigoplus_{f \in F, \pi \in \Pi(\{q_0\}, x, \{f\})} \lambda \otimes \omega[\pi] \otimes \rho(f)$$

$[\![A]\!](x) = \overline{0}$, if $\Pi(\{q_0\}, x, F) = \emptyset$.

Weighted Finite-state Automata

What does the Master formula mean?

- For each path labeled with x, starting at the start state and reaching a final state, we combine a constant weight λ, the weight of the individual transitions and the weight assigned to that final state by ***abstract multiplication***.
- There may be more than one path for x, if A is non-deterministic. In that case, we ***abstractly add*** the weights of all paths.
- This in turn means: if A is deterministic there is at most on path for x in A. The contribution of $\bigoplus$ can then be neglected.
- If there is no path for x, the weight assigned to x is $\overline{0}$. Thus, $\overline{0}$ – the identity element of $\oplus$ – signals, that a certain word x is ***not*** accepted by A.
- Note that $[\![A]\!](x)$ is a function mapping $x \in \Sigma^*$ to elements in $\mathcal{K}$.

Weighted Finite-state Automata

What does the Master formula mean?

- For each path labeled with x, starting at the start state and reaching a final state, we combine a constant weight λ, the weight of the individual transitions and the weight assigned to that final state by ***abstract multiplication***.
- There may be more than one path for x, if A is non-deterministic. In that case, we ***abstractly add*** the weights of all paths.
- This in turn means: if A is deterministic there is at most on path for x in A. The contribution of $\bigoplus$ can then be neglected.
- If there is no path for x, the weight assigned to x is $\overline{0}$. Thus, $\overline{0}$ – the identity element of $\oplus$ – signals, that a certain word x is ***not*** accepted by A.
- Note that $[\![A]\!](x)$ is a function mapping $x \in \Sigma^*$ to elements in $\mathcal{K}$.

Weighted Finite-state Automata

What does the Master formula mean?

- For each path labeled with x, starting at the start state and reaching a final state, we combine a constant weight λ, the weight of the individual transitions and the weight assigned to that final state by ***abstract multiplication***.
- There may be more than one path for x, if A is non-deterministic. In that case, we ***abstractly add*** the weights of all paths.
- This in turn means: if A is deterministic there is at most on path for x in A. The contribution of $\bigoplus$ can then be neglected.
- If there is no path for x, the weight assigned to x is $\overline{0}$. Thus, $\overline{0}$ – the identity element of $\oplus$ – signals, that a certain word x is ***not*** accepted by A.
- Note that $[\![A]\!](x)$ is a function mapping $x \in \Sigma^*$ to elements in $\mathcal{K}$.

Weighted Finite-state Automata

What does the Master formula mean?

- For each path labeled with x, starting at the start state and reaching a final state, we combine a constant weight λ, the weight of the individual transitions and the weight assigned to that final state by ***abstract multiplication***.
- There may be more than one path for x, if A is non-deterministic. In that case, we ***abstractly add*** the weights of all paths.
- This in turn means: if A is deterministic there is at most on path for x in A. The contribution of $\bigoplus$ can then be neglected.
- If there is no path for x, the weight assigned to x is $\overline{0}$. Thus, $\overline{0}$ – the identity element of $\oplus$ – signals, that a certain word x is ***not*** accepted by A.
- Note that $[\![A]\!](x)$ is a function mapping $x \in \Sigma^*$ to elements in $\mathcal{K}$.

Weighted Finite-state Automata

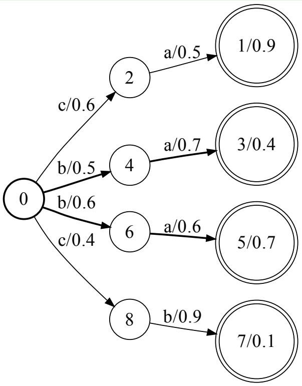
What does the Master formula mean?

- For each path labeled with x, starting at the start state and reaching a final state, we combine a constant weight λ, the weight of the individual transitions and the weight assigned to that final state by ***abstract multiplication***.
- There may be more than one path for x, if A is non-deterministic. In that case, we ***abstractly add*** the weights of all paths.
- This in turn means: if A is deterministic there is at most on path for x in A. The contribution of $\bigoplus$ can then be neglected.
- If there is no path for x, the weight assigned to x is $\overline{0}$. Thus, $\overline{0}$ – the identity element of $\oplus$ – signals, that a certain word x is ***not*** accepted by A.
- Note that $[\![A]\!](x)$ is a function mapping $x \in \Sigma^*$ to elements in $\mathcal{K}$.

