Automata & transducers with weights & data storage

after papers by Jonathan Goldstine*

talk by Tobias Denkinger

Freitagsseminar 2018-06-15

* [Goldstine 79, 80]

Overview

1. Why AFA-theory almost died (according to Goldstine)

2. Stripping down PDAs

3. Monoids and rational sets

4. Automata with data storage

5. Weighted transducers for free (work in progress)

Why AFA-theory almost died

<u>Ex.</u> $L = \{a^n b^n \mid n \in \mathbb{N}\}$

Why AFA-theory almost died

<u>Ex.</u> $L = \{a^n b^n \mid n \in \mathbb{N}\}$

- $G = (\{S\}, \{a,b\}, S, \{S \to aSb, S \to \varepsilon\})$

Why AFA-theory almost died

Ex. $L = \{a^n b^n \mid n \in \mathbb{N}\}$

- $G = (\{S\}, \{a,b\}, S, \{S \to aSb, S \to \varepsilon\})$

- $A = (\{q_0, q_1, q_2\}, \{a,b\}, \{Z_0, Z\}, \delta, q_0, Z_0, \{q_2\})$ where

Why AFA-theory almost died

<u>Ex.</u> $L = \{a^n b^n \mid n \in \mathbb{N}\}$

- $G = (\{S\}, \{a,b\}, S, \{S \to aSb, S \to \varepsilon\})$
- $A = (\{q_0, q_1, q_2\}, \{a, b\}, \{Z_0, Z\}, \delta, q_0, Z_0, \{q_2\})$ where

$$
\begin{aligned}
\delta: (q_0, \varepsilon, Z_0) &\mapsto \{(q_2, \varepsilon)\} \\
(q_0, a, Z_0) &\mapsto \{(q_0, Z_0 Z), (q_1, Z_0 Z)\} \\
(q_0, a, Z) &\mapsto \{(q_0, ZZ), (q_1, ZZ)\} \\
(q_1, b, Z) &\mapsto \{(q_1, \varepsilon)\} \\
(q_1, \varepsilon, Z_0) &\mapsto \{(q_2, \varepsilon)\} \\
(\_, \_, \_) &\mapsto \emptyset
\end{aligned}
$$

Why AFA-theory almost died

Ex. $L = \{a^n b^n \mid n \in \mathbb{N}\}$

- $G = (\{S\}, \{a,b\}, S, \{S \to aSb, S \to \varepsilon\})$
- $A = (\{q_0, q_1, q_2\}, \{a, b\}, \{Z_0, Z\}, \delta, q_0, Z_0, \{q_2\})$ where

$$
\begin{aligned}
\delta: (q_0, \varepsilon, Z_0) &\mapsto \{(q_2, \varepsilon)\} \\
(q_0, a, Z_0) &\mapsto \{(q_0, Z_0 Z), (q_1, Z_0 Z)\} \\
(q_0, a, Z) &\mapsto \{(q_0, ZZ), (q_1, ZZ)\} \\
(q_1, b, Z) &\mapsto \{(q_1, \varepsilon)\} \\
(q_1, \varepsilon, Z_0) &\mapsto \{(q_2, \varepsilon)\} \\
(\_, \_, \_) &\mapsto \emptyset
\end{aligned}
$$

"... these automata are awkward to write down or manipulate largely because the formalism is bad."

[Goldstine 80]

Stripping down PDAs

"... there is nothing trivial about matters of notation, and I do not apologize for stressing the importance of good notation."

[Goldstine 80]

Stripping down PDAs

"... there is nothing trivial about matters of notation, and I do not apologize for stressing the importance of good notation."

[Goldstine 80]

original definition:

$A = (\{q_0, q_1, q_2\}, \{a, b\}, \{Z_0, Z\}, \delta, q_0, Z_0, \{q_2\})$ where

$$
\begin{aligned}
\delta: (q_0, \varepsilon, Z_0) &\mapsto \{(q_2, \varepsilon)\} \\
(q_0, a, Z_0) &\mapsto \{(q_0, Z_0 Z), (q_1, Z_0 Z)\} \\
(q_0, a, Z) &\mapsto \{(q_0, ZZ), (q_1, ZZ)\} \\
(q_1, b, Z) &\mapsto \{(q_1, \varepsilon)\} \\
(q_1, \varepsilon, Z_0) &\mapsto \{(q_2, \varepsilon)\} \\
(\_, \_, \_) &\mapsto \emptyset
\end{aligned}
$$

Stripping down PDAs

"... there is nothing trivial about matters of notation, and I do not apologize for stressing the importance of good notation."

