Lipschitz Robustness of Finite-state Transducers

Roopsha Samanta
IST Austria

Joint work with Tom Henzinger and Jan Otop

December 16, 2014
Problem Overview

Computational systems in physical environments $\rightarrow$ uncertain data
Problem Overview

- Corrupt data from sensors in avionics software or medical devices
- Incomplete DNA strings in computational biology
- Wrongly spelled inputs to text processors
Problem Overview

- Functional correctness not enough
- System behaviour must degrade smoothly given input perturbation
- We need continuity or robustness
Informal definition

- Focus on finite-state transducers
- Robustness of transducers $\sim$ Lipschitz continuity
- Transducer $\mathcal{T}$ is $K$-Lipschitz robust if for all inputs $s, t$:
  $$d(s, t) < \infty \Rightarrow d(\mathcal{T}(s), \mathcal{T}(t)) \leq Kd(s, t)$$
Example

\[ d(s, t) = \text{Hamming distance}(s, t) \text{ if } |s| = |t|, \text{ else } \infty \]

Inputs: \( d(a^{k+1}, ba^k) = 1 \)
Outputs: \( d(a^{k+1}, b^{k+1}) = k + 1 \)

Inputs: \( d(a^{k+1}, ba^k) = 1 \)
Outputs: \( d(b^{k+1}, ab^k) = 1 \)
Main contributions

- Undecidability of $K$-robustness of deterministic transducers
- Characterization of functional transducers with decidable $K$-robustness
  - Consider distances computable by weighted automata
  - Reduce $K$-robustness to emptiness problem for weighted automata
  - Polynomial-time decision procedure
- Formalization/study of $K$-robustness of nondeterministic transducers
- All results hold for transducers over finite or infinite words
Main contributions

- Undecidability of $K$-robustness of deterministic transducers
- Characterization of functional transducers with decidable $K$-robustness
  - Consider distances computable by weighted automata
  - Reduce $K$-robustness to emptiness problem for weighted automata
  - Polynomial-time decision procedure
- Formalization/study of $K$-robustness of nondeterministic transducers
- All results hold for transducers over finite or infinite words
Main contributions

- Undecidability of $K$-robustness of deterministic transducers
- Characterization of functional transducers with decidable $K$-robustness
  - Consider distances computable by weighted automata
  - Reduce $K$-robustness to emptiness problem for weighted automata
  - Polynomial-time decision procedure
- Formalization/study of $K$-robustness of nondeterministic transducers
- All results hold for transducers over finite or infinite words
Main contributions

- Undecidability of $K$-robustness of deterministic transducers
- Characterization of functional transducers with decidable $K$-robustness
  - Consider distances computable by weighted automata
  - Reduce $K$-robustness to emptiness problem for weighted automata
  - Polynomial-time decision procedure
- Formalization/study of $K$-robustness of nondeterministic transducers
- All results hold for transducers over finite or infinite words
Main contributions

- Undecidability of $K$-robustness of *deterministic* transducers
- Characterization of *functional* transducers with decidable $K$-robustness
  - Consider distances computable by weighted automata
  - Reduce $K$-robustness to emptiness problem for weighted automata
  - Polynomial-time decision procedure
- Formalization/study of $K$-robustness of *nondeterministic* transducers
- All results hold for transducers over finite or infinite words
Transducers ($\mathcal{T}$)

- **Finite-state device with two tapes**
  - In each step, a transducer $\mathcal{T}$
    - reads an input letter (alphabet $\Sigma$)
    - writes a finite word (alphabet $\Gamma$)
    - nondeterministically changes state
  - Output of $\mathcal{T}$: defined only if run is *accepting*
    - $\text{dom}(\mathcal{T}) \subseteq \Sigma^*$
    - Transduction $\lbrack \mathcal{T} \rbrack \subseteq \text{dom}(\mathcal{T}) \times \Gamma^*$
  - Functional: at most one output word for every input word
    - $s' = \lbrack \mathcal{T} \rbrack(s)$
  - Mealy machines: deterministic, letter-to-letter, all states accepting
Transducers ($\mathcal{T}$)

- Finite-state device with two tapes
- In each step, a transducer $\mathcal{T}$
  - reads an input letter (alphabet $\Sigma$)
  - writes a finite word (alphabet $\Gamma$)
  - nondeterministically changes state
- Output of $\mathcal{T}$: defined only if run is *accepting*
  - $\text{dom}(\mathcal{T}) \subseteq \Sigma^*$
  - Transduction $\mathcal{[}\mathcal{T}] \subseteq \text{dom}(\mathcal{T}) \times \Gamma^*$
- Functional: at most one output word for every input word
  - $s' = \mathcal{[}\mathcal{T}](s)$
- Mealy machines: deterministic, letter-to-letter, all states accepting
Transducers ($\mathcal{T}$)