Weighted Finite-state Automata

Master formula

Example (Weight interpretation)



WFSA A with numerical weights
($\lambda = \overline{1}$ in all cases)

$[\![A]\!](ba) =$

Weights interpreted in the tropical semiring:
$min(\{0.5+0.7+0.4, 0.6+0.6+0.7\}) = 1.6$

Weights interpreted in the probabilistic SR:
$0.5 \cdot 0.7 \cdot 0.4 + 0.6 \cdot 0.6 \cdot 0.7 = 0.392$

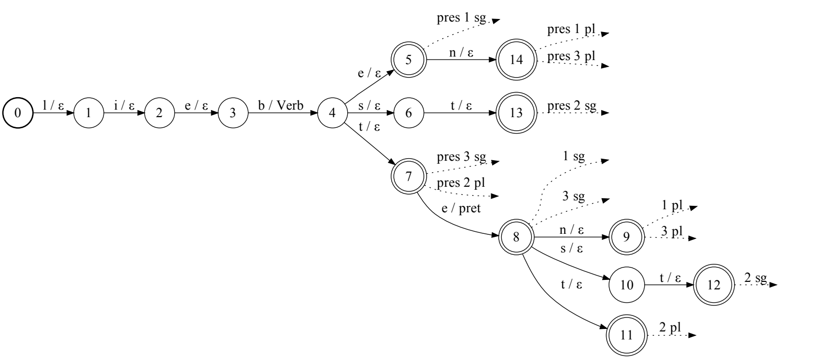
Weights interpreted in the Viterbi semiring:
$max(\{0.5 \cdot 0.7 \cdot 0.4, 0.6 \cdot 0.6 \cdot 0.7\}) = 0.252$

Weights interpreted in the (max, min) SR:
$max(\{min(\{0.5, 0.7, 0.4\}),$
$min(\{0.6, 0.6, 0.7\})\}) = 0.6$

Weighted Finite-state Automata

Master formula

Example (A_{lieben} representing some forms of German verb *lieben*)



$$[\![A_{lieben}]\!](liebst) = \varepsilon \cdot \varepsilon \cdot \varepsilon \cdot Verb \cdot \varepsilon \cdot \varepsilon \cdot pres\ 2\ sg$$

Weighted Finite-state Automata

Why are semirings a suitable weight structure for finite-state automata?

- The fact that the semiring is composed out of two monoids gives us:
  - two identity elements $\overline{0}$ and $\overline{1}$: $\overline{0}$ signals that some string is not part of the language accepted by the FSM, and $\overline{1}$, which gives us a notion of a "trivial" weight
  - a high degree of freedom in which order to compute weights (by associativity)
- Commutativity of $\oplus$ gives us the freedom to consider the paths according to the master formula in any order
- Distributivity of $\otimes$ enables us to factor out common "weight prefixes" of different paths (this property is for example important for weighted determinization)
- The annihilation relationship between $\otimes$ and $\overline{0}$ causes any path which contains a transition with weight $\overline{0}$ to be invalid

Weighted Finite-state Automata

Why are semirings a suitable weight structure for finite-state automata?

- The fact that the semiring is composed out of two monoids gives us:
  - two identity elements $\overline{0}$ and $\overline{1}$: $\overline{0}$ signals that some string is not part of the language accepted by the FSM, and $\overline{1}$, which gives us a notion of a "trivial" weight
  - a high degree of freedom in which order to compute weights (by associativity)
- Commutativity of $\oplus$ gives us the freedom to consider the paths according to the master formula in any order
- Distributivity of $\otimes$ enables us to factor out common "weight prefixes" of different paths (this property is for example important for weighted determinization)
- The annihilation relationship between $\otimes$ and $\overline{0}$ causes any path which contains a transition with weight $\overline{0}$ to be invalid

Weighted Finite-state Automata

Why are semirings a suitable weight structure for finite-state automata?

- The fact that the semiring is composed out of two monoids gives us:
  - two identity elements $\overline{0}$ and $\overline{1}$: $\overline{0}$ signals that some string is not part of the language accepted by the FSM, and $\overline{1}$, which gives us a notion of a "trivial" weight
  - a high degree of freedom in which order to compute weights (by associativity)
- Commutativity of $\oplus$ gives us the freedom to consider the paths according to the master formula in any order
- Distributivity of $\otimes$ enables us to factor out common "weight prefixes" of different paths (this property is for example important for weighted determinization)
- The annihilation relationship between $\otimes$ and $\overline{0}$ causes any path which contains a transition with weight $\overline{0}$ to be invalid

Weighted Finite-state Automata

Why are semirings a suitable weight structure for finite-state automata?

- The fact that the semiring is composed out of two monoids gives us:
  - two identity elements $\overline{0}$ and $\overline{1}$: $\overline{0}$ signals that some string is not part of the language accepted by the FSM, and $\overline{1}$, which gives us a notion of a "trivial" weight
  - a high degree of freedom in which order to compute weights (by associativity)
- Commutativity of $\oplus$ gives us the freedom to consider the paths according to the master formula in any order
- Distributivity of $\otimes$ enables us to factor out common "weight prefixes" of different paths (this property is for example important for weighted determinization)
- The annihilation relationship between $\otimes$ and $\overline{0}$ causes any path which contains a transition with weight $\overline{0}$ to be invalid

Weighted Finite-state Automata

Why are semirings a suitable weight structure for finite-state automata?