[Goldstine 80]

original definition:

$A = (\{q_0, q_1, q_2\}, \{a, b\}, \{Z_0, Z\}, \delta, q_0, Z_0, \{q_2\})$ where

$\delta: (q_0, \varepsilon, Z_0) \mapsto \{(q_2, \varepsilon)\}$

$(q_0, a, Z_0) \mapsto \{(q_0, Z_0Z), (q_1, Z_0Z)\}$

$(q_0, a, Z) \mapsto \{(q_0, ZZ), (q_1, ZZ)\}$

$(q_1, b, Z) \mapsto \{(q_1, \varepsilon)\}$

$(q_1, \varepsilon, Z_0) \mapsto \{(q_2, \varepsilon)\}$

$(\_, \_, \_) \mapsto \emptyset$

$\delta(p, x, y) \ni (q, z) \ldots$

Stripping down PDAs

"... there is nothing trivial about matters of notation, and I do not apologize for stressing the importance of good notation."
[Goldstine 80]

original definition:

$A = (\{q_0, q_1, q_2\}, \{a, b\}, \{Z_0, Z\}, \delta, q_0, Z_0, \{q_2\})$ where

$\delta: (q_0, \varepsilon, Z_0) \mapsto \{(q_2, \varepsilon)\}$
$(q_0, a, Z_0) \mapsto \{(q_0, Z_0 Z), (q_1, Z_0 Z)\}$
$(q_0, a, Z) \mapsto \{(q_0, ZZ), (q_1, ZZ)\}$
$(q_1, b, Z) \mapsto \{(q_1, \varepsilon)\}$
$(q_1, \varepsilon, Z_0) \mapsto \{(q_2, \varepsilon)\}$
$(\_, \_, \_) \mapsto \emptyset$

$\delta(p, x, y) \ni (q, z) \ldots$

make it a tuple

$(p, x, y, z, q) \ldots$

Stripping down PDAs

"... there is nothing trivial about matters of notation, and I do not apologize for stressing the importance of good notation."

[Goldstine 80]

original definition:

$A = (\{q_0, q_1, q_2\}, \{a, b\}, \{z_0, z\}, \delta, q_0, z_0, \{q_2\})$ where

$\delta: (q_0, \varepsilon, z_0) \mapsto \{(q_2, \varepsilon)\}$

$(q_0, a, z_0) \mapsto \{(q_0, z_0 z), (q_1, z_0 z)\}$

$(q_0, a, z) \mapsto \{(q_0, zz), (q_1, zz)\}$

$(q_1, b, z) \mapsto \{(q_1, \varepsilon)\}$

$(q_1, \varepsilon, z_0) \mapsto \{(q_2, \varepsilon)\}$

$(\_, \_, \_) \mapsto \emptyset$

$\delta(p, x, y) \ni (q, z) \ldots$

make it a tuple

$(p, x, y, z, q) \ldots$

$\mathrm{pop}_y = \mathrm{push}_y^{-1}$

$(p, x, y^{-1} z, q) \ldots$

Stripping down PDAs

"... there is nothing trivial about matters of notation, and I do not apologize for stressing the importance of good notation."

[Goldstine 80]

original definition:

$A = (\{q_0, q_1, q_2\}, \{a, b\}, \{z_0, z\}, \delta, q_0, z_0, \{q_2\})$ where

$$\begin{aligned}
\delta: (q_0, \varepsilon, z_0) &\mapsto \{(q_2, \varepsilon)\}\\
(q_0, a, z_0) &\mapsto \{(q_0, z_0 z), (q_1, z_0 z)\}\\
(q_0, a, z) &\mapsto \{(q_0, zz), (q_1, zz)\}\\
(q_1, b, z) &\mapsto \{(q_1, \varepsilon)\}\\
(q_1, \varepsilon, z_0) &\mapsto \{(q_2, \varepsilon)\}\\
(_, _, _) &\mapsto \emptyset
\end{aligned}$$

$\delta(p, x, y) \ni (q, z) \ldots$

make it a tuple

$(p, x, y, z, q) \ldots$

$pop_y = push_y^{-1}$

$(p, x, y^{-1}z, q) \ldots$

remove states

$\langle y^{-1}z, x \rangle$

Stripping down PDAs

"... there is nothing trivial about matters of notation, and I do not apologize for stressing the importance of good notation."