- Finite-state device with two tapes
- In each step, a transducer $\mathcal{T}$
  - reads an input letter (alphabet $\Sigma$)
  - writes a finite word (alphabet $\Gamma$)
  - nondeterministically changes state
- Output of $\mathcal{T}$: defined only if run is accepting
  - $\text{dom}(\mathcal{T}) \subseteq \Sigma^*$
  - Transduction $[\mathcal{T}] \subseteq \text{dom}(\mathcal{T}) \times \Gamma^*$
- Functional: at most one output word for every input word
  - $s' = [\mathcal{T}](s)$
- Mealy machines: deterministic, letter-to-letter, all states accepting
Transducers ($\mathcal{T}$)

- Finite-state device with two tapes
- In each step, a transducer $\mathcal{T}$
  - reads an input letter (alphabet $\Sigma$)
  - writes a finite word (alphabet $\Gamma$)
  - nondeterministically changes state
- Output of $\mathcal{T}$: defined only if run is *accepting*
  - $\text{dom}(\mathcal{T}) \subseteq \Sigma^*$
  - Transduction $\mathcal{T} \subseteq \text{dom}(\mathcal{T}) \times \Gamma^*$
- Functional: at most one output word for every input word
  - $s' = \mathcal{T}(s)$
- Mealy machines: deterministic, letter-to-letter, all states accepting
Transducers ($\mathcal{T}$)

- Finite-state device with two tapes
- In each step, a transducer $\mathcal{T}$
  - reads an input letter (alphabet $\Sigma$)
  - writes a finite word (alphabet $\Gamma$)
  - nondeterministically changes state
- Output of $\mathcal{T}$: defined only if run is accepting
  - $\text{dom}(\mathcal{T}) \subseteq \Sigma^*$
  - Transduction $\llbracket \mathcal{T} \rrbracket \subseteq \text{dom}(\mathcal{T}) \times \Gamma^*$
- Functional: at most one output word for every input word
  - $s' = \llbracket \mathcal{T} \rrbracket(s)$
- Mealy machines: deterministic, letter-to-letter, all states accepting
Weighted automata

- Finite automaton $A$ with weighted transitions (weights $\in \mathbb{Q}$)
- A (weighted) run $\pi$ is a sequence $c(\pi)$ of weights
- Given value-function $f$, value of a run $\pi$: $f(\pi) = f(c(\pi))$
  - Examples for value-functions: sum, discounted sum, limit-average
  - Notation: SUM-WA, DISC$\delta$-WA, LIMAVG-WA
- Value of a word $s$: $\mathcal{L}_A(s) = \inf_{\pi \in \text{Acc}(s)} f(\pi)$
Weighted automata

- Finite automaton $A$ with weighted transitions (weights $\in \mathbb{Q}$)
- A (weighted) run $\pi$ is a sequence $c(\pi)$ of weights
- Given value-function $f$, value of a run $\pi$: $f(\pi) = f(c(\pi))$
  - Examples for value-functions: sum, discounted sum, limit-average
  - Notation: $\text{SUM-WA}$, $\text{DISC}_\delta$-WA, $\text{LIMAVG}$-WA
- Value of a word $s$: $\mathcal{L}_A(s) = \inf_{\pi \in \text{Acc}(s)} f(\pi)$
Weighted automata

- Finite automaton $A$ with weighted transitions (weights $\in \mathbb{Q}$)
- A (weighted) run $\pi$ is a sequence $c(\pi)$ of weights
- Given value-function $f$, value of a run $\pi$: $f(\pi) = f(c(\pi))$
  - Examples for value-functions: sum, discounted sum, limit-average
  - Notation: $\text{SUM-WA}$, $\text{DISC}_\delta\text{-WA}$, $\text{LIMAVG-WA}$
- Value of a word $s$: $\mathcal{L}_A(s) = \inf_{\pi \in \text{Acc}(s)} f(\pi)$
Weighted automata

- Finite automaton $A$ with weighted transitions (weights $\in \mathbb{Q}$)
- A (weighted) run $\pi$ is a sequence $c(\pi)$ of weights
- Given value-function $f$, value of a run $\pi$: $f(\pi) = f(c(\pi))$
  - Examples for value-functions: sum, discounted sum, limit-average
  - Notation: $\text{SUM-WA}$, $\text{DISC}_\delta$-WA, $\text{LIMAVG-WA}$
- Value of a word $s$: $\mathcal{L}_A(s) = \inf_{\pi \in \text{Acc}(s)} f(\pi)$
Weighted automata

Given an $f$-wa $\mathcal{A}$ and a threshold $\lambda$:

- Emptiness: $\exists s : \mathcal{L}_{\mathcal{A}}(s) < \lambda$?
- Universality: $\forall s : \mathcal{L}_{\mathcal{A}}(s) < \lambda$?

