Ternary is Still Good for Parikh Matrices

TheorieTag 2024

Robert Mercaș   Wen Chean Teh

20th September 2024

Loughborough University

1. Preliminaries
2. ℙ-Distinguishability
3. Minimal Hamming Distances
4. Conclusions

Compression

Compression

Wikipedia

In signal processing, data compression, source coding, or bit-rate reduction involves encoding information using fewer bits than the original presentation.

Compression

Wikipedia

In signal processing, data compression, source coding, or bit-rate reduction involves encoding information using fewer bits than the original presentation. Compression can be either lossy or lossless.

- Lossless compression reduces bits by identifying and eliminating statistical redundancy (no information is lost in compression).
- Lossy compression reduces bits by removing unnecessary or less important information.

Compression

Wikipedia

In signal processing, data compression, source coding, or bit-rate reduction involves encoding information using fewer bits than the original presentation. Compression can be either lossy or lossless.

- Lossless compression reduces bits by identifying and eliminating statistical redundancy (no information is lost in compression).
- Lossy compression reduces bits by removing unnecessary or less important information.

Lossless techniques

- Lempel-Ziv factorization LZ77 [Lempel and Ziv, 1976, Lempel and Ziv, 1977];
- Straight Line Programs [Kieffer and Yang, 2000];
- run-length Burrows-Wheeler transform [Burrows and Wheeler, 1994];
- Compact Directed Acyclic Word Graph [Blumer et al., 1987], [Crochemore and Vérin, 1997].

Parikh matrices

- Histograms;
- Parikh vectors [Parikh, 1966] represent a type histograms specific to the analysis of sequences of symbols;
- compression is easily computed and guaranteed to be logarithmic in the size of the word

Parikh matrices

- Histograms;
- Parikh vectors [Parikh, 1966] represent a type histograms specific to the analysis of sequences of symbols;
- compression is easily computed and guaranteed to be logarithmic in the size of the word, but ambiguous (multiple words share a Parikh vector).

Parikh matrices

- Histograms;
- Parikh vectors [Parikh, 1966] represent a type histograms specific to the analysis of sequences of symbols;
- compression is easily computed and guaranteed to be logarithmic in the size of the word, but ambiguous (multiple words share a Parikh vector).

- Parikh matrices [Mateescu et al., 2000] are a refinement of the vectors;
- same asymptotic compactness as Parikh vector and significantly smaller number of words associated to it;
- but it does not normally remove ambiguity (entirely).

Example

Definition

Let $w \in \Sigma_k^*$. The *Parikh matrix*, denoted $\Psi(w)$, that w is associated with has size $(k+1) \times (k+1)$. The diagonal of the matrix is populated with 1's and all elements below it are 0. The count of all subwords that consist of consecutive letters in Σ_k and are of length n in the word are found on the n-diagonal, for $1 \leq n \leq k$, i. e., the Parikh vector is found on the 1-diagonal.

Example

Definition

Let $w \in \Sigma_k^*$. The *Parikh matrix*, denoted $\Psi(w)$, that w is associated with has size $(k+1) \times (k+1)$. The diagonal of the matrix is populated with 1's and all elements below it are 0. The count of all subwords that consist of consecutive letters in Σ_k and are of length n in the word are found on the n-diagonal, for $1 \leq n \leq k$, i. e., the Parikh vector is found on the 1-diagonal.

For $\Sigma_3 = \{a < b < c\}$ we consider all of the factors of abc.

Thus, for $w \in \Sigma_3$ the Parikh matrix is of size 4×4 and we count all occurrences of the above factors as (scattered) subwords:

$$\Psi(aabcbaa) = \begin{pmatrix} 1 & |w|_a & |w|_{ab} & |w|_{abc} \\ 0 & 1 & |w|_b & |w|_{bc} \\ 0 & 0 & 1 & |w|_c \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 4 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Example

Consider the word $w = abca$ defined over the alphabet $\Sigma_3 = \{a < b < c\}$.

Example

Consider the word $w = abca$ defined over the alphabet $\Sigma_3 = \{a < b < c\}$. The Parikh matrix is of size 4×4 and we have:

$$\Psi(abca) = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Example

Consider the word $w = abca$ defined over the alphabet $\Sigma_3 = \{a < b < c\}$. The Parikh matrix is of size 4×4 and we have:

$$\Psi(abca) = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Example

Consider the word $w = abca$ defined over the alphabet $\Sigma_3 = \{a < b < c\}$. The Parikh matrix is of size 4×4 and we have:

$$\begin{aligned}\Psi(abca) &= \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \Psi(a) \cdot \Psi(b) \cdot \Psi(c) \cdot \Psi(a)\end{aligned}$$

Example

Consider the word $w = abca$ defined over the alphabet $\Sigma_3 = \{a < b < c\}$.
The Parikh matrix is of size 4×4 and we have:

$$
\begin{aligned}
\Psi(abca) &= \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \Psi(a) \cdot \Psi(b) \cdot \Psi(c) \cdot \Psi(a)
\end{aligned}
$$

$$
\Psi(abca) = \left\langle \begin{matrix} 2 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{matrix} \right\rangle
$$

A big mess (on integer coefficients)

A big mess (on integer coefficients)

Parikh vectors are not injective.

$\phi(abbcabac) = (3, 3, 2) = \phi(aaabbbcc)$

A big mess (on integer coefficients)

Parikh vectors are not injective.

$$\phi(abbcabac) = (3,3,2) = \phi(aaabbbcc)$$

Parikh matrices are not injective

$$\Psi(abca) = \left\langle \begin{matrix} 2 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{matrix} \right\rangle = \Psi(abac)$$

A big mess (on integer coefficients)

Parikh vectors are not injective.

$\phi(abbcabac) = (3, 3, 2) = \phi(aaabbbcc)$

Parikh matrices are not injective, but also not surjective.