- The fact that the semiring is composed out of two monoids gives us:
  - two identity elements $\overline{0}$ and $\overline{1}$: $\overline{0}$ signals that some string is not part of the language accepted by the FSM, and $\overline{1}$, which gives us a notion of a "trivial" weight
  - a high degree of freedom in which order to compute weights (by associativity)
- Commutativity of $\oplus$ gives us the freedom to consider the paths according to the master formula in any order
- Distributivity of $\otimes$ enables us to factor out common "weight prefixes" of different paths (this property is for example important for weighted determinization)
- The annihilation relationship between $\otimes$ and $\overline{0}$ causes any path which contains a transition with weight $\overline{0}$ to be invalid

Weighted Finite-state Automata

Why are semirings a suitable weight structure for finite-state automata?

- The fact that the semiring is composed out of two monoids gives us:
  - two identity elements $\overline{0}$ and $\overline{1}$: $\overline{0}$ signals that some string is not part of the language accepted by the FSM, and $\overline{1}$, which gives us a notion of a "trivial" weight
  - a high degree of freedom in which order to compute weights (by associativity)
- Commutativity of $\oplus$ gives us the freedom to consider the paths according to the master formula in any order
- Distributivity of $\otimes$ enables us to factor out common "weight prefixes" of different paths (this property is for example important for weighted determinization)
- The annihilation relationship between $\otimes$ and $\overline{0}$ causes any path which contains a transition with weight $\overline{0}$ to be invalid

Weighted Finite-state Automata

Semirings play an import role in processing tasks based on finite-state automata, but are not limited to that.

- The ***tropical semiring*** is the classical semiring of all kinds of shortest-distance / best-analysis problems
- The ***probabilistic semiring*** plays a important role in statistical language processing for computing string probabilities.
- The same is true for the ***Viterbi semiring***, since it is related to processing tasks where we are interested in the analysis with the highest probability.
- The ***string semiring*** is used in string processing tasks, computational morphology.

Weighted Finite-state Automata

Significance of semirings

Semirings play an import role in processing tasks based on finite-state automata, but are not limited to that.

- The ***tropical semiring*** is the classical semiring of all kinds of shortest-distance / best-analysis problems
- The ***probabilistic semiring*** plays a important role in statistical language processing for computing string probabilities.
- The same is true for the ***Viterbi semiring***, since it is related to processing tasks where we are interested in the analysis with the highest probability.
- The ***string semiring*** is used in string processing tasks, computational morphology.

Semirings play an import role in processing tasks based on finite-state automata, but are not limited to that.

- The ***tropical semiring*** is the classical semiring of all kinds of shortest-distance / best-analysis problems
- The ***probabilistic semiring*** plays a important role in statistical language processing for computing string probabilities.
- The same is true for the ***Viterbi semiring***, since it is related to processing tasks where we are interested in the analysis with the highest probability.
- The ***string semiring*** is used in string processing tasks, computational morphology.

Weighted Finite-state Automata

Significance of semirings

Semirings play an import role in processing tasks based on finite-state automata, but are not limited to that.

- The ***tropical semiring*** is the classical semiring of all kinds of shortest-distance / best-analysis problems
- The ***probabilistic semiring*** plays a important role in statistical language processing for computing string probabilities.
- The same is true for the ***Viterbi semiring***, since it is related to processing tasks where we are interested in the analysis with the highest probability.
- The ***string semiring*** is used in string processing tasks, computational morphology.

Weighted rational languages as monoid homomorphisms

Weighted Finite-state Automata

Weighted regular relations

Definition (Weighted regular relation)

1.
2.

Outline

Semiring properties

Closure

An important operation on semiring elements is the *closure* operation $*$.

Definition (Closure)

Let $\mathcal{K} = \langle W, \oplus, \otimes, \overline{0}, \overline{1} \rangle$ be a semiring. The *closure* a^* of $a \in W$ is defined as:

$$a^* = \bigoplus_{n=0}^{\infty} a^n = \overline{1} \oplus a \oplus (a \otimes a) \oplus (a \otimes a \otimes a) \ldots = \overline{1} \oplus a \oplus a^2 \oplus a^3 \ldots$$

Example (Closures in different semirings)

Semiring	**a**	$\mathbf{a}^*$
tropical	0.5	$0.5^* = 0 \;\min\; 0.5 \;\min\; (0.5 + 0.5) \;\min\; \ldots = 0$
Viterbi	0.5	$0.5^* = 1 \;\max\; 0.5 \;\max\; (0.5 \cdot 0.5) \;\max\; \ldots = 1$
real	$\frac{1}{2}$	$\left(\frac{1}{2}\right)^* = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = 2$
string	*cat*	$cat^* = \varepsilon \wedge cat \wedge cat \cdot cat \wedge \ldots = \varepsilon$
concatenation	$\{a, b\}$	$\{a, b\}^* = \{\varepsilon\} \cup \{a, b\} \cup (\{a, b\} \cdot \{a, b\}) \cup \ldots = \{\varepsilon, a, b, aa, ab, ba, bb, \ldots\}$
unification	$\{[f : a]\}$	$\{[f : a]\}^* = \{\bot\} \cup \{[f : a]\} \cup \ldots = \{\bot, [f : a]\}$

Semiring properties

Semirings differ with respect a number of formal properties.