[Goldstine 80]

original definition:

$A = (\{q_0, q_1, q_2\}, \{a, b\}, \{z_0, z\}, \delta, q_0, z_0, \{q_2\})$ where

$\delta: (q_0, \varepsilon, z_0) \mapsto \{(q_2, \varepsilon)\}$
$(q_0, a, z_0) \mapsto \{(q_0, z_0 z), (q_1, z_0 z)\}$
$(q_0, a, z) \mapsto \{(q_0, zz), (q_1, zz)\}$
$(q_1, b, z) \mapsto \{(q_1, \varepsilon)\}$
$(q_1, \varepsilon, z_0) \mapsto \{(q_2, \varepsilon)\}$
$(\_, \_, \_) \mapsto \emptyset$

regular expression for L:

$A' = \langle z, a\rangle^* \circ \langle z^{-1}, b\rangle^* \circ \langle z_0^{-1}, \varepsilon\rangle$

$\delta(p, x, y) \ni (q, z) \ldots$

$\Big\downarrow$ make it a tuple

$(p, x, y, z, q) \ldots$

$\Big\downarrow$ $pop_y = push_y^{-1}$

$(p, x, y^{-1}z, q) \ldots$

$\Big\downarrow$ remove states

$\langle y^{-1}z, x\rangle$

Monoids and rational sets

- monoid $(M, \odot, \mathbb{1})$ $\quad \odot : M \times M \to M$ assoc., $\mathbb{1}$ neutral

Monoids and rational sets

- monoid $(M, \odot, \mathbb{1})$ $\odot : M \times M \to M$ assoc., $\mathbb{1}$ neutral
- free monoid $(\Sigma^*, \cdot, \varepsilon)$ $\cdot$ concatenation

Monoids and rational sets

- monoid $(M, \odot, \mathbb{1})$ $\quad \odot : M \times M \to M$ assoc., $\mathbb{1}$ neutral
- free monoid $(\Sigma^*, \circ, \varepsilon)$ $\quad \circ$ concatenation
- product monoid $M_1 \times M_2 = (M_1 \times M_2, \odot_1 \times \odot_2, (\mathbb{1}_1, \mathbb{1}_2))$

Monoids and rational sets

- monoid $(M, \odot, \mathbb{1})$ $\quad \odot : M \times M \to M$ assoc., $\mathbb{1}$ neutral
- free monoid $(\Sigma^*, \cdot, \varepsilon)$ $\quad \cdot$ concatenation
- product monoid $M_1 \times M_2 = (M_1 \times M_2, \odot_1 \times \odot_2, (\mathbb{1}_1, \mathbb{1}_2))$
- rational subsets of a monoid M, $Rat(M)$: smallest set R s.t.
  - R contains all finite subsets of M
  - R contains all unions and products of its elements
  - R contains all Kleene-$*$s of its elements

Monoids and rational sets

- monoid $(M, \odot, \mathbb{1})$ $\quad \odot : M \times M \to M$ assoc., $\mathbb{1}$ neutral
- free monoid $(\Sigma^*, \cdot, \varepsilon)$ $\quad \cdot$ concatenation
- product monoid $M_1 \times M_2 = (M_1 \times M_2, \odot_1 \times \odot_2, (\mathbb{1}_1, \mathbb{1}_2))$
- rational subsets of a monoid M, $Rat(M)$: smallest set R s.t.
  - R contains all finite subsets of M
  - R contains all unions and products of its elements
  - R contains all Kleene-$*$s of its elements
- rational transducers $T \in Rat(\langle \Sigma^* \times \Delta^* \rangle^*)$

Automata with data storage

- data storage mechanism (C, R, In, Out) where
  - C is a set (configurations)
  - $R \subseteq \mathcal{P}(C \times C)$ (instructions)
  - $In \subseteq C$, $Out \subseteq C$ (input & output configurations)