$\Psi(abca) = \left\langle \begin{smallmatrix} 2 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{smallmatrix} \right\rangle = \Psi(abac)$

there exists no w with $\Psi(w) = \left\langle \begin{smallmatrix} 1 & 0 & 1 \\ & 1 & 0 \\ & & 1 \end{smallmatrix} \right\rangle$

A big mess (on integer coefficients)

Parikh vectors are not injective.

$\phi(abbcabac) = (3, 3, 2) = \phi(aaabbbcc)$

Parikh matrices are not injective, but also not surjective.

$\Psi(abca) = \left\langle \begin{smallmatrix} 2 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{smallmatrix} \right\rangle = \Psi(abac)$

there exists no w with $\Psi(w) = \left\langle \begin{smallmatrix} 1 & 0 & 1 \\ & 1 & 0 \\ & & 1 \end{smallmatrix} \right\rangle$

Note

If two words share a Parikh matrix, we say that they are *amiable* (use $\equiv$).

History

History

- alternative to the Parikh matrix concept that would make a mapping from a word injective, or less ambiguous in general [Șerbănuță, 2004], [Egecioglu, 2004], [Egecioglu and Ibarra, 2004], [Egecioglu and Ibarra, 2007], [Alazemi and Černý, 2011], [Alazemi and Černý, 2013], [Bera and Mahalingam, 2016]

History

- alternative to the Parikh matrix concept that would make a mapping from a word injective, or less ambiguous in general [Şerbănuţă, 2004], [Egecioglu, 2004], [Egecioglu and Ibarra, 2004], [Egecioglu and Ibarra, 2007], [Alazemi and Černý, 2011], [Alazemi and Černý, 2013], [Bera and Mahalingam, 2016];
- for Parikh matrices the focus was on investigating uniqueness on binary [Atanasiu et al., 2001], [Atanasiu et al., 2002], [Atanasiu et al., 2008], [Salomaa and Yu, 2010], [Atanasiu and Teh, 2016]

History

- alternative to the Parikh matrix concept that would make a mapping from a word injective, or less ambiguous in general [Şerbănuţă, 2004], [Egecioglu, 2004], [Egecioglu and Ibarra, 2004], [Egecioglu and Ibarra, 2007], [Alazemi and Černý, 2011], [Alazemi and Černý, 2013], [Bera and Mahalingam, 2016];
- for Parikh matrices the focus was on investigating uniqueness on binary [Atanasiu et al., 2001], [Atanasiu et al., 2002], [Atanasiu et al., 2008], [Salomaa and Yu, 2010], [Atanasiu and Teh, 2016] and ternary [Şerbănuţă, 2009], [Mahalingam and Subramanian, 2012], [Atanasiu, 2014], [Poovanandran and Teh, 2017]

History

- alternative to the Parikh matrix concept that would make a mapping from a word injective, or less ambiguous in general [Șerbănuță, 2004], [Egecioglu, 2004], [Egecioglu and Ibarra, 2004], [Egecioglu and Ibarra, 2007], [Alazemi and Černý, 2011], [Alazemi and Černý, 2013], [Bera and Mahalingam, 2016];
- for Parikh matrices the focus was on investigating uniqueness on binary [Atanasiu et al., 2001], [Atanasiu et al., 2002], [Atanasiu et al., 2008], [Salomaa and Yu, 2010], [Atanasiu and Teh, 2016] and ternary [Șerbănuță, 2009], [Mahalingam and Subramanian, 2012], [Atanasiu, 2014], [Poovanandran and Teh, 2017];
- *reducing* ambiguity, investigation was based on either gathering more information about the specific word by altering the order of the alphabet (dual order) or by considering the reverse image of the word [Mateescu et al., 2000], [Atanasiu et al., 2002]

History

- alternative to the Parikh matrix concept that would make a mapping from a word injective, or less ambiguous in general [Șerbănuță, 2004], [Egecioglu, 2004], [Egecioglu and Ibarra, 2004], [Egecioglu and Ibarra, 2007], [Alazemi and Černý, 2011], [Alazemi and Černý, 2013], [Bera and Mahalingam, 2016];
- for Parikh matrices the focus was on investigating uniqueness on binary [Atanasiu et al., 2001], [Atanasiu et al., 2002], [Atanasiu et al., 2008], [Salomaa and Yu, 2010], [Atanasiu and Teh, 2016] and ternary [Șerbănuță, 2009], [Mahalingam and Subramanian, 2012], [Atanasiu, 2014], [Poovanandran and Teh, 2017];
- *reducing* ambiguity, investigation was based on either gathering more information about the specific word by altering the order of the alphabet (dual order) or by considering the reverse image of the word [Mateescu et al., 2000], [Atanasiu et al., 2002];
- *additional* matrices used to remove ambiguity, by altering the word itself: projection morphisms and conjugates [Dick et al., 2021]

History

- alternative to the Parikh matrix concept that would make a mapping from a word injective, or less ambiguous in general [Şerbănuţă, 2004], [Egecioglu, 2004], [Egecioglu and Ibarra, 2004], [Egecioglu and Ibarra, 2007], [Alazemi and Černý, 2011], [Alazemi and Černý, 2013], [Bera and Mahalingam, 2016];
- for Parikh matrices the focus was on investigating uniqueness on binary [Atanasiu et al., 2001], [Atanasiu et al., 2002], [Atanasiu et al., 2008], [Salomaa and Yu, 2010], [Atanasiu and Teh, 2016] and ternary [Şerbănuţă, 2009], [Mahalingam and Subramanian, 2012], [Atanasiu, 2014], [Poovanandran and Teh, 2017];
- *reducing* ambiguity, investigation was based on either gathering more information about the specific word by altering the order of the alphabet (dual order) or by considering the reverse image of the word [Mateescu et al., 2000], [Atanasiu et al., 2002];
- *additional* matrices used to remove ambiguity, by altering the word itself: projection morphisms and conjugates [Dick et al., 2021];
- *injectivity* was recently settled for ternary [Baek et al., 2024], [Salomaa, 2010], and higher alphabets (but using rational exponents) [Hahn et al., 2023].