- Emptiness is decidable in polynomial time for $\text{SUM-}$, $\text{DISC}_\delta$-, $\text{LIMAVG}$-WA
- Universality is undecidable for $\text{SUM}$-WA.
Similarity functions

\[ d : S \times S \to \mathbb{Q} \cup \infty \] such that \( \forall x, y \in S:\)

- \( d(x, y) \geq 0 \)
- \( d(x, y) = d(y, x) \)

Example: Manhattan distance
\[ d(s, t) = |\{ s[i] \neq t[i] \}| \]
Similarity functions

Pairing: \( ab \otimes bcaa = (\begin{array}{c}a \\ b \end{array}) (\begin{array}{c}b \\ c \end{array}) (\#) (\#) \)
Similarity functions

Pairing: \( ab \otimes bcaa = \left( \begin{array}{c} a \\ b \\ c \end{array} \right) \left( \begin{array}{c} b \\ a \end{array} \right) \)

\( d : S \times S \rightarrow \mathbb{Q}^{\infty} \) is computed by a weighted automaton \( \mathcal{A} \) if:
\[
\forall s, t \in S : \quad d(s, t) = \mathcal{L}_\mathcal{A}(s \otimes t)
\]
Similarity functions

Pairing: $ab \otimes bcaa = (a)(b)(\#)(\#)$

$d : S \times S \rightarrow \mathbb{Q}^\infty$ is computed by a weighted automaton $A$ if:

$\forall s, t \in S : d(s, t) = \mathcal{L}_A(s \otimes t)$

Example: Manhattan distance

\textbf{SUM-WA $A$}:

\begin{align*}
(q_0) & \quad (a) \\
(a) & \quad (b) \\
(b) & \quad (a) \\
0 & \quad q_0 \\
1 & \quad (a) \\
\end{align*}

$\mathcal{L}_A((s[1])(s[2]) \ldots) = |\{s[i] \neq t[i]\}|$
Robust transducers

Given:
- a functional transducer $T$, with $[T] \subseteq \Sigma^* \times \Gamma^*$
- similarity function $d_\Sigma : \Sigma^* \times \Sigma^* \rightarrow Q^\infty$
- similarity function $d_\Gamma : \Gamma^* \times \Gamma^* \rightarrow Q^\infty$
- constant $K \in \mathbb{Q}$ with $K > 0$

$T$ is defined to be $K$-Lipschitz robust w.r.t $d_\Sigma, d_\Gamma$ if:

$$\forall s, t \in \text{dom}(T) : d_\Sigma(s, t) < \infty \Rightarrow d_\Gamma([T](s), [T](t)) \leq Kd_\Sigma(s, t).$$
Robust transducers

Given:

- a functional transducer $\mathcal{T}$, with $\mathcal{T} \subseteq \Sigma^* \times \Gamma^*$
- similarity function $d_\Sigma : \Sigma^* \times \Sigma^* \rightarrow \mathbb{Q}^\infty$
- similarity function $d_\Gamma : \Gamma^* \times \Gamma^* \rightarrow \mathbb{Q}^\infty$
- constant $K \in \mathbb{Q}$ with $K > 0$

$\mathcal{T}$ is defined to be $K$-Lipschitz robust w.r.t $d_\Sigma, d_\Gamma$ if:

$$\forall s, t \in \text{dom}(\mathcal{T}) : d_\Sigma(s, t) < \infty \Rightarrow d_\Gamma([\mathcal{T]}(s), [\mathcal{T]}(t)) \leq Kd_\Sigma(s, t).$$
Problem definition

Given:
- a functional transducer $\mathcal{T}$, with $[\mathcal{T}] \subseteq \Sigma^* \times \Gamma^*$
- similarity function $d_{\Sigma} : \Sigma^* \times \Sigma^* \rightarrow \mathbb{Q}^\infty$
- similarity function $d_{\Gamma} : \Gamma^* \times \Gamma^* \rightarrow \mathbb{Q}^\infty$
- constant $K \in \mathbb{Q}$ with $K > 0$

Check if $\mathcal{T}$ is $K$-Lipschitz robust w.r.t $d_{\Sigma}$, $d_{\Gamma}$.
Undecidability

\( K \)-robustness of functional transducers is undecidable.

1-robustness of deterministic transducers w.r.t. generalized Manhattan distances is undecidable.

