Weighted Two-Way Transducers

Fan Feng and Andreas Maletti

Leipzig University

October 28, 2022

Introduction

Two-way transducers received a lot of interest recently. They are a natural generalization of **finite-state transducers** extended with the facility to move left and right along the input string like a Turing machine but without the Turing machine's facility to write onto that tape. They can similarly be considered as an extension of **two-way automata** with the facility to generate output. Recently, two-way automata were generalized to **weighted two-way automata**, in which transitions carry weights in order to model success probabilities, multiplicities, or other quantitative aspects. In this contribution we introduce **weighted two-way transducers** that similarly extend two-way transducers with the facility to charge weights on each transition.

Definition 1.1

A **(commutative) semiring** is an algebraic structure $(S, +, \cdot, 0, 1)$ such that $(S, +, 0)$ and $(S, \cdot, 1)$ are commutative monoids, $s \cdot 0 = 0$ for every $s \in S$, and the distributive law $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$ holds for all $x, y, z \in S$.

Definition 1.2

A **complete semiring** $(S, +, \cdot, 0, 1, \sum)$ is a semiring $(S, +\cdot, 0, 1)$ in which for any set I and mapping $\alpha : I \to S$ the infinite sum $\sum \alpha$ is defined and the satisfies the following three **axioms**:

- If $I = \{i, j\}$ has two different elements, then $\sum \alpha = \alpha(i) + \alpha(j)$.
- $\sum \alpha = \sum \beta$ for every partition $\prod : J \to \mathcal{P}(I)$ of I, where $\beta : J \to S$ is such that $\beta(j) = \sum \alpha|_{\prod(j)}$ for all $j \in J$.
- $s \cdot (\sum \alpha) = \sum \alpha_s$ for every $s \in S$, where $\alpha_s : I \to S$ is given by $\alpha_s(i) = s \cdot \alpha(i)$ for every $i \in I$.

To avoid the explicit definition of α we often write $\sum_{i \in I} a_i$ instead of $\sum \alpha$. It follows from the last axiom that $\sum_{i \in I} 0 = 0$.

Example 1.3

$(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1, \sum)$ with the usual addition and multiplication extended to ∞ $(i.e., 0 \cdot \infty = 0)$ is a **complete semiring**, where for every $\alpha : I \to S$ we have $\sum \alpha = \sum_{i \in supp(\alpha)} \alpha(i)$ if $supp(\alpha)$ is finite and $\sum \alpha = \infty$ otherwise.

$(S, +, \cdot, 0, 1, \sum)$: **complete semiring**

$A_{\vdash\dashv} = A \cup \{\vdash, \dashv\}$: **input alphabet** A with $A \cap \{\vdash, \dashv\} = \emptyset$.
($\vdash$ and $\dashv$ are used as **left- and right-end marker**)

$B^{\leq 1} = B \cup \{\varepsilon\}$: **output alphabet**.
(ε is used as **empty output**)

Definition 2.1

A **weighted two-way finite-state transducer** (for short: w2fst) is a tuple $\mathfrak{A} = (Q^{\rightarrow}, Q^{\leftarrow}, A, B, T, I, F)$, in which $Q^{\rightarrow}$ and $Q^{\leftarrow}$ are disjoint finite sets of **forward** and **backward states**, resp., A and B are alphabets of **input** and **output symbols**, resp., $I, F \colon Q^{\rightarrow} \to S$ are **initial** and **final weights**, resp., and $T \colon Q \times A_{\vdash\dashv} \times B^{\leq 1} \times Q \to S$ assigns **weights to transitions**, where $Q = Q^{\rightarrow} \cup Q^{\leftarrow}$. It is **deterministic** (for short: wd2fst) if $|supp(I)| \leqslant 1$ and for every $q \in Q$ and input symbol $a \in A_{\vdash\dashv}$ there exists at most one pair $(b, q') \in B^{\leq 1} \times Q$ with $(q, a, b, q') \in supp(T)$. Similarly, it is **co-deterministic** if $|supp(F)| \leq 1$ and for every $q' \in Q$ and input symbol $a \in A_{\vdash\dashv}$ there exists at most one pair $(b, q) \in B^{\leq 1} \times Q$ such that $(q, a, b, q') \in supp(T)$. The w2fst is **reversible** (for short: wr2fst) if it is deterministic as well as co-deterministic.