Definition (Semiring properties)

Let $\mathcal{K} = \langle W, \oplus, \otimes, \overline{0}, \overline{1} \rangle$ be a semiring.

- **Finiteness**: $\mathcal{K}$ is said to be *finite* if W is finite.
- **Commutativity**: $\mathcal{K}$ is said to be *commutative* if $\otimes$ is commutative: $\forall x, y \in W : x \otimes y = y \otimes x$. Note that $\oplus$ is commutative by definition.
- **Idempotency**: $\mathcal{K}$ is said to be *idempotent* if $\oplus$ is idempotent: $\forall x \in W : x \oplus x = x$
- **Boundedness**: $\mathcal{K}$ is said to be *bounded* if $\overline{1}$ is an annihilator for $\oplus$: $\forall x \in W : \overline{1} \oplus x = x \oplus \overline{1} = \overline{1}$
- k**-Closedness**: $\mathcal{K}$ is said to be *k-closed* (for a fixed integer k) if

$$\forall x \in W : \bigoplus_{n=0}^{k} x^n = \bigoplus_{n=0}^{k+1} x^n$$

Semiring properties

Properties of individual semirings

The tropical, Viterbi and string semirings are idempotent, bounded and 0-closed:

	+idempotent	**+bounded**	**0-closed**
tropical	$a \text{ min } a = a$	$a \text{ min } 0 = 0$	$a^0 = a^1 = 0 \text{ min } a = 0$
Viterbi	$a \text{ max } a = a$	$a \text{ max } 1 = 1$	$a^0 = a^1 = 1 \text{ max } a = 1$
string	$a \wedge a = a$	$a \wedge \varepsilon = \varepsilon$	$a^0 = a^1 = \varepsilon \wedge a = \varepsilon$

The real semiring is not idempotent, not bounded and not k-closed:

	−idempotent	**−bounded**	**−k-closed**
Real	$a + a = 2a \neq a$	$a + 1 \neq a$	$\sum_{n=0}^{\infty} a^n = \begin{cases} \infty & \text{if } n \geq 1 \\ \frac{1}{1-a} & \text{if } n < 1 \end{cases}$

Semiring properties

Closed semiring

Definition (Closed semiring)

Let $\mathcal{K} = \langle W, \oplus, \otimes, \overline{0}, \overline{1} \rangle$ be a semiring. $\mathcal{K}$ is called *closed* if

- Every finite or infinite sum of elements $a_i \in W$ is in W too:
  $$\forall \{a_1, \ldots, a_n\} \subseteq W : \bigoplus_{i=1}^{n} a^i \in W$$
- Associativity and commutativity hold for countable sums:
- Distributivity hold for countable sums:
  $$\forall \{a_1, \ldots, a_n\}, \{b_1, \ldots, b_m\} \subseteq W \colon \bigoplus_{i=1}^{n} a^i \otimes \bigoplus_{j=1}^{m} b^j = \bigoplus_{i=1}^{n} \Big(\bigoplus_{j=1}^{m} (a^i \otimes b^j) \Big)$$

Note

Closed semirings play an important role in shortest-distance algorithms.

Semiring properties

Definition (Semiring properties)

Let $\mathcal{K} = \langle W, \oplus, \otimes, \overline{0}, \overline{1} \rangle$ be a semiring.

- **Left divisibility**: $\mathcal{K}$ is said to be *left divisible* if $\forall x \in W, x \neq \overline{0}$, $\exists y \in W$ such that $y \otimes x = \overline{1}$ (that is, every element has a *left inverse*)
- **Weak left divisibility**: $\mathcal{K}$ is said to be *weakly left divisible* if $\forall x, y \in W, x \oplus y \neq \overline{0}, \exists z \in W : x = (x \oplus y) \otimes z$.
  If z is unique we write $z = (x \oplus y)^{-1} \otimes x$.
- **$\overline{0}$-divisor-freeness**: $\mathcal{K}$ is said to be $\overline{0}$-*divisor-free* if $\neg \exists x, y \in W, x, y \neq \overline{0}$, such that $x \otimes y = \overline{0}$
- **$\overline{0}$-sum-freeness**: $\mathcal{K}$ is said to be $\overline{0}$-*sum-free* if $\neg \exists x, y \in W, x, y \neq \overline{0}$, such that $x \oplus y = \overline{0}$