Automata with data storage

- data storage mechanism (C, R, In, Out) where
  - C is a set (configurations)
  - $R \subseteq \mathcal{P}(C \times C)$ (instructions)
  - $In \subseteq C$, $Out \subseteq C$ (input & output configurations)
- automaton with data storage $\mathcal{A} \in Rat(\langle R^* \times \Sigma^* \rangle^*)$

Automata with data storage

- data storage mechanism (C, R, In, Out) where
  - C is a set (configurations)
  - $R \subseteq \mathcal{P}(C \times C)$ (instructions)
  - $In \subseteq C$, $Out \subseteq C$ (input & output configurations)

- automaton with data storage $\mathcal{A} \in Rat(\langle R^* \times \Sigma^* \rangle^*)$
  - $L(\mathcal{A}) = \{ u_1 \cdots u_k \mid \langle r_1, u_1 \rangle \cdots \langle r_k, u_k \rangle \in \mathcal{A} :$
    $(r_1; \cdots ; r_k) \cap In \times Out \neq \emptyset \}$

Automata with data storage

- data storage mechanism (C, R, In, Out) where
  - C is a set (configurations)
  - $R \subseteq \mathcal{P}(C \times C)$ (instructions)
  - $In \subseteq C$, $Out \subseteq C$ (input & output configurations)

- automaton with data storage $\mathcal{A} \in Rat(\langle R^* \times \Sigma^* \rangle^*)$
  - $L(\mathcal{A}) = \{ u_1 \cdots u_k \mid \langle r_1, u_1 \rangle \cdots \langle r_k, u_k \rangle \in \mathcal{A} :$
    $(r_1; \ldots; r_k) \cap In \times Out \neq \emptyset \}$
    $= \{ w \mid \langle r, w \rangle \in \tau(\mathcal{A}) : r \cap In \times Out \neq \emptyset \}$

Automata with data storage

- data storage mechanism (C, R, In, Out) where
  - C is a set (configurations)
  - $R \subseteq \mathcal{P}(C \times C)$ (instructions)
  - $In \subseteq C$, $Out \subseteq C$ (input & output configurations)

- automaton with data storage $\mathcal{A} \in Rat(\langle R^* \times \Sigma^* \rangle^*)$
  - $L(\mathcal{A}) = \{ u_1 \cdots u_k \mid \langle r_1, u_1 \rangle \cdots \langle r_k, u_k \rangle \in \mathcal{A} :$
    $(r_1; \cdots; r_k) \cap In \times Out \neq \emptyset \}$

    $= \{ w \mid \langle r, w \rangle \in \tau(\mathcal{A}) : r \cap In \times Out \neq \emptyset \}$

    $= \tau(\mathcal{A})(B_{\mathcal{A}})$ where

Automata with data storage

- data storage mechanism (C, R, In, Out) where
  - C is a set (configurations)
  - $R \subseteq \mathcal{P}(C \times C)$ (instructions)
  - $In \subseteq C$, $Out \subseteq C$ (input & output configurations)

- automaton with data storage $\mathcal{A} \in Rat(\langle R^* \times \Sigma^* \rangle^*)$
  - $L(\mathcal{A}) = \{ u_1 \cdots u_k \mid \langle r_1, u_1 \rangle \cdots \langle r_k, u_k \rangle \in \mathcal{A} :$

$$(r_1; \ldots; r_k) \cap In \times Out \neq \emptyset \}$$

$$= \{ w \mid \langle r, w \rangle \in \tau(\mathcal{A}) : r \cap In \times Out \neq \emptyset \}$$

$$= \tau(\mathcal{A})(B_{\mathcal{A}}) \qquad \text{where}$$

$$B_{\mathcal{A}} = \{ r_1 \cdots r_k \mid k \in \mathbb{N}, r_1, \ldots, r_k \in R, \; (r_1; \ldots; r_k) \cap In \times Out \neq \emptyset \}$$

"storage behaviour"

A definition of weighted transducers

(using way too much algebra)

storage in out weight

$$T = \langle \mathrm{inc}, a, \alpha, 2 \rangle^* \quad \langle \mathrm{inc}, a, \gamma, 4 \rangle^* \quad \langle \mathrm{dec}, b, \beta, 1 \rangle^*$$