Example 2.2

We utilize the complete semiring $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1, \sum)$ and consider the w2fst $\mathfrak{A} = (Q^{\rightarrow}, Q^{\leftarrow}, A, A, T, I, F)$ with input and output alphabet $A = \{a, b\}$, forward states $Q^{\rightarrow} = \{q_0, q_1, q_2, q_5, q_\mathsf{f}\}$ and backward states $Q^{\leftarrow} = \{q_3, q_4\}$, and the following nonzero-weighted transitions, nonzero initial and final weights $I(q_0) = 1 = F(q_\mathsf{f})$.

$$T(q_1, b, \varepsilon, q_2) = 2$$
$$T(q_0, \vdash, \varepsilon, q_1) = T(q_1, a, a, q_1) = T(q_1, \dashv, \varepsilon, q_\mathsf{f}) = T(q_2, a, \varepsilon, q_3) = 1$$
$$T(q_3, b, \varepsilon, q_4) = T(q_4, a, b, q_4) = T(q_4, b, \varepsilon, q_5) = T(q_4, \vdash, \varepsilon, q_5) = 1$$
$$T(q_5, a, \varepsilon, q_5) = T(q_5, b, \varepsilon, q_1) = 1$$

The w2fst $\mathfrak{A}$ is a wr2fst. □

Transition $\tau = (q, a, b, q') \in supp(T)$

$q, q' \in Q^{\rightarrow} \Rightarrow$ the head moves one step to the **right**

$q, q' \in Q^{\leftarrow} \Rightarrow$ the head moves one step to the **left**

$q \in Q^{\rightarrow}, q' \in Q^{\leftarrow} \Rightarrow$ the head **does not move**

$q \in Q^{\leftarrow}, q' \in Q^{\rightarrow} \Rightarrow$ the head **does not move**

For $w2fst\mathfrak{A} = (Q^{\rightarrow}, Q^{\leftarrow}, A, B, T, I, F)$ with $Q = Q^{\rightarrow} \cup Q^{\leftarrow}$

A **configuration** on w: $(\ell, q, r) \in A^*_{\vdash\dashv} \times Q \times A^*_{\vdash\dashv}$ with $\ell r = \vdash w \dashv$
The set of all configurations on w: $\mathcal{C}_{\mathfrak{A}}(w)$
Transition $\tau = (q, a, b, q') \in supp(T)$

For every configuration $(\ell, q, ar) \in \mathcal{C}_{\mathfrak{A}}(w)$:
If $q \in Q^{\rightarrow}, q' \in Q^{\rightarrow}$, then we let $(\ell, q, ar) \xrightarrow{\tau}_{\mathfrak{A}} (\ell a, q', r)$;
if $q \in Q^{\rightarrow}, q' \in Q^{\leftarrow}$, then we let $(\ell, q, ar) \xrightarrow{\tau}_{\mathfrak{A}} (\ell, q', ar)$.

For every configuration $(\ell a, q, r) \in \mathcal{C}_{\mathfrak{A}}(w)$:
If $q \in Q^{\leftarrow}, q' \in Q^{\leftarrow}$, then we let $(\ell a, q, r) \xrightarrow{\tau}_{\mathfrak{A}} (\ell, q', ar)$;
If $q \in Q^{\leftarrow}, q' \in Q^{\rightarrow}$, then we let $(\ell a, q, r) \xrightarrow{\tau}_{\mathfrak{A}} (\ell a, q', r)$.

Configurations $C_0, \cdots, C_n \in \mathcal{C}_{\mathfrak{A}}(w)$

Transitions $\tau_1, \cdots, \tau_n \in supp(T) \Rightarrow C_0 \xrightarrow{\tau_1}_{\mathfrak{A}} C_1 \xrightarrow{\tau_2}_{\mathfrak{A}} \cdots \xrightarrow{\tau_n}_{\mathfrak{A}} C_n$

For $C_0 \xrightarrow{\rho}_{\mathfrak{A}} C_n$ where $\rho = (\tau_1, \cdots, \tau_n)$ is a **run** from C_0 to C_n, we denote the set of all such runs by $Run_{\mathfrak{A}}(C_0, C_n)$.