Elementary equivalence preserving rules

Let $\Sigma = \{a_1 < a_2 < \cdots < a_s\}$ and $w, w' \in \Sigma^*$.

*E*1. If $w = xa_ia_jy$ and $w' = xa_ja_iy$, for $x, y \in \Sigma^*$ & $|i-j| \geq 2$, then $w \equiv w'$.

$$\psi(aaaacb) = \psi(aaacab) = \psi(aacaab) = \psi(acaaab) = \psi(caaaab)$$

*E*2. If $w = xa_ja_{j+1}ya_{j+1}a_jz$ and $w' = xa_{j+1}a_jya_ja_{j+1}z$, for $x, y, z \in \Sigma^*$ such that $|y|_{a_{j-1}} = |y|_{a_{j+2}} = 0$, then $w \equiv w'$.

$$\psi(aaabbaac) = \psi(aabaabac) = \psi(abaaaabc)$$

Elementary equivalence preserving rules

Let $\Sigma = \{a_1 < a_2 < \cdots < a_s\}$ and $w, w' \in \Sigma^*$.

*E*1. If $w = xa_ia_jy$ and $w' = xa_ja_iy$, for $x, y \in \Sigma^*$ & $|i-j| \geq 2$, then $w \equiv w'$.

$$\psi(aaaacb) = \psi(aaacab) = \psi(aacaab) = \psi(acaaab) = \psi(caaaab)$$

*E*2. If $w = xa_ja_{j+1}ya_{j+1}a_jz$ and $w' = xa_{j+1}a_jya_ja_{j+1}z$, for $x, y, z \in \Sigma^*$ such that $|y|_{a_{j-1}} = |y|_{a_{j+2}} = 0$, then $w \equiv w'$.

$$\psi(aaabbaac) = \psi(aabaabac) = \psi(abaaaabc)$$

Theorem ([Atanasiu, 2007], [Mateescu and Salomaa, 2004])

Let $\Sigma = \{a < b\}$ and $w, w' \in \Sigma^*$. Then $w \equiv w'$ if and only if w can be obtained from w' by finitely many applications of Rule E_2. As a consequence, the following are all unambiguous words over Σ:

$$\lambda, a^\alpha, b^\alpha, a^\alpha b^\beta, b^\alpha a^\beta, a^\alpha ba^\beta, b^\alpha ab^\beta, a^\alpha bab^\beta, b^\alpha aba^\beta, \quad \alpha, \beta \geq 1.$$

1 Preliminaries

2 $\mathbb{P}$-Distinguishability

3 Minimal Hamming Distances

4 Conclusions

ℙ-Parikh Matrix

Definition

Let $S \subset \Sigma_n$ such that $S = \{a_{k_1}, a_{k_2}, \ldots a_{k_m}\}$. We define the ℙ-Parikh matrix of the word w with respect to S as $\Psi_S(\pi_S(w))$, where the morphism $\pi : \Sigma_n^* \to \Sigma_m^*$ is defined as

$$\pi_S(a_i) := \begin{cases} a_i & : a_i \in S, \\ \varepsilon & : a_i \notin S. \end{cases}$$

$\mathbb{P}$-Parikh Matrix

Definition

Let $S \subset \Sigma_n$ such that $S = \{a_{k_1}, a_{k_2}, \ldots a_{k_m}\}$. We define the $\mathbb{P}$-Parikh matrix of the word w with respect to S as $\Psi_S(\pi_S(w))$, where the morphism $\pi : \Sigma_n^* \to \Sigma_m^*$ is defined as

$$\pi_S(a_i) := \begin{cases} a_i & : a_i \in S, \\ \varepsilon & : a_i \notin S. \end{cases}$$

Example

Let $\Sigma_5 = \{a, b, c, d, e\}$, $S = \{a, d, e\}$, and $w = bacbebda$.

ℙ-Parikh Matrix

Definition

Let $S \subset \Sigma_n$ such that $S = \{a_{k_1}, a_{k_2}, \ldots a_{k_m}\}$. We define the ℙ-Parikh matrix of the word w with respect to S as $\Psi_S(\pi_S(w))$, where the morphism $\pi : \Sigma_n^* \to \Sigma_m^*$ is defined as

$$\pi_S(a_i) := \begin{cases} a_i & : a_i \in S, \\ \varepsilon & : a_i \notin S. \end{cases}$$

Example

Let $\Sigma_5 = \{a, b, c, d, e\}$, $S = \{a, d, e\}$, and $w = bacbebda$.

$$\pi_S(a) = a, \qquad \pi_S(d) = b, \qquad \pi_S(e) = c$$

ℙ-Parikh Matrix

Definition

Let $S \subset \Sigma_n$ such that $S = \{a_{k_1}, a_{k_2}, \dots a_{k_m}\}$. We define the ℙ-Parikh matrix of the word w with respect to S as $\Psi_S(\pi_S(w))$, where the morphism $\pi : \Sigma_n^* \to \Sigma_m^*$ is defined as

$$\pi_S(a_i) := \begin{cases} a_i & : a_i \in S, \\ \varepsilon & : a_i \notin S. \end{cases}$$

Example

Let $\Sigma_5 = \{a, b, c, d, e\}$, $S = \{a, d, e\}$, and $w = bacbebda$.

$$\pi_S(a) = a, \qquad \pi_S(d) = b, \qquad \pi_S(e) = c$$

$$\pi_S(w) = \pi_S(b) \cdot \pi_S(a) \cdot \pi_S(c) \cdot \pi_S(b) \cdot \pi_S(e) \cdot \pi_S(b) \cdot \pi_S(d) \cdot \pi_S(a)$$