Semiring properties

Example (Weak left divisibility)

$\forall \mathbf{x}, \mathbf{y} \in \mathbf{W}, \mathbf{x} \oplus \mathbf{y} \neq \overline{\mathbf{0}}, \exists \mathbf{z} \in \mathbf{W} : \mathbf{x} = (\mathbf{x} \oplus \mathbf{y}) \otimes \mathbf{z}.$

- *In the tropical semiring:* $x = 0.5, y = 0.2, z = 0.3$
  $(0.5 \min 0.2) + 0.3 = 0.5$
  $-(0.5 \min 0.2) + 0.5 = -0.2 + 0.5 = 0.3$
  $\mathbf{a^{-1}} = -\mathbf{a}$

- *In the real semiring:* $x = \frac{3}{10}, y = \frac{2}{10}, z = \frac{3}{5}$
  $(\frac{3}{10} + \frac{2}{10}) \cdot \frac{3}{5} = \frac{3}{10}$
  $\frac{1}{\frac{3}{10}+\frac{2}{10}} \cdot \frac{3}{10} = 2 \cdot \frac{3}{10} = \frac{3}{5}$
  $\mathbf{a^{-1}} = \frac{\mathbf{1}}{\mathbf{a}}$

- *In the string semiring:* $x = cats, y = cars, z = ts$
  $(cats \wedge car) \cdot ts = ca \cdot ts = cats$
  $(cats \wedge car)^{-1} \cdot cats = ca^{-1} \cdot cats = ts$
  $\mathbf{a^{-1}} = \textit{inverse strings}$

Semiring properties

Left and right semirings

Definition (Left and right semirings)

A semiring $\mathcal{K}$ is called a *left semiring* if it is left distributive.
A semiring $\mathcal{K}$ is called a *right semiring* if it is right distributive.

Example (String semiring)

The string semiring is only a left semiring:

$x \cdot (cat \wedge car) = xcat \wedge xcar = xca$

$(cat \wedge car) \cdot s = ca \cdot s = cas \neq cats \wedge cars = ca$

Semiring properties

Natural order

Definition (Natural order)

Let $\mathcal{K}$ be an idempotent semiring. We may define a partial order $\leq_{\mathcal{K}}$ – called the *natural order* – on the elements in $\mathcal{K}$ in the following way:

$$a \leq_{\mathcal{K}} b \equiv a \oplus b = a$$

Theorem ($\leq_{\mathcal{K}}$ is a partial order)

$\leq_{\mathcal{K}}$ is reflexive, antisymmetric and transitive.

Proof.

1. Reflexivity:
2. Antisymmetry:
3. Transitivity:

□

Semiring properties

Natural order

Based on the notion of a natural order we can define more semiring properties:

Definition

Let $\mathcal{K} = \langle W, \oplus, \otimes, \overline{0}, \overline{1} \rangle$ be a semiring. Let $\leq_{\mathcal{K}}$ be a partial order on the elements in $\mathcal{K}$.

- **Negativity**: $\mathcal{K}$ is said to be *negative* if $\overline{1} \leq_{\mathcal{K}} \overline{0}$.
  $\mathcal{K}$ is said to be *positive* if $\overline{0} \leq_{\mathcal{K}} \overline{1}$.
- **Monotonicity**: $\mathcal{K}$ is said to be *monotonic* if for all $x, y, z \in W$:
  $(x \leq_{\mathcal{K}} y) \rightarrow (x \oplus z \leq_{\mathcal{K}} y \oplus z)$
  $(x \leq_{\mathcal{K}} y) \rightarrow (x \otimes z \leq_{\mathcal{K}} y \otimes z)$
  $(x \leq_{\mathcal{K}} y) \rightarrow (z \otimes x \leq_{\mathcal{K}} z \otimes y)$
- **Superiority**: $\mathcal{K}$ is said to be *superior* if for all $x, y \in W$:
  $x \leq_{\mathcal{K}} x \otimes y$ and
  $y \leq_{\mathcal{K}} x \otimes y$

Semiring properties

Natural order: significance

Superiority and Monotonicity are important properties for shortest-distance algorithms on WFSA:

- Superiority, that is $x \ \leq_{\mathcal{K}} \ x \otimes y$, means that the weight x of a string will not get better if you multiply it with the weight y of another transition.
- Monotonicity ensures an optimal substructure of shortest-distance problems: a problem where we haven to decide whether $x \otimes z \leq_{\mathcal{K}} y \otimes z$ holds can be reduced to the question whether $x \leq_{\mathcal{K}} y$ holds.

Semiring properties

Theorem

1. *Every superior semiring is negative.*
2. *Every bounded semiring is idempotent.*

Proof.