We define the **weight** $wt_{\mathfrak{A}}(\rho) = \prod_{i=1}^{n} T(\tau_i)$ and the **output** $out(\rho) = \pi_3(\tau_1) \cdots \pi_3(\tau_n)$ of the run ρ, where π_3 is the projection to the third component. The w2fst $\mathfrak{A}$ computes the mapping $\| \mathfrak{A} \| \colon A^* \times B^* \to S$ given for every $w \in A^*$ and $v \in B^*$ by

$$\| \mathfrak{A} \| (w, v) = \sum_{\substack{q_0, q_{\mathsf{f}} \in Q^{\rightarrow} \\ C = (\varepsilon, q_0, \vdash w \dashv),\, C' = (\vdash w \dashv, q_{\mathsf{f}}, \varepsilon) \\ \rho \in Run_{\mathfrak{A}}(C, C'),\, out(\rho) = v}} I(q_0) \cdot wt_{\mathfrak{A}}(\rho) \cdot F(q_{\mathsf{f}}) \ .$$

A mapping $f\colon A^* \times B^* \to S$ is *w2fst-**computable*** (resp. *wd2fst-computable* and *wr2fst-computable*) if there exists a w2fst $\mathfrak{A}$ (resp. wd2fst $\mathfrak{A}$ and wr2fst $\mathfrak{A}$) that computes f (i.e., $f = ||\mathfrak{A}||$). If $S = \mathbb{B}$ is the BOOLEAN semiring

$$\mathbb{B} = (\{0,1\}, \vee, \wedge, 0, 1, \bigvee) \enspace ,$$

then we can identify mappings $f\colon A^* \times B^* \to \mathbb{B}$ with relations $supp(f) \subseteq A^* \times B^*$ and obtain the ***two-way definable*** (resp. *deterministic two-way definable* and *reversible two-way definable*) relations.

Example 2.3

Recall the wr2fst $\mathfrak{A}$ of Example 2.2. Let us illustrate a run without the transitions on top of $\to_{\mathfrak{A}}$ and configurations (ℓ, q, r) written simply as $\ell\, q\, r$.

$$\begin{aligned} & q_0 \vdash aaba\dashv \to_{\mathfrak{A}} \vdash q_1\, aaba\dashv \to_{\mathfrak{A}} \vdash a\, q_1\, aba\dashv \to_{\mathfrak{A}} \vdash aa\, q_1\, ba\dashv \\ \to_{\mathfrak{A}} & \vdash aab\, q_2\, a\dashv \to_{\mathfrak{A}} \vdash aab\, q_3\, a\dashv \to_{\mathfrak{A}} \vdash aa\, q_4\, ba\dashv \to_{\mathfrak{A}} \vdash a\, q_4\, aba\dashv \\ \to_{\mathfrak{A}} & \vdash q_4\, aaba\dashv \to_{\mathfrak{A}} \vdash q_5\, aaba\dashv \to_{\mathfrak{A}} \vdash a\, q_5\, aba\dashv \to_{\mathfrak{A}} \vdash aa\, q_5\, ba\dashv \\ \to_{\mathfrak{A}} & \vdash aab\, q_1\, a\dashv \to_{\mathfrak{A}} \vdash aaba\, q_1 \dashv \to_{\mathfrak{A}} \vdash aaba\dashv q_{\mathsf{f}} \end{aligned}$$

The output of this run is $aabba$ and its weight is 2 because each utilized transition has weight 1 except for the fourth transition, which has weight 2. The nonzero entries in $||\mathfrak{A}||$. are $||\mathfrak{A}||(a^{i_1}ba^{i_2}b\cdots a^{i_n}ba^{i_{n+1}},\, a^{i_1}b^{i_1}a^{i_2}b^{i_2}\cdots a^{i_n}b^{i_n}a^{i_{n+1}}) = 2^n$ for every $n \in \mathbb{N}$ and $i_1, \cdots, i_{n+1} \in \mathbb{N}$.

Theorem 3.1

For every w2fst-computable mapping $f\colon A^* \times B^* \to S$ the support $supp(f)$ is two-way definable. This extends to the deterministic and reversible case.

