Finite automata modelling

Antoine Grimod

2022

Outline of the course
- weighted automata
- sequential transducers
- minimization
- learning algorithm L^* and variations

1 Lecture 1

Definition (Semiring)**.** $(\mathbb{K}, +, \cdot, 0, 1)$ is a set $\mathbb{K}$equipped with a commutative monoid structure $(\mathbb{K}, +, 0)$, a second monoid structure $(\mathbb{K}, \cdot, 1)$ such that the following axioms holds :
- $\forall x, y, z, \in \mathbb{K}, x \cdot (y + z) = xy + xz$ and $(y + z)x = yx + zx$
- $\forall x \in \mathbb{K}, 0 \cdot x = x \cdot 0 = 0$

Examples
- Boolean semiring $\mathbb{B} = \{0, 1\}$
- $\mathbb{N}, \mathbb{Z}, \mathbb{Q}_+, \mathbb{Q}, \mathbb{R}$
- Tropical semirings : $\mathbb{N}_{min} = (\mathbb{N} \cup \{+\infty\}, min, +, +\infty, 0)$
- if $(M, \cdot, 1)$ is a monoid then $(\mathcal{P}(M), \cup, \cdot, \emptyset, \{1_M\})$ is a semiring where $\cup$ is the union for $A, B \in \mathcal{P}(M), A \cdot B = \{a \cdot b | a \in A, b \in B\}$
- in particular $\mathcal{P}(A^*)$ has a semiring structure $Rat(A^*) \subseteq (A^*)$ is a subsemiring of $\mathcal{P}(A^*)$

Definition (semiring morphism)**.** A morphism between semiring $(A, +_A, \cdot_A, 0_A, 1_A)$, $(B, +_B, \cdot_B, 0_B, 1_B)$ is a function $f : A \to B$ such that $\forall x, y \in A$
- $f(x +_A y) = f(x) +_B f(y)$
- $f(x \cdot_A y) = f(x) \cdot_B f(y)$
- $f(1_A) = 1_B$
- $f(0_K) = 0_B$

Definition (Finite automaton)**.** A finite automaton over a finite alphabet A is $(Q, (\delta_a : Q \to Q)_{a \in A}), q_0, F)$ where Q is a finite set of states, $q_0 \in Q$ the initial state $F \subseteq Q$ the set of accepting states

Definition (Weighted automaton)**.** over A and a semiring $\mathbb{K}$ $(Q, (\delta_a : Q \to Q)_{a \in A}, i, f)$ where
- Q is a module over $\mathbb{K}$
- δ_a is a linear transformation for all $a \in A$
- $i : \mathbb{K} \to Q$ is the initial linear transformation
- $f : Q \to \mathbb{K}$ is the final linear transformation

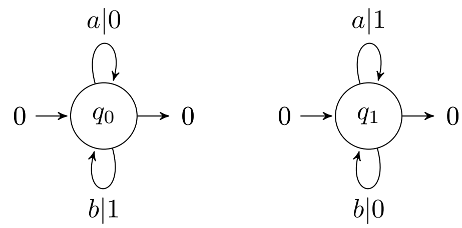
Definition (Another definition of weighted automaton)**.** $\mathcal{A} = (S, i, f, E \subseteq S \times A \times K \times Q)$ where E is the graph of a partial function from $S \times A \times S$ to $\mathbb{K} \setminus \{0_K\}$, with S finite, $i, f \in \mathbb{K}^S$

Notations and conventions
- we do not write output arrows from a state q if $f(q) = 0$ and likewise for $i(q) = 0$
- 2 notations for the graph, either the edge are labelled by wa where w is the weight and a the letter, or by $a|w$
- $(q_1, a, k, q_2) \in E$ can also be written $q_0 \xrightarrow{ka \text{ or } k|a} q_1$
- we omit writing the unit 1 in the case of the numerical semirings.