$\mathbb{P}$-Parikh Matrix

Definition

Let $S \subset \Sigma_n$ such that $S = \{a_{k_1}, a_{k_2}, \ldots a_{k_m}\}$. We define the $\mathbb{P}$-Parikh matrix of the word w with respect to S as $\Psi_S(\pi_S(w))$, where the morphism $\pi : \Sigma_n^* \to \Sigma_m^*$ is defined as

$$\pi_S(a_i) := \begin{cases} a_i & : a_i \in S, \\ \varepsilon & : a_i \notin S. \end{cases}$$

Example

Let $\Sigma_5 = \{a, b, c, d, e\}$, $S = \{a, d, e\}$, and $w = bacbebda$.

$$\pi_S(a) = a, \qquad \pi_S(d) = b, \qquad \pi_S(e) = c$$

$$\pi_S(w) = \pi_S(b) \cdot \pi_S(a) \cdot \pi_S(c) \cdot \pi_S(b) \cdot \pi_S(e) \cdot \pi_S(b) \cdot \pi_S(d) \cdot \pi_S(a)$$

$$= \varepsilon\ a\ \varepsilon\ \varepsilon\ c\ \varepsilon\ b\ a = acba$$

$\mathbb{P}$-Parikh Matrix

Definition

Let $S \subset \Sigma_n$ such that $S = \{a_{k_1}, a_{k_2}, \ldots a_{k_m}\}$. We define the $\mathbb{P}$-Parikh matrix of the word w with respect to S as $\Psi_S(\pi_S(w))$, where the morphism $\pi : \Sigma_n^* \to \Sigma_m^*$ is defined as

$$\pi_S(a_i) := \begin{cases} a_i & : a_i \in S, \\ \varepsilon & : a_i \notin S. \end{cases}$$

Example

Let $\Sigma_5 = \{a, b, c, d, e\}$, $S = \{a, d, e\}$, and $w = bacbebda$.

$$\pi_S(a) = a, \qquad \pi_S(d) = b, \qquad \pi_S(e) = c$$

$$\pi_S(w) = \pi_S(b) \cdot \pi_S(a) \cdot \pi_S(c) \cdot \pi_S(b) \cdot \pi_S(e) \cdot \pi_S(b) \cdot \pi_S(d) \cdot \pi_S(a)$$
$$= \varepsilon\ a\ \varepsilon\ \varepsilon\ c\ \varepsilon\ b\ a = acba$$

$$\Psi_S(bacbebda) = \Psi_S(\pi_{\{a,d,e\}}(bacbebda)) = \Psi_S(acba) = \left\langle \begin{matrix} 2 & 1 & 0 \\ & 1 & 0 \\ & & 1 \end{matrix} \right\rangle$$

When do $\mathbb{P}$-Parikh Matrices reduce ambiguity, and when not?

Example

Let $w = abd$, and $w' = adb$ in Σ_4. Then w and w' are amiable. Note that w contains the factor bd, and we choose $S = \{b, d\}$.

When do ℙ-Parikh Matrices reduce ambiguity, and when not?

Example

Let $w = abd$, and $w' = adb$ in Σ_4. Then w and w' are amiable. Note that w contains the factor bd, and we choose $S = \{b, d\}$.

$$\pi_S(abd) = bd$$

When do $\mathbb{P}$-Parikh Matrices reduce ambiguity, and when not?

Example

Let $w = abd$, and $w' = adb$ in Σ_4. Then w and w' are amiable. Note that w contains the factor bd, and we choose $S = \{b, d\}$.

$$\pi_S(abd) = bd \implies \qquad \Psi_S(w) = \Psi(\pi_S(abd)) = \Psi(bd) = \left\langle \begin{matrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{matrix} \right\rangle$$

When do ℙ-Parikh Matrices reduce ambiguity, and when not?

Example

Let $w = abd$, and $w' = adb$ in Σ_4. Then w and w' are amiable. Note that w contains the factor bd, and we choose $S = \{b, d\}$.

$$\pi_S(abd) = bd \implies \qquad \Psi_S(w) = \Psi(\pi_S(abd)) = \Psi(bd) = \left\langle \begin{smallmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{smallmatrix} \right\rangle$$

$$\pi_S(adb) = db$$

When do ℙ-Parikh Matrices reduce ambiguity, and when not?

Example

Let $w = abd$, and $w' = adb$ in Σ_4. Then w and w' are amiable. Note that w contains the factor bd, and we choose $S = \{b, d\}$.

$$\pi_S(abd) = bd \implies \quad \Psi_S(w) = \Psi(\pi_S(abd)) = \Psi(bd) = \left\langle \begin{smallmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{smallmatrix} \right\rangle$$

$$\pi_S(adb) = db \implies \quad \Psi_S(w') = \Psi(\pi_S(adb)) = \Psi(db) = \left\langle \begin{smallmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 1 \end{smallmatrix} \right\rangle$$

When do $\mathbb{P}$-Parikh Matrices reduce ambiguity, and when not?

Example

Let $w = abd$, and $w' = adb$ in Σ_4. Then w and w' are amiable. Note that w contains the factor bd, and we choose $S = \{b, d\}$.

$$\pi_S(abd) = bd \implies \Psi_S(w) = \Psi(\pi_S(abd)) = \Psi(bd) = \left\langle \begin{matrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{matrix} \right\rangle$$

$$\pi_S(adb) = db \implies \Psi_S(w') = \Psi(\pi_S(adb)) = \Psi(db) = \left\langle \begin{matrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 1 \end{matrix} \right\rangle$$

Example

Consider $w = acbbca$ and $w' = cabbac$. Then $\Psi(w) = \Psi(w') = \left\langle \begin{matrix} 2 & 2 & 2 \\ & 2 & 2 \\ & & 2 \end{matrix} \right\rangle$.

When do $\mathbb{P}$-Parikh Matrices reduce ambiguity, and when not?