A definition of weighted transducers

(using way too much algebra)

storage in out weight

$$T = \langle \mathrm{inc}, a, \alpha, 2\rangle^* \quad \langle \mathrm{inc}, a, \gamma, 4\rangle^* \quad \langle \mathrm{dec}, b, \beta, 1\rangle^*$$

$$= (\mathrm{inc} \times a \times \alpha \times 2)^* (\mathrm{inc} \times a \times \gamma \times 4)^* (\mathrm{dec} \times b \times \beta \times 1)^*$$

A definition of weighted transducers

(using way too much algebra)

$$T = \langle \overset{\text{storage}}{inc}, \overset{\text{in}}{a}, \overset{\text{out}}{\alpha}, \overset{\text{weight}}{2} \rangle^* \; \langle inc, a, \gamma, 4 \rangle^* \; \langle dec, b, \beta, 1 \rangle^*$$

$$= (inc \times a \times \alpha \times 2)^* (inc \times a \times \gamma \times 4)^* (dec \times b \times \beta \times 1)^*$$

$$= \sum_{x,y,z} (inc \times a \times \alpha \times 2)^x (inc \times a \times \gamma \times 4)^y (dec \times b \times \beta \times 1)^z$$

A definition of weighted transducers

(using way too much algebra)

storage in out weight

$$T = \langle inc, a, \alpha, 2\rangle^* \quad \langle inc, a, \gamma, 4\rangle^* \quad \langle dec, b, \beta, 1\rangle^*$$

$$= (inc \times a \times \alpha \times 2)^* (inc \times a \times \gamma \times 4)^* (dec \times b \times \beta \times 1)^*$$

$$= \sum_{x,y,z} (inc \times a \times \alpha \times 2)^x (inc \times a \times \gamma \times 4)^y (dec \times b \times \beta \times 1)^z$$

$$\tau(T) = \sum_{x,y,z} (inc^x; inc^y; dec^z) \times a^{x+y} b^z \times \alpha^x \gamma^y \beta^z \times 2^{x+2y}$$

A definition of weighted transducers

(using way too much algebra)

storage in out weight

$$T = \langle \mathrm{inc}, a, \alpha, 2\rangle^* \; \langle \mathrm{inc}, a, \gamma, 4\rangle^* \; \langle \mathrm{dec}, b, \beta, 1\rangle^*$$

$$= (\mathrm{inc} \times a \times \alpha \times 2)^* (\mathrm{inc} \times a \times \gamma \times 4)^* (\mathrm{dec} \times b \times \beta \times 1)^*$$

$$= \sum_{x,y,z} (\mathrm{inc} \times a \times \alpha \times 2)^x (\mathrm{inc} \times a \times \gamma \times 4)^y (\mathrm{dec} \times b \times \beta \times 1)^z$$

$$\tau(T) = \sum_{x,y,z} \underbrace{(\mathrm{inc}^x ; \mathrm{inc}^y ; \mathrm{dec}^z)}_{\{(n, n+x+y-z) \mid n \in \mathbb{N}\} \longrightarrow \sum_n n \times (n+x+y-z)} \times a^{x+y} b^z \times \alpha^x \gamma^y \beta^z \times 2^{x+2y}$$

A definition of weighted transducers

(using way too much algebra)

$$T = \langle \overset{\text{storage}}{\mathrm{inc}}, \overset{\text{in}}{a}, \overset{\text{out}}{\alpha}, \overset{\text{weight}}{2} \rangle^* \quad \langle \mathrm{inc}, a, \gamma, 4 \rangle^* \quad \langle \mathrm{dec}, b, \beta, 1 \rangle^*$$

$$= (\mathrm{inc} \times a \times \alpha \times 2)^* (\mathrm{inc} \times a \times \gamma \times 4)^* (\mathrm{dec} \times b \times \beta \times 1)^*$$

$$= \sum_{x,y,z} (\mathrm{inc} \times a \times \alpha \times 2)^x (\mathrm{inc} \times a \times \gamma \times 4)^y (\mathrm{dec} \times b \times \beta \times 1)^z$$

$$\tau(T) = \sum_{x,y,z} \underbrace{(\mathrm{inc}^x; \mathrm{inc}^y; \mathrm{dec}^z)}_{\{(n, n+x+y-z) \mid n \in \mathbb{N}\} \longrightarrow \sum_n n \times (n+x+y-z)} \times a^{x+y} b^z \times \alpha^x \gamma^y \beta^z \times 2^{x+2y}$$

$$= \sum_{x,y,z,n} n \times (n+x+y-z) \times a^{x+y} b^z \times \alpha^x \gamma^y \beta^z \times 2^{x+2y}$$