Proof(1/2)

Let $\mathfrak{A} = (Q^{\to}, Q^{\leftarrow}, A, B, T, I, F)$ be a w2fst computing f. Moreover, let $Q = Q^{\to} \cup Q^{\leftarrow}$ and W be the finite set $W = \{0\} \cup ran(T) \cup ran(I) \cup ran(F) \subseteq S$. We consider the monoid $(\mathbb{N}^W, +, \mathbf{0})$ with point-wise addition and the homomorphism $h \colon \mathbb{N}^W \to S$ into the multiplicative monoid $(S, \cdot, 1)$ of S given by

$$h(\varphi) = \prod_{s \in W} s^{\varphi(s)}$$

for every $\varphi \colon W \to \mathbb{N}$, where we assume that $0^0 = 1$. By DICKSON's lemma the set $\min h^{-1}(0)$ is finite, where the partial order, for which the minimal elements are determined, is the standard pointwise order on $\mathbb{N}^W$. Consequently, there exists $u \in \mathbb{N}$ such that $\min h^{-1}(0) \subseteq \{0, \ldots, u\}^W = U$. We define the mapping $\oplus \colon U^2 \to U$ by $(\varphi \oplus \varphi')(s) = \min(\varphi(s) + \varphi'(s), u)$

Proof(2/2)

for every $\varphi, \varphi' \in U$ and $s \in W$. Moreover, for every $s \in W$ we let $\overline{s} \in U$ be such that $\overline{s}(s) = 1$ and $\overline{s}(s') = 0$ for all $s' \in S \setminus \{s\}$. Let $V = U \setminus h^{-1}(0)$. We construct the equivalent w2fst $\mathfrak{A}' = (Q^{\rightarrow} \times V, Q^{\leftarrow} \times V, A, B, T', I', F')$ with $P = Q \times V$ and

$$T'(\langle q, \varphi \rangle, a, b, \langle q', \varphi' \rangle) = \begin{cases} T(q, a, b, q') & \text{if } \varphi' = \varphi \oplus \overline{T(q, a, b, q')} \\ 0 & \text{otherwise} \end{cases}$$

$$I'(\langle q'', \varphi'' \rangle) = \begin{cases} I(q'') & \text{if } \varphi'' = \overline{I(q'')} \\ 0 & \text{otherwise} \end{cases}$$

$$F'(\langle q'', \varphi'' \rangle) = F(q'')$$

for every $\langle q, \varphi \rangle, \langle q', \varphi' \rangle \in P$, $\langle q'', \varphi'' \rangle \in Q^{\rightarrow} \times V$, $a \in A_{\vdash\dashv}$, and $b \in B^{\leq 1}$. □

Theorem 4.1

*The **sum** $(f+g)\colon A^* \times B^* \to S$ of two w2fst-computable mappings $f, g\colon A^* \times B^* \to S$ is again w2fst-computable.*

Proof.

Let $\mathfrak{A} = (Q^{\rightarrow}, Q^{\leftarrow}, A, B, T, I, F)$ and $\mathfrak{A}' = (P^{\rightarrow}, P^{\leftarrow}, A, B, T', I', F')$ be w2fst that compute f and g, respectively. Without loss of generality, suppose that $(Q^{\rightarrow} \cup Q^{\leftarrow}) \cap (P^{\rightarrow} \cup P^{\leftarrow}) = \emptyset$. The w2fst $\mathfrak{A} + \mathfrak{A}'$

$$\mathfrak{A} + \mathfrak{A}' = (Q^{\rightarrow} \cup P^{\rightarrow}, Q^{\leftarrow} \cup P^{\leftarrow}, A, B, T \cup T', I \cup I', F \cup F')$$

obviously computes $f + g$. □

Definition 4.2

For every set X and all mappings $f, g \colon X \to S$, their **point-wise product** $(f \cdot g) \colon X \to S$ is defined by $(f \cdot g)(x) = f(x) \cdot g(x)$ for every $x \in X$.

Theorem 4.3

If w2fst emptiness is decidable (for the semiring S), then the class of w2fst-computable mappings is not closed under products.

Proof.

We consider an instance (h, h') of the undecidable Post Correspondence Problem (PCP), so $h, h' \colon A \to B^*$ are mappings for some alphabets A and B that extend to homomorphisms $h, h' \colon A^* \to B^*$. This instance is solvable if there exists $w \in A^* \backslash \{\varepsilon\}$ such that $h(w) = h'(w)$.
Using the fact that PCP is undecidable, we achieve the result. □

Definition 4.4

Let $f\colon A^* \times B^* \to S$ and $g\colon B^* \times C^* \to S$ be two mappings. Their **composition** $(f\,;g)\colon A^* \times C^* \to S$ is defined for every $w \in A^*$ and $u \in C^*$ by $(f\,;g)(w,u) = \sum_{v \in B^*} f(w,v) \cdot g(v,u)$.