1. $$\begin{array}{ll} \forall a \in \mathcal{K}: \ \overline{1} \leq \overline{1} \otimes a = a & \text{(by superiority)} \\ \forall a \in \mathcal{K}: \ a \leq \overline{0} \otimes a = \overline{0} & \text{(by superiority and annihilator property)} \\ \forall a \in \mathcal{K}: \ \overline{1} \leq a \leq \overline{0} & \end{array}$$

2. $$\begin{array}{ll} a \oplus \overline{1} = a & \text{(Boundedness)} \\ \overline{1} \oplus \overline{1} = \overline{1} & (a = \overline{1}) \\ a \oplus a = a & \text{(multiplication with } a\text{)} \end{array}$$

□

Semiring properties

Summary: Semirings and their properties

$\langle \mathbf{W}, \oplus, \otimes, \bar{\mathbf{0}}, \bar{\mathbf{1}} \rangle$	**Properties**
$\langle \{t, f\}, \vee, \wedge, f, t \rangle$	*boolean semiring*: finite, commutative, positive, idempotent, bounded, 0-closed, closed
$\langle \mathbb{R}, +, \cdot, 0, 1 \rangle$	*real semiring*: infinite, commutative, non-idempotent, monotonic, non-bounded, non-k-closed, non-closed
$\langle \mathbb{R}_\infty, \oplus_{log}, +, \infty, 0 \rangle$	*log semiring*: infinite, commutative, non-idempotent, monotonic, non-bounded, non-k-closed, non-closed
$\langle \mathbb{R}_\infty, min, +, \infty, 0 \rangle$	*tropical semiring*: infinite, commutative, negative, idempotent, monotonic, 0-closed, bounded
$\langle [0,1], max, \cdot, 0, 1 \rangle$	*Viterbi semiring*: infinite, commutative, negative, idempotent, monotonic, 0-closed, bounded
$\langle \mathbb{R}_{-\infty}, max, +, -\infty, 0 \rangle$	*arctic semiring*
$\langle [0,1], max, min, 0, 1 \rangle$	*fuzzy semiring*

Semiring properties

Summary: Semirings and their properties

$\langle \mathbf{W}, \oplus, \otimes, \bar{\mathbf{0}}, \bar{\mathbf{1}} \rangle$	**Properties**
$\langle 2^{\Sigma^*}, \cup, \cdot, \emptyset, \{\varepsilon\} \rangle$	*concatenation semiring*: infinite, non-commutative, negative, idempotent, non-k-closed, closed, natural order: $\subseteq$
$\langle \Sigma^*, \wedge, \cdot, \emptyset, \varepsilon \rangle$	*string semiring*: infinite, left semiring, non-commutative, negative, idempotent, monotonic, bounded, 0-closed, natural order: is_prefix_of
$\langle 2^{\mathcal{F}}, \cup, \sqcup, \emptyset, \{\perp\} \rangle$	*unification semiring*: infinite, left semiring, commutative, negative, idempotent, monotonic, bounded, 1-closed, natural order: $\sqsubseteq$ (subsumption)

Semiring properties

Semiring properties and FSM algebra

Some operations of the FSM algebra require certain properties of the underlying semiring:

- Intersection and composition are only defined for FSMs based on commutative semirings.
- Composition of WFSTs with $\varepsilon : x$ or $x : \varepsilon$ transitions must be parameterized for non-idempotent semirings.
- Shortest-distance algorithms usually require idempotent and negative semirings
- Weighted determinization is only defined for weakly left-divisible semirings.
- Making an FSM connected (remove all paths with path weight $\overline{0}$) usually requires $\overline{0}$-divisor-freeness

Outline

Closure properties and algebra of WFSA

Weighted finite-state acceptors are closed under the following operations:

- Union
- Concatenation
- Closure
- Intersection
- Difference with unweighted finite-state acceptors
- Reversal
- Substitution / homomorphism
- Cross product

Weighted finite-state acceptors are not closed under:

- Complementation

Closure properties and algebra of WFSM

The set of weighted finite state transducers is closed under

- Union
- Concatenation
- Closure
- Reversal
- Projection (note that this leads to FSAs)
- Composition
- Inversion

Weighted finite state transducers are **not** closed under

- Complementation
- Intersection (but acyclic and ε-free transducers are)
- Difference

Closure properties and algebra of WFSM

The set of weighted finite state transducers is closed under

- Union
- Concatenation
- Closure
- Reversal
- Projection (note that this leads to FSAs)
- Composition
- Inversion

Weighted finite state transducers are **not** closed under

- Complementation
- Intersection (but acyclic and ε-free transducers are)
- Difference

Definition (Weighted closure)