Example

Let $w = abd$, and $w' = adb$ in Σ_4. Then w and w' are amiable. Note that w contains the factor bd, and we choose $S = \{b, d\}$.

$$\pi_S(abd) = bd \implies \Psi_S(w) = \Psi(\pi_S(abd)) = \Psi(bd) = \left\langle \begin{smallmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{smallmatrix} \right\rangle$$

$$\pi_S(adb) = db \implies \Psi_S(w') = \Psi(\pi_S(adb)) = \Psi(db) = \left\langle \begin{smallmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 1 \end{smallmatrix} \right\rangle$$

Example

Consider $w = acbbca$ and $w' = cabbac$. Then $\Psi(w) = \Psi(w') = \left\langle \begin{smallmatrix} 2 & 2 & 2 \\ & 2 & 2 \\ & & 2 \end{smallmatrix} \right\rangle$.
Now let us find $\Psi_S(w)$ and $\Psi_S(w')$, for all $S \subseteq \Sigma_3$.

When do $\mathbb{P}$-Parikh Matrices reduce ambiguity, and when not?

Example

Let $w = abd$, and $w' = adb$ in Σ_4. Then w and w' are amiable. Note that w contains the factor bd, and we choose $S = \{b, d\}$.

$$\pi_S(abd) = bd \implies \quad \Psi_S(w) = \Psi(\pi_S(abd)) = \Psi(bd) = \left\langle \begin{smallmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{smallmatrix} \right\rangle$$

$$\pi_S(adb) = db \implies \quad \Psi_S(w') = \Psi(\pi_S(adb)) = \Psi(db) = \left\langle \begin{smallmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 1 \end{smallmatrix} \right\rangle$$

Example

Consider $w = acbbca$ and $w' = cabbac$. Then $\Psi(w) = \Psi(w') = \left\langle \begin{smallmatrix} 2 & 2 & 2 \\ & 2 & 2 \\ & & 2 \end{smallmatrix} \right\rangle$. Now let us find $\Psi_S(w)$ and $\Psi_S(w')$, for all $S \subseteq \Sigma_3$.

$$S = \{a\}, S = \{b\}, S = \{c\} : \ \Psi_S(w) = \langle \begin{smallmatrix} 1 & 2 \\ 1 & 1 \end{smallmatrix} \rangle = \Psi_S(w')$$

When do ℙ-Parikh Matrices reduce ambiguity, and when not?

Example

Let $w = abd$, and $w' = adb$ in Σ_4. Then w and w' are amiable. Note that w contains the factor bd, and we choose $S = \{b, d\}$.

$$\pi_S(abd) = bd \implies \quad \Psi_S(w) = \Psi(\pi_S(abd)) = \Psi(bd) = \left\langle \begin{smallmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{smallmatrix} \right\rangle$$

$$\pi_S(adb) = db \implies \quad \Psi_S(w') = \Psi(\pi_S(adb)) = \Psi(db) = \left\langle \begin{smallmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 1 \end{smallmatrix} \right\rangle$$

Example

Consider $w = acbbca$ and $w' = cabbac$. Then $\Psi(w) = \Psi(w') = \left\langle \begin{smallmatrix} 2 & 2 & 2 \\ & 2 & 2 \\ & & 2 \end{smallmatrix} \right\rangle$.
Now let us find $\Psi_S(w)$ and $\Psi_S(w')$, for all $S \subseteq \Sigma_3$.

$$S = \{a\}, S = \{b\}, S = \{c\} : \ \Psi_S(w) = \langle \begin{smallmatrix} 1 & 2 \\ 1 & 1 \end{smallmatrix} \rangle = \Psi_S(w')$$

$$S = \{a, b\}, S = \{a, c\} : \ \Psi_S(w) = \left\langle \begin{smallmatrix} 1 & 2 & 2 \\ & 1 & 2 \\ & & 1 \end{smallmatrix} \right\rangle = \Psi_S(w')$$

When do ℙ-Parikh Matrices reduce ambiguity, and when not?

Example

Let $w = abd$, and $w' = adb$ in Σ_4. Then w and w' are amiable. Note that w contains the factor bd, and we choose $S = \{b, d\}$.

$$\pi_S(abd) = bd \implies \quad \Psi_S(w) = \Psi(\pi_S(abd)) = \Psi(bd) = \left\langle \begin{smallmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{smallmatrix} \right\rangle$$

$$\pi_S(adb) = db \implies \quad \Psi_S(w') = \Psi(\pi_S(adb)) = \Psi(db) = \left\langle \begin{smallmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 1 \end{smallmatrix} \right\rangle$$

Example

Consider $w = acbbca$ and $w' = cabbac$. Then $\Psi(w) = \Psi(w') = \left\langle \begin{smallmatrix} 2 & 2 & 2 \\ & 2 & 2 \\ & & 2 \end{smallmatrix} \right\rangle$.
Now let us find $\Psi_S(w)$ and $\Psi_S(w')$, for all $S \subseteq \Sigma_3$.

$$S = \{a\}, S = \{b\}, S = \{c\} : \ \Psi_S(w) = \langle \begin{smallmatrix} 1 & 2 \\ 1 & 1 \end{smallmatrix} \rangle = \Psi_S(w')$$

$$S = \{a, b\}, S = \{a, c\} : \ \Psi_S(w) = \left\langle \begin{smallmatrix} 1 & 2 & 2 \\ & 1 & 2 \\ & & 1 \end{smallmatrix} \right\rangle = \Psi_S(w')$$

$$S = \{b, c\} : \ \pi_S(w) = cbbc = \pi_S(w')$$

Ambiguity and equivalence

Definition ([Dick et al., 2021])

Let Σ be an ordered alphabet and $w, w' \in \Sigma^*$. We say that w and w' are $\mathbb{P}$-*distinct* if $\Psi_S(\pi_S(w)) \neq \Psi_S(\pi_S(w'))$, for some $S \subseteq \Sigma$. When w is ambiguous, we say that $[w]$ is $\mathbb{P}$-*distinguishable* (or simply distinguishable) if there exists $w' \in [w]$ such that w and w' are $\mathbb{P}$-distinct. Otherwise, $[w]$ is $\mathbb{P}$-*indistinguishable*.