Transducer semantics (application)

$$\mathcal{I}(T) = \sum_{x,y,z,n} n \times (n+x+y-z) \times a^{x+y} b^{z} \times \alpha^{x} \gamma^{y} \beta^{z} \times 2^{x+2y}$$

Transducer semantics (application)

$$\boxed{\mathcal{I}(T) = \sum_{x,y,z,n} n \times (n+x+y-z) \times a^{x+y} b^{z} \times \alpha^{x} \gamma^{y} \beta^{z} \times 2^{x+2y}}$$

$$\mathcal{I}(T)(0 \times 0) = \sum_{x,y} a^{x+y} b^{x+y} \times \alpha^{x} \gamma^{y} \beta^{x+y} \times 2^{x+2y}$$

Transducer semantics (application)

$$\boxed{\mathcal{I}(T) = \sum_{x,y,z,n} n \times (n+x+y-z) \times a^{x+y} b^{z} \times \alpha^{x} \gamma^{y} \beta^{z} \times 2^{x+2y}}$$

$$\mathcal{I}(T)(0 \times 0) = \underbrace{\sum_{x,y} a^{x+y} b^{x+y} \times \alpha^{x} \gamma^{y} \beta^{x+y} \times 2^{x+2y}}_{\text{weighted transduction}} =: [\![T]\!]$$

Transducer semantics (application)

$$\boxed{\tau(T) = \sum_{x,y,z,n} n \times (n+x+y-z) \times a^{x+y} b^{z} \times \alpha^{x} \gamma^{y} \beta^{z} \times 2^{x+2y}}$$

$$\tau(T)(0 \times 0) = \underbrace{\sum_{x,y} a^{x+y} b^{x+y} \times \alpha^{x} \gamma^{y} \beta^{x+y} \times 2^{x+2y}}_{\text{weighted transduction}} =: [\![T]\!]$$

$$\tau(T)(0 \times 0 \times a^{2} b^{2}) = \tau(T)(0 \times 0)(a^{2} b^{2})$$

Transducer semantics (application)

$$\boxed{\tau(T) = \sum_{x,y,z,n} n \times (n+x+y-z) \times a^{x+y} b^{z} \times \alpha^{x} \gamma^{y} \beta^{z} \times 2^{x+2y}}$$

$$\tau(T)(0 \times 0) = \underbrace{\sum_{x,y} a^{x+y} b^{x+y} \times \alpha^{x} \gamma^{y} \beta^{x+y} \times 2^{x+2y}}_{\text{weighted transduction}} =: [\![T]\!]$$

$$\tau(T)(0 \times 0 \times a^2 b^2) = \tau(T)(0 \times 0)(a^2 b^2)$$

$$= \alpha^2 \beta^2 \times 2^2 + \alpha\gamma\beta^2 \times 2^3 + \gamma^2 \beta^2 \times 2^4$$

Transducer semantics (application)

$$\boxed{\tau(T) = \sum_{x,y,z,n} n \times (n+x+y-z) \times a^{x+y} b^{z} \times \alpha^{x} \gamma^{y} \beta^{z} \times 2^{x+2y}}$$

$$\tau(T)(0 \times 0) = \underbrace{\sum_{x,y} a^{x+y} b^{x+y} \times \alpha^{x} \gamma^{y} \beta^{x+y} \times 2^{x+2y}}_{\text{weighted transduction}} =: [\![T]\!]$$

$$\tau(T)(0 \times 0 \times a^{2} b^{2}) = \tau(T)(0 \times 0)(a^{2} b^{2})$$

$$= \underbrace{\alpha^{2}\beta^{2} \times 2^{2} + \alpha\gamma\beta^{2} \times 2^{3} + \gamma^{2}\beta^{2} \times 2^{4}}_{\text{weighted language}}$$

References

[Goldstine 79] J. Goldstine: A rational theory of AFLs, 1979. DOI.

[Goldstine 80] J. Goldstine: Formal languages and their relation to automata – What Hopcroft & Ullman didn't tell us, 1980. ISBN: 0-12-115350-9.