Proof hint: Reduction from PCP.
From $K$-robustness to robustness

$T$ is robust w.r.t. $d_\Sigma, d_r$ if $\exists K: T$ is $K$-robust w.r.t. $d_\Sigma, d_r$.

Let $d_\Sigma, d_r$ be generalized Manhattan distances.

1. Robustness of $T$ is decidable in $\text{co-NP}$.
2. One can compute $K_T$ such that $T$ is robust iff $T$ is $K_T$-robust.
Synchronized transducers

A transducer $T$ with $\mathcal{T}(T) \subseteq \Sigma^\omega \times \Gamma^\omega$ is synchronized iff:

There exists an automaton $A_T$ over $\Sigma \otimes \Gamma$ recognizing $\{s \otimes \mathcal{T}(T)(s) : s \in \text{dom}(T)\}$.

Synchronicity of a functional transducer can be decided in polynomial time.

Example: Mealy machines are synchronized transducers.
$q \xrightarrow{a \otimes a'} q' \in A_T$ iff $(q, a, a', q') \in T$
Robustness of synchronized transducers

If $d_\Sigma$, $d_\Gamma$ are similarity functions computed by functional $f$-WA $A_{d_\Sigma}$, $A_{d_\Gamma}$, $K$-robustness of synchronized $T$ w.r.t. $d_\Sigma$, $d_\Gamma$ is decidable in PTIME.

Proof hint:
Construct $A$ over $s \otimes t \otimes s' \otimes t'$:

$$
\bar{A}_{d_\Sigma}^K \times \bar{A}_T^L \times \bar{A}_T^R \times \bar{A}_{d_\Gamma}^{-1}
$$

$A_T$ accepts $s \otimes s'$  $A_T$ accepts $t \otimes t'$

$\mathcal{L}_{A_{d_\Sigma}}(s \otimes t)$  $\mathcal{L}_{A_{d_\Gamma}}(s' \otimes t')$

$\exists w : \mathcal{L}_A(w) < 0$ iff $T$ is not $K$-robust w.r.t. $d_\Sigma$, $d_\Gamma$
Example

Recall:

- Mealy machines are synchronized transducers
- Manhattan distances are computable by functional $\text{SUM-wa}$

$K$-robustness of Mealy machines w.r.t. Manhattan distances is decidable.
Nondeterministic transducers

For a nondeterministic transducer $\mathcal{T}$, $|\mathcal{T}(s)| \geq 1$
Nondeterministic transducers

Given:

- a transducer $T$, with $\llbracket T \rrbracket \subseteq \Sigma^* \times \Gamma^*$,
- similarity function $d_\Sigma : \Sigma^* \times \Sigma^* \rightarrow \mathbb{Q}^\infty$
- set-similarity function $D_\Gamma : 2^{\Gamma^* \times \Gamma^*} \rightarrow \mathbb{Q}^\infty$
- constant $K \in \mathbb{Q}$ with $K > 0$

$T$ is defined to be $K$-robust w.r.t $d_\Sigma$, $D_\Gamma$ if:

$$\forall s, t \in \text{dom}(T) : d_\Sigma(s, t) < \infty \Rightarrow D_\Gamma(\llbracket T \rrbracket(s), \llbracket T \rrbracket(t)) \leq Kd_\Sigma(s, t).$$
Nondeterministic transducers

<table>
<thead>
<tr>
<th>Set-similarity functions $D(A, B)$ induced by $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hausdorff:</strong> $\max \left{ \sup_{s \in A} \inf_{t \in B} d(s, t), \sup_{s \in B} \inf_{t \in A} d(s, t) \right}$</td>
</tr>
<tr>
<td><strong>Inf-inf:</strong> $\inf_{s \in A} \inf_{t \in B} d(s, t)$</td>
</tr>
<tr>
<td><strong>Sup-sup:</strong> $\sup_{s \in A} \sup_{t \in B} d(s, t)$</td>
</tr>
</tbody>
</table>
Nondeterministic transducers

<table>
<thead>
<tr>
<th>Set-similarity functions $D(A, B)$ induced by $d$</th>
<th>$K$-robustness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hausdorff: $\max { \sup_{s \in A} \inf_{t \in B} d(s, t) , \sup_{s \in B} \inf_{t \in A} d(s, t) }$</td>
<td>Undecidable</td>
</tr>
<tr>
<td>Inf-inf: $\inf_{s \in A} \inf_{t \in B} d(s, t)$</td>
<td>Undecidable</td>
</tr>
<tr>
<td>Sup-sup: $\sup_{s \in A} \sup_{t \in B} d(s, t)$</td>
<td>Decidable</td>
</tr>
</tbody>
</table>
Thank you.