Definition (path)**.** A path in the automaton $\mathcal{A}$ is a sequence of transitions $l_1, \ldots, l_n$ of the form $l_i : p_i \xrightarrow{a_i|k_i} p_{i+1}$
The label of a path is $a_1 \ldots a_n$.
The weight of a path is the product in the semiring $\mathbb{K}$, $\mathfrak{w} = i(p_0) \cdot k_1 \cdot \ldots \cdot k_n \cdot f(p_{n+1})$

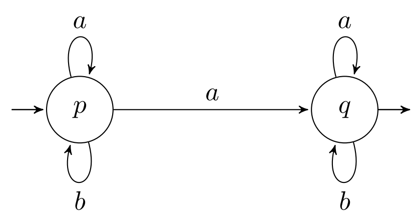
Definition. The language accepted by $\mathcal{A}$weighted over the semiring $\mathbb{K}$is a function $\mathcal{L}(A) : A^* \to \mathbb{K}$ computed as follow : $\mathcal{L}(A)(w) = \sum_{d \text{ path labelled by } w} \mathfrak{w}(d)$

Examples



$\mathcal{A}_1$ over $\ltimes_{min}$

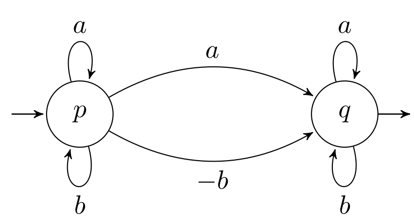
$\mathcal{L}(\mathcal{A}_1)(w) = min(|w|_a, |w|_b)$



$\mathcal{A}_2$ over $\mathbb{N}$

$\mathcal{L}(\mathcal{A}_2) = |w|_a$, all the paths have weight 1 (all weights are 1) and there is one path for every a in w.

Exercise Fin an automaton weighted over $\mathbb{Z}$ such that $\mathcal{L}(\mathcal{A}_3)(w) = |w|_a - |w|_b$



$\mathcal{A}_3$ over $\mathbb{Z}$

Definition ($\mathbb{K}$-series)**.** A $\mathbb{K}$-series over A^* is a function $s : A^* \to \mathbb{K}$.
The set of $\mathbb{K}$-series over A^* is denoted by $\mathbb{K}\langle\langle A^* \rangle\rangle$. $\mathbb{K}\langle\langle A^* \rangle\rangle$ has a $\mathbb{K}$-algebra structure such that we have the following operations :

Sum $s, t \in \mathbb{K}\langle\langle A^*\rangle\rangle$, $s+t$ is defined by $(s+t)(w) = s(w) + t(w)$ for $w \in A^*$
External left and right multiplication $\forall k \in \mathbb{K}, \forall s \in \mathbb{K}\langle\langle A^*\rangle\rangle$ we define $(s \cdot k)(w) = s(w) \cdot k$ and $(k \cdot s)(w) = k \cdot s(w)$
Cauchy product for $s, t \in \mathbb{K}\langle\langle A^*\rangle\rangle$ we define $(s \cdot t)(w) = \sum_{u,v \in A^* w=uv} s(u)t(v)$

We have the following properties :
— $\forall s, t, r \in \mathbb{K}\langle\langle A^*\rangle\rangle, (s+t) \cdot r = s \cdot r + t \cdot r$ and $r \cdot (s+t) = r \cdot s + r \cdot t$
— $\forall k, k' \in \mathbb{K}, s, t \in \mathbb{K}\langle\langle A^*\rangle\rangle$:
  — $k \cdot (s+t) = k \cdot s + k \cdot t$
  — $(s+t) \cdot k = s \cdot k + t \cdot k$
  — $k \cdot (k' \cdot s) = (k \cdot_{\mathbb{K}}) \cdot s$
  — $k \cdot (s \cdot t) = (k \cdot s) \cdot t$
  — $(s \cdot k) \cdot k' = s \cdot (k \cdot_{\mathbb{K}} k')$
  — $(s \cdot t) \cdot k = s \cdot (t \cdot k)$

Definition (support)**.** Given $s \in \mathbb{K} << A^* >>$ the support of s is defined as $supp(s) = \{w \in A^* | s(w) \neq 0_{\mathbb{K}}\}$

Given a $\mathbb{K}$-automaton $\mathcal{A}$ we get a Boolean automaton by replacing every non-zero transition in $\mathcal{A}$ with 1. Denote this automaton by $supp(\mathcal{A})$.