Let $A = \langle Q, \Sigma, q_0, F, E, \lambda, \rho \rangle$ be a weighted finite-state acceptor with underlying semiring $\mathcal{K} = \langle W, \oplus, \otimes, \overline{0}, \overline{1} \rangle$.
The *closure* of A – denoted by A^* – is defined as:

$$A^* = \langle Q \cup \{q'_0\}, \Sigma, q'_0, F \cup \{q'_0\}, E \cup E_\varepsilon, \lambda, \rho' \rangle$$

with

$$E_\varepsilon = \{\langle q, \varepsilon, \rho(q), q'_0 \rangle \mid \forall q \in F\} \cup \{\langle q'_0, \varepsilon, \overline{1}, q_0 \rangle\}$$

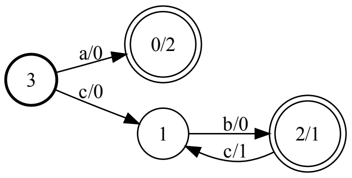
and

$$\rho'(q) = \begin{cases} \rho(q) & \forall q \in F \\ \overline{1} & q = q'_0 \end{cases}$$

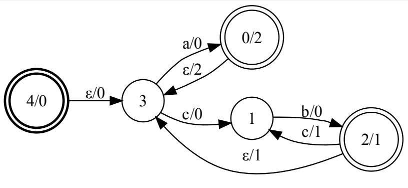
Closure properties and algebra of WFSA

Weighted closure

Example (Weighted closure in the tropical semiring)



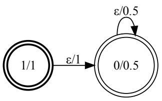
A



A^*

Example (Weighted closure in the probabilistic semiring)

A



A^*

Closure properties and algebra of WFSA

Weighted intersection

Definition (Intersection of two weighted regular languages)

Let A_1 and A_2 be two weighted finite-state acceptors over alphabets Σ_1 and Σ_2, resp. The intersection $A_1 \cap A_2$ is defined in the following way:

$$\forall x \in \Sigma_1^* \cap \Sigma_2^* \;\;:\;\; [\![A_1 \cap A_2]\!](x) = [\![A_1]\!](x) \otimes [\![A_2]\!](x)$$

Closure properties and algebra of WFSA

Weighted intersection

Definition (Weighted intersection of two finite-state acceptors)

Let $A_1 = \langle Q_1, \Sigma_1, q_{0_1}, F_1, E_1, \lambda_1, \rho_1 \rangle$ and $A_2 = \langle Q_2, \Sigma_2, q_{0_2}, F_2, E_2, \lambda_2, \rho_2 \rangle$ be two FSAs. $A_1 \cap A_2$, the intersection of A_1 and A_2, is an acceptor:

$$A = \langle Q_1 \times Q_2, \Sigma_1 \cap \Sigma_2, \langle q_{0_1}, q_{0_2} \rangle, F_1 \times F_2, E, \lambda_1 \otimes \lambda_2, \rho \rangle \text{ where}$$

- $\langle \langle p, q \rangle, a, w_1 \otimes w_2, \langle p', q' \rangle \rangle \in E$
  if $\langle p, a, w_1, p' \rangle \in E_1$ and $\langle q, a, w_2, q' \rangle \in E_2$, for all $a \in \Sigma_1 \cap \Sigma_2$.
- $\rho(\langle p, q \rangle) = \rho_1(p) \otimes \rho_2(q), \forall p \in F_1, \forall q \in F_2$

Comments

- The individual weights of the two operands are abstractly multiplied during intersection.
- Note that weighted intersection – unlike unweighted intersection – is no longer idempotent: $A \cap A \neq A$.
- Weighted composition is defined analogously.

Closure properties and algebra of WFSA

Weighted intersection

Example (Weighted intersection)

Closure properties and algebra of WFSA

Weighted intersection: why is it only possible in commutative semirings?

Closure properties and algebra of WFSA

Weighted composition in non-idempotent semirings

Definition (Composition of two weighted regular relations)

Let T_1 and T_2 be two weighted finite-state transducers over alphabets Σ_1, Δ_1 and Σ_2, Δ_2, resp. The composition $T_1 \circ T_2$ is defined in the following way:

$$\forall x \in \Sigma_1^*, y \in \Delta_2^* \;:\; [\![T_1 \circ T_2]\!](x, y) = \bigoplus_{z \in \Delta_1^* \cap \Sigma_2^*} [\![T_1]\!](x, z) \otimes [\![T_2]\!](z, y)$$

Example

Consider the two weighted relations in the real semiring
$R_1 = \{\langle abcd, ad\rangle \mapsto 1\}$ and
$R_2 = \{\langle ad, dea\rangle \mapsto 1\}$.
According to the definition above,
$R_1 \circ R_2 = \{\langle abcd, dea\rangle \mapsto 1\}$ with $z = ad$.