We note that for wd2fst-computable mappings f and g the sum in the definition of $f\,;g$ is always finite.

Definition 4.5

A wr2fst $\mathfrak{A} = (Q^{\rightarrow}, Q^{\leftarrow}, A, B, T, I, F)$ has a single initial state $q_0 \in supp(I)$ and a single final state $q_{\mathsf{f}} \in supp(F)$. It is **normalized** if
(i) $(q_0, \vdash, b, q) \in supp(T)$ implies $b = \varepsilon$ and
(ii) $(q, \dashv, b, q_{\mathsf{f}}) \in supp(T)$ implies $b = \varepsilon$.

In other words, the initial and final transition produce no output. Without loss of generality, we can assume that a wr2fst is normalized.

For a normalized wr2fst $\mathfrak{A} = (Q^{\rightarrow}, Q^{\leftarrow}, A, B, T, I, F)$ we can construct the wr2fst $S(\mathfrak{A}) = \big(Q^{\rightarrow} \cup P^{\rightarrow}, Q^{\leftarrow} \cup P^{\leftarrow} \cup \{\bot\}, A, B_{\vdash\dashv}, T', I', F'\big)$ which separates the weight and the output of $\mathfrak{A}$.

Phase I: $S(\mathfrak{A})$ charges only the **weights**, without producing any output;
Phase II: $S(\mathfrak{A})$ produces the **output**, without charging the weight.

Theorem 4.6

The composition $(f\,;g)\colon A^* \times C^* \to S$ *of two wr2fst-computable mappings* $f\colon A^* \times B^* \to S$ *and* $g\colon B^* \times C^* \to S$ *is again wr2fst-computable.*

Proof(1/2)

Let $\mathfrak{A}$ and $\mathfrak{A}' = (P^{\rightarrow}, P^{\leftarrow}, B, C, T', I', F')$ be wr2fst that compute the mappings f and g, respectively. Moreover, let $S(\mathfrak{A}) = (Q^{\rightarrow}, Q^{\leftarrow}, A, B, T, I, F)$, $Q = Q^{\rightarrow} \cup Q^{\leftarrow}$ and $P = P^{\rightarrow} \cup P^{\leftarrow}$.

The main idea:

We can construct the wr2fst $\mathfrak{B} = \Big((Q^{\rightarrow} \times P^{\rightarrow}) \cup (Q^{\leftarrow} \times P^{\leftarrow}), (Q^{\rightarrow} \times P^{\leftarrow}) \cup (Q^{\leftarrow} \times P^{\rightarrow}), A, C, T'', I'', F''\Big)$ with initial and final weights

$I''(\langle q, p\rangle) = I(q) \cdot I'(p)$ and $F''(\langle q, p\rangle) = F(q) \cdot F'(p)$ for every $q \in Q$ and $p \in P$, and for every $q, q' \in Q$, $p, p' \in P$, and $a \in A_{\vdash\dashv}$ we let

Proof(2/2)

$$T''\Big(\langle q,p\rangle, a, c, \langle q',p'\rangle\Big) = \begin{cases} \sum_{b\in B^{\leq 1}} T(q,a,b,q')\cdot T'(p,b,c,p') & \text{if } p,p'\in P^{\rightarrow} \\ \sum_{b\in B^{\leq 1}} T(q',a,b,q)\cdot T'(p,b,c,p') & \text{if } p,p'\in P^{\leftarrow} \\ \sum_{b\in B, q''\in Q} T(q,a,b,q'')\cdot T'(p,b,c,p') & \text{if } p\in P^{\rightarrow}, p'\in P^{\leftarrow}, q=q' \\ \sum_{b\in B, q''\in Q} T(q'',a,b,q)\cdot T'(p,b,c,p') & \text{if } p\in P^{\leftarrow}, p'\in P^{\rightarrow}, q=q' \\ 0 & \text{otherwise} \end{cases}$$

Theorem 4.7

For every wd2fst there exists an equivalent wr2fst.

Corollary 4.8

The classes of wd2fst-computable and wr2fst-computable mappings coincide and they are both closed under composition.