Exercise

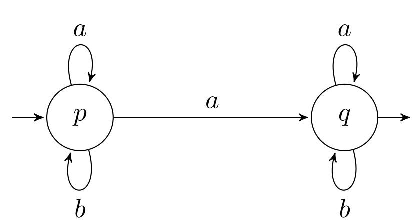
1. show that $supp(\mathcal{L}(\mathcal{A})) \subseteq \mathcal{L}(supp(\mathcal{A}))$
2. find a sufficient condition so that the equality holds

Answers

1. Let $w \in \text{supp}(\mathcal{L}(\mathcal{A}))$, then $\text{supp}(\mathcal{L}(\mathcal{A})) = \sum_{dlabelledbyw} \mathfrak{w}(d) \neq 0$, therefore there exists a path d labelled by w in $\mathcal{A}$ such that $\mathfrak{w}(d) \neq 0$ and $d : \xrightarrow{a_1|k_1} p_1 \to \ldots \xrightarrow{a_n|k_n} p_n$, thus $\mathfrak{w}(d) = i(p_0) \cdot k_1 \cdot \ldots \cdot k_n \cdot f(p_n)$, from which we can deduce that $k_i \neq 0$ for all i, therefore the path is also valid in $\text{supp}(\mathcal{L}(\mathcal{A}))$ and thus $w \in \mathcal{L}(\text{supp}(\mathcal{A}))$.
2. Having weights such that for 2 paths d, d' with $\mathfrak{w}(d) \neq 0$ and $\mathfrak{w}(d') \neq 0$, $\mathfrak{w}(d) + \mathfrak{w}(d') \neq 0$ is a sufficient to have the equality

Definition (Matrix representation)**.** The matrix representation of an automaton $\mathcal{A}$ is the matrix Δ such that for every states p and q, if there is k edges from p to q labelled by $l_1, \ldots, l_k$, some linear transformations, then $\Delta_{(p,q)} = \sum_{i \leq k} l_i$

Example Given the following automaton :



$\mathcal{A}$

its matrix representation is $\begin{pmatrix} a+b & a \\ 0 & 2a+b \end{pmatrix}$

Remark. This initial map $I : Q \to \mathbb{K}$ can be seen as a row vector and the final map $F : Q \to \mathbb{K}$ as a column vector

Lemma. (Δ, I, F) *the matrix representation of an automaton :*

$$\mathcal{L}(A)(w) = (I \cdot \Delta^{|w|} \cdot F)(w)$$

Property.

- $s \in \mathbb{K}\langle\langle A^* \rangle\rangle$ *can also be written as* $\sum_{w \in A^*} s(w) \cdot w$
- Δ^n *is the matrix of sums of "weighted labels" of paths of length* n

$$\forall p, q \in Q, (\Delta^{n+1})_{p,q} = \sum_{s \in Q} (\Delta^n)_{p,s} \cdot \Delta_{s,q}$$

- $\mathcal{L}(\mathcal{A}) = \sum_{n \in \mathbb{N}} I \cdot \Delta^n \cdot F = I \cdot (\sum_{n \in \mathbb{N}} \Delta^n) \cdot F$

Given a semiring $\mathbb{K}$, we would like to define the operation $(\cdot)^*$ by $k^* = \sum_{n \geq 0} k^n$. This is not always defined.

Definition. A family $(s_i)_{i \in I}$ of $\mathbb{K}\langle\langle A^* \rangle\rangle$ is locally finite when $\forall w \in A^*, \{i \in I | s_i(w) = 0\}$ is finite.

Theorem 1. *If* $(s_i)_{i \in I}$ *is a locally finite family of series, then we can define* $\sum_{i \in I} s_i$

Definition. A series $s \in \mathbb{K}\langle\langle A^* \rangle\rangle$ is proper if $s(\epsilon) = 0_{\mathbb{K}}$

If $s \in \mathbb{K}\langle\langle A^* \rangle\rangle$ is proper then the family $(s^n)_{n \geq 0}$ is locally finite ($s^n(w) = 0$ if $|w| < n$ (Cauchy product)). Thus for a proper series $s \in \mathbb{K}\langle\langle A^* \rangle\rangle$ we can define s^*.

Definition. A subset of $\mathbb{K}\langle\langle A^* \rangle\rangle$ is called rationally closed if it closed under

- left and right externe multiplication
- point wise sum
- Cauchy product
- under *-operator when it is defined

Definition. A polynomial is a series of finite support
The set of of polynomials over A^* is denoted by $\mathbb{K}\langle A^* \rangle \subseteq \mathbb{K}\langle\langle A^* \rangle\rangle$
The set of rational series is the rational closure of $\mathbb{K}\langle A^* \rangle$