Definition

Let Σ be an ordered alphabet and $w, w' \in \Sigma^*$. We say that w and w' are $\mathbb{P}$-*equivalent* if $\Psi_S(\pi_S(w)) = \Psi_S(\pi_S(w'))$, for every $S \subseteq \Sigma$. Equivalently, in this case, w and w' are not $\mathbb{P}$-distinct.

Ambiguity reduction

Conjecture ([Dick et al., 2021])

Let Σ be an ordered alphabet and $w \in \Sigma^*$ an ambiguous word. Then the class $[w]$ is $\mathbb{P}$-distinguishable if and only if $[w]$ is trivially $\mathbb{P}$-distinguishable.

Ambiguity reduction

Conjecture ([Dick et al., 2021])

Let Σ be an ordered alphabet and $w \in \Sigma^*$ an ambiguous word. Then the class $[w]$ is $\mathbb{P}$-distinguishable if and only if $[w]$ is trivially $\mathbb{P}$-distinguishable.

Trivially distinguishable means that no application of Rule $E1$ is possible:

$\psi(aaabcccbbabbabc)$ is not $\mathbb{P}$-distinguishable;

$\psi(aaabcccbbabbacc)$ is $\mathbb{P}$-distinguishable .

Main result

Theorem

If $[w]$ is $\mathbb{P}$-distinguishable, for $w \in \Sigma_3^*$, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

Proof Idea.

If ac^na or ca^nc factos of $\pi_{a,c}(w)$, $n \geq 2$, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

Main result

Theorem

If $[w]$ is $\mathbb{P}$-distinguishable, for $w \in \Sigma_3^*$, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

Proof Idea.

If ac^na or ca^nc factos of $\pi_{a,c}(w)$, $n \geq 2$, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

If $w = a_{i_1}^{p_1} a_{i_2}^{p_2} \cdots a_{i_n}^{p_n}$, for $a_{i_1}, a_{i_2}, \ldots, a_{i_n} \in \Sigma^*$ and integers $p_1, p_2, \ldots, p_n \geq 0$ such that $a_{i_j} \neq a_{i_{j+1}}$, then the *print* of w is $\text{print}(w) = a_{i_1} a_{i_2} \cdots a_{i_n}$.

Main result

Theorem

If $[w]$ is $\mathbb{P}$-distinguishable, for $w \in \Sigma_3^*$, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

Proof Idea.

If ac^na or ca^nc factos of $\pi_{a,c}(w)$, $n \geq 2$, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

If $w = a_{i_1}^{p_1} a_{i_2}^{p_2} \cdots a_{i_n}^{p_n}$, for $a_{i_1}, a_{i_2}, \ldots, a_{i_n} \in \Sigma^*$ and integers $p_1, p_2, \ldots, p_n \geq 0$ such that $a_{i_j} \neq a_{i_{j+1}}$, then the *print* of w is $\text{print}(w) = a_{i_1} a_{i_2} \cdots a_{i_n}$.

If $|\text{print}(\pi_{a,c}(w))| = 2$ and $[w]$ not trivially $\mathbb{P}$-distinguishable, then for $w' \in [w]$, $\pi_{a,c}(w) = a^m c^n$ and $\pi_{a,c}(w') = c^n a^m$. But, $|w|_{abc} \geq 1$ and $|w'|_{abc} = 0$.

Main result

Theorem

If $[w]$ is $\mathbb{P}$-distinguishable, for $w \in \Sigma_3^*$, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

Proof Idea.

If ac^na or ca^nc factos of $\pi_{a,c}(w)$, $n \geq 2$, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

If $w = a_{i_1}^{p_1} a_{i_2}^{p_2} \cdots a_{i_n}^{p_n}$, for $a_{i_1}, a_{i_2}, \ldots, a_{i_n} \in \Sigma^*$ and integers $p_1, p_2, \ldots, p_n \geq 0$ such that $a_{i_j} \neq a_{i_{j+1}}$, then the *print* of w is $\text{print}(w) = a_{i_1} a_{i_2} \cdots a_{i_n}$.

If $|\text{print}(\pi_{a,c}(w))| = 2$ and $[w]$ not trivially $\mathbb{P}$-distinguishable, then for $w' \in [w]$, $\pi_{a,c}(w) = a^m c^n$ and $\pi_{a,c}(w') = c^n a^m$. But, $|w|_{abc} \geq 1$ and $|w'|_{abc} = 0$.

If $|\text{print}(\pi_{a,c}(w))| = 3$, then $[w]$ is trivially $\mathbb{P}$-distinguishable or unambiguous.

If $|\text{print}(\pi_{a,c}(w))| = 5$, then $[w]$ is trivially $\mathbb{P}$-distinguishable or unambiguous.

If $|\text{print}(\pi_{a,c}(w))| = 6$, then $[w]$ is trivially $\mathbb{P}$-distinguishable or unambiguous.

Main result

Theorem

If $[w]$ is $\mathbb{P}$-distinguishable, for $w \in \Sigma_3^*$, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

Proof Idea.

If ac^na or ca^nc factos of $\pi_{a,c}(w)$, $n \geq 2$, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

If $w = a_{i_1}^{p_1} a_{i_2}^{p_2} \cdots a_{i_n}^{p_n}$, for $a_{i_1}, a_{i_2}, \ldots, a_{i_n} \in \Sigma^*$ and integers $p_1, p_2, \ldots, p_n \geq 0$ such that $a_{i_j} \neq a_{i_{j+1}}$, then the *print* of w is $\text{print}(w) = a_{i_1} a_{i_2} \cdots a_{i_n}$.

If $|\text{print}(\pi_{a,c}(w))| = 2$ and $[w]$ not trivially $\mathbb{P}$-distinguishable, then for $w' \in [w]$, $\pi_{a,c}(w) = a^m c^n$ and $\pi_{a,c}(w') = c^n a^m$. But, $|w|_{abc} \geq 1$ and $|w'|_{abc} = 0$.