Closure properties and algebra of WFSA

Weighted composition in non-idempotent semirings

Definition (Composition of two weighted finite-state transducers)

Let $T_1 = \langle Q_1, \Sigma_1, \Delta_1, q_{0_1}, F_1, E_1, \lambda_1, \rho_1 \rangle$ and $T_2 = \langle Q_2, \Sigma_2, \Delta_2, q_{0_2}, F_2, E_2, \lambda_2, \rho_2 \rangle$ be two WFSTs.
$T_1 \circ T_2$, the composition of T_1 and T_2, is a transducer:

$$T = \langle Q_1 \times Q_2, \Sigma_1, \Delta_2, \langle q_{0_1}, q_{0_2} \rangle, F_1 \times F_2, E \cup E_\varepsilon \cup E_{i,\varepsilon} \cup E_{o,\varepsilon}, \lambda_1 \otimes \lambda_2, \rho \rangle$$

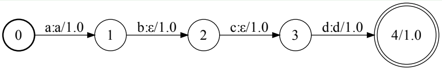
where

1. $E = \{\langle \langle p, q \rangle, a, b, w_1 \otimes w_2, \langle p', q' \rangle \rangle \mid$ $\exists c \in \Delta_1 \cap \Sigma_2 : \langle p, a, c, w_1, p' \rangle \in E_1 \wedge \langle q, c, b, w_2, q' \rangle \in E_2\}$
2. $E_\varepsilon = \{\langle \langle p, q \rangle, a, b, w_1 \otimes w_2, \langle p', q' \rangle \rangle \mid$ $\langle p, a, \varepsilon, w_1, p' \rangle \in E_1 \wedge \langle q, \varepsilon, b, w_2, q' \rangle \in E_2\}$
3. $E_{i,\varepsilon} = \{\langle \langle p, q \rangle, \varepsilon, a, w, \langle p, q' \rangle \rangle \mid \langle q, \varepsilon, a, w, q' \rangle \in E_2 \wedge p \in Q_1\}$
4. $E_{o,\varepsilon} = \{\langle \langle p, q \rangle, a, \varepsilon, w, \langle p', q \rangle \rangle \mid \langle p, a, \varepsilon, w, p' \rangle \in E_1 \wedge q \in Q_2\}$
5. $\rho(\langle p, q \rangle) = \rho_1(p) \otimes \rho_2(q), \forall p \in F_1, \forall q \in F_2$

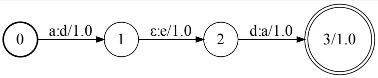
Closure properties and algebra of WFSA

Weighted composition in non-idempotent semirings

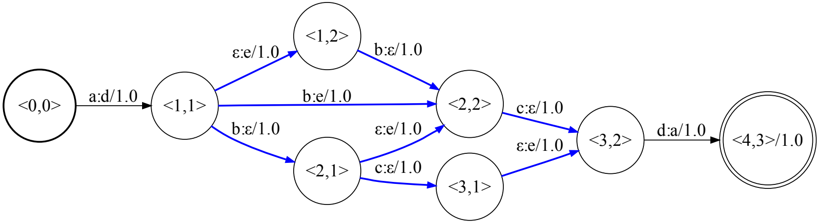
Example



T_1



T_2



$T_1 \circ T_2$

$T_1 \circ T_2$ corresponds to the weighted relation $\{\langle abcd, dea\rangle \mapsto 4\}$

Closure properties and algebra of WFSA

Weighted composition: ε-filter FSTs

- The problem is, that there are 4 paths deriving the string pair $bc : e$ with weight 1.0
- By the master formula, paths weights are added
- In idempotent semirings this doesn't cause harm, since $a \oplus a = a$. But in non-idempotent semirings, we are faced with a problem
- **Solution**: introduce a filter such that all paths except one are deleted

Closure properties and algebra of WFSA

Weighted composition: ε-filter FSTs

- The problem is, that there are 4 paths deriving the string pair $bc : e$ with weight 1.0
- By the master formula, paths weights are added
- In idempotent semirings this doesn't cause harm, since $a \oplus a = a$. But in non-idempotent semirings, we are faced with a problem
- **Solution**: introduce a filter such that all paths except one are deleted

Closure properties and algebra of WFSA

Weighted composition: ε-filter FSTs

- The problem is, that there are 4 paths deriving the string pair $bc : e$ with weight 1.0
- By the master formula, paths weights are added
- In idempotent semirings this doesn't cause harm, since $a \oplus a = a$. But in non-idempotent semirings, we are faced with a problem
- **Solution**: introduce a filter such that all paths except one are deleted

Closure properties and algebra of WFSA

Weighted composition: ε-filter FSTs

- The problem is, that there are 4 paths deriving the string pair $bc : e$ with weight 1.0
- By the master formula, paths weights are added
- In idempotent semirings this doesn't cause harm, since $a \oplus a = a$. But in non-idempotent semirings, we are faced with a problem
- **Solution**: introduce a filter such that all paths except one are deleted

Closure properties and algebra of WFSA

Weighted composition: ε-filter FSTs

Outline

Outline

Versions

- 29.10.2008: version 0.1 (initial version)