If $|\text{print}(\pi_{a,c}(w))| = 3$, then $[w]$ is trivially $\mathbb{P}$-distinguishable or unambiguous.

If $|\text{print}(\pi_{a,c}(w))| = 5$, then $[w]$ is trivially $\mathbb{P}$-distinguishable or unambiguous.

If $|\text{print}(\pi_{a,c}(w))| = 6$, then $[w]$ is trivially $\mathbb{P}$-distinguishable or unambiguous.

If $|\text{print}(\pi_{a,c}(w))| \geq 7$, then $[w]$ is trivially $\mathbb{P}$-distinguishable. □

Proof sketch: $|\text{print}(\pi_{a,c}(w))| = 4$

Assume $\pi_{a,c}(w) = a^m cac^n$ and that $[w]$ is not trivially $\mathbb{P}$-distinguishable.

Without loss of generality, for some $\beta, \gamma, \delta \geq 1$ and $\alpha, \nu \geq 0$, we may assume

$$w = a^{m-1} b^{\alpha} ab^{\beta} cb^{\gamma} ab^{\delta} cb^{\nu} c^{n-1}.$$

Proof sketch: $|\text{print}(\pi_{a,c}(w))| = 4$

Assume $\pi_{a,c}(w) = a^m cac^n$ and that $[w]$ is not trivially $\mathbb{P}$-distinguishable.

Without loss of generality, for some $\beta, \gamma, \delta \geq 1$ and $\alpha, \nu \geq 0$, we may assume

$$w = a^{m-1} b^\alpha ab^\beta cb^\gamma ab^\delta cb^\nu c^{n-1}.$$

Case

$\pi_{a,c}(w') = c^{n+1}a^{m+1}$ $|w'|_{abc} \neq |w|_{abc}$

Case

$\pi_{a,c}(w') = a^{m+1}c^{n+1}$

contradiction on equality over numbers of ab and bc in the suffix

Case

$\pi_{a,c}(w') = c^n aca^m$

contradiction on equality over numbers of ab and bc in words

Thus, if $[w]$ is $\mathbb{P}$-distinguishable, then $[w]$ is trivially $\mathbb{P}$-distinguishable.

Four is not good

Example

Consider the ordered alphabet $\Sigma = \{a < b < c < d\}$ and the words $w = bcbabcbcdcbcbbabcbccdccbb$ and $v = cbbabbcbcdccbcbabcbcdcbcb$,

$$\Psi_\Sigma(w) = \Psi_\Sigma(v) = \begin{pmatrix} 1 & 2 & 13 & 51 & 36 \\ 0 & 1 & 11 & 61 & 51 \\ 0 & 0 & 1 & 10 & 11 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then $[w] = \{w, v\}$ is not trivially $\mathbb{P}$-distinguishable.

Four is not good

Example

Consider the ordered alphabet $\Sigma = \{a < b < c < d\}$ and the words $w = bcbabcbcdcbcbbabcbccdccbb$ and $v = cbbabbcbcdccbcbabcbcdcbcb$,

$$\Psi_\Sigma(w) = \Psi_\Sigma(v) = \begin{pmatrix} 1 & 2 & 13 & 51 & 36 \\ 0 & 1 & 11 & 61 & 51 \\ 0 & 0 & 1 & 10 & 11 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then $[w] = \{w, v\}$ is not trivially $\mathbb{P}$-distinguishable.

However, $[w]$ is $\mathbb{P}$-distinguishable since $|w|_x \neq |v|_x$, for $x \in \{ac, bd, acd, abd\}$.

Hamming distances for Parikh equivalence classes [Atanasiu et al., 2022]

Definition

Let Σ be an alphabet and $w, w' \in \Sigma^*$ with $|w| = |w'|$. The *Hamming distance* between w and w', denoted $d_H(w, w')$, is the number of positions in which the corresponding letters of w and w' differ.

Definition

Let $\Sigma = \{a_1 < a_2 < \cdots < a_s\}$ and take $w \in \Sigma^*$ to be ambiguous. The *minimal Hamming distance* of $[w]$, denoted $d_H([w])$, is defined by:

$$d_H([w]) = \min\{\, d_H(w', w'') \mid w', w'' \in [w] \text{ and } w' \neq w'' \,\}.$$

Minimal Hamming distance

Example

We have $d_H(aaaacb, aaacab) = 2$ and $d_H([aaaacb]) = 2$, since

$$\psi(aaaacb) = \psi(aaacab) = \psi(aacaab) = \psi(acaaab) = \psi(caaaab).$$

Minimal Hamming distance

Example

We have $d_H(aaaacb, aaacab) = 2$ and $d_H([aaaacb]) = 2$, since

$$\psi(aaaacb) = \psi(aaacab) = \psi(aacaab) = \psi(acaaab) = \psi(caaaab).$$

Then $d_H([aaabbaac]) = 4$, since

$$\psi(aaabbaac) = \psi(aabaabac) = \psi(abaaaabc).$$

Minimal Hamming distance

Example

We have $d_H(aaaacb, aaacab) = 2$ and $d_H([aaaacb]) = 2$, since

$$\psi(aaaacb) = \psi(aaacab) = \psi(aacaab) = \psi(acaaab) = \psi(caaaab).$$

Then $d_H([aaabbaac]) = 4$, since

$$\psi(aaabbaac) = \psi(aabaabac) = \psi(abaaaabc).$$

And then $d_H([abbac]) = 2$, since

$$\psi(abbac) = \psi(abbca) = \psi(baabc).$$

Minimal Hamming distance

Example

We have $d_H(aaaacb, aaacab) = 2$ and $d_H([aaaacb]) = 2$, since

$$\psi(aaaacb) = \psi(aaacab) = \psi(aacaab) = \psi(acaaab) = \psi(caaaab).$$

Then $d_H([aaabbaac]) = 4$, since

$$\psi(aaabbaac) = \psi(aabaabac) = \psi(abaaaabc).$$

And then $d_H([abbac]) = 2$, since

$$\psi(abbac) = \psi(abbca) = \psi(baabc).$$

Consider $w = ab^{11}cbabcb^5$. Then $d_H([w]) = 7$, where

$$[w] = \{\, b^i ab^{11-2i} cbab^{1+2i} cb^{5-i} \mid 0 \leq i \leq 5 \,\}$$

see [Atanasiu et al., 2022, Example 5.1 & Theorem 5.2].

Second Main Result

Theorem (Conjectured in [Atanasiu et al., 2022])

Let $w \in \Sigma_3^*$ be ambiguous. Then $d_H([w]) \in \{2, 4, 7, 8\}$.

Proof.

$[w]$ is trivially $\mathbb{P}$-distinguishable if and only if $d_H([w]) = 2$.

Second Main Result

Theorem (Conjectured in [Atanasiu et al., 2022])

Let $w \in \Sigma_3^*$ be ambiguous. Then $d_H([w]) \in \{2, 4, 7, 8\}$.

Proof.

$[w]$ is trivially $\mathbb{P}$-distinguishable if and only if $d_H([w]) = 2$.

If $[w]$ is not trivially $\mathbb{P}$-distinguishable, then

Second Main Result

Theorem (Conjectured in [Atanasiu et al., 2022])

Let $w \in \Sigma_3^*$ be ambiguous. Then $d_H([w]) \in \{2, 4, 7, 8\}$.

Proof.

$[w]$ is trivially $\mathbb{P}$-distinguishable if and only if $d_H([w]) = 2$.

If $[w]$ is not trivially $\mathbb{P}$-distinguishable, then

- clearly, $|\text{print}(\pi_{a,c}(w))| \geq 1$ or else w is unambiguous;

Second Main Result

Theorem (Conjectured in [Atanasiu et al., 2022])

Let $w \in \Sigma_3^*$ be ambiguous. Then $d_H([w]) \in \{2, 4, 7, 8\}$.

Proof.

$[w]$ is trivially $\mathbb{P}$-distinguishable if and only if $d_H([w]) = 2$.

If $[w]$ is not trivially $\mathbb{P}$-distinguishable, then

- clearly, $|\text{print}(\pi_{a,c}(w))| \geq 1$ or else w is unambiguous;
- using previous Main Result we get $|\text{print}(\pi_{a,c}(w))| \leq 4$;

Second Main Result

Theorem (Conjectured in [Atanasiu et al., 2022])

Let $w \in \Sigma_3^*$ be ambiguous. Then $d_H([w]) \in \{2, 4, 7, 8\}$.

Proof.

$[w]$ is trivially $\mathbb{P}$-distinguishable if and only if $d_H([w]) = 2$.

If $[w]$ is not trivially $\mathbb{P}$-distinguishable, then

- clearly, $|\text{print}(\pi_{a,c}(w))| \geq 1$ or else w is unambiguous;
- using previous Main Result we get $|\text{print}(\pi_{a,c}(w))| \leq 4$;
- if $\text{print}(\pi_{a,c}(w)) \in \{a, c\}$, then $w \in \{a, b\}^*$ or $w \in \{b, c\}^*$, and $d_H([w]) = 4$;

Second Main Result

Theorem (Conjectured in [Atanasiu et al., 2022])

Let $w \in \Sigma_3^*$ be ambiguous. Then $d_H([w]) \in \{2, 4, 7, 8\}$.

Proof.

$[w]$ is trivially $\mathbb{P}$-distinguishable if and only if $d_H([w]) = 2$.

If $[w]$ is not trivially $\mathbb{P}$-distinguishable, then

- clearly, $|\text{print}(\pi_{a,c}(w))| \geq 1$ or else w is unambiguous;
- using previous Main Result we get $|\text{print}(\pi_{a,c}(w))| \leq 4$;
- if $\text{print}(\pi_{a,c}(w)) \in \{a, c\}$, then $w \in \{a, b\}^*$ or $w \in \{b, c\}^*$, and $d_H([w]) = 4$;
- if $|\text{print}(\pi_{a,c}(w))| \in \{2, 3\}$, then $d_H([w]) = 4$ unless w is unambiguous;

Second Main Result

Theorem (Conjectured in [Atanasiu et al., 2022])

Let $w \in \Sigma_3^*$ be ambiguous. Then $d_H([w]) \in \{2, 4, 7, 8\}$.

Proof.

$[w]$ is trivially $\mathbb{P}$-distinguishable if and only if $d_H([w]) = 2$.

If $[w]$ is not trivially $\mathbb{P}$-distinguishable, then

- clearly, $|\text{print}(\pi_{a,c}(w))| \geq 1$ or else w is unambiguous;
- using previous Main Result we get $|\text{print}(\pi_{a,c}(w))| \leq 4$;
- if $\text{print}(\pi_{a,c}(w)) \in \{a, c\}$, then $w \in \{a, b\}^*$ or $w \in \{b, c\}^*$, and $d_H([w]) = 4$;
- if $|\text{print}(\pi_{a,c}(w))| \in \{2, 3\}$, then $d_H([w]) = 4$ unless w is unambiguous;
- if $|\text{print}(\pi_{a,c}(w))| = 4$, then $d_H([w]) \in \{4, 7, 8\}$ unless w is unambiguous. □

Conclusions (how come 3?

It was conjectured in [Şerbănuţă and Şerbănuţă, 2006], that for any $u, v \in \Sigma^*$ and $a \in \Sigma$, for ordered Σ, if uav is ambiguous, then $uaav$ is also ambiguous.

Conclusions (how come 3?

It was conjectured in [Şerbănuţă and Şerbănuţă, 2006], that for any $u, v \in \Sigma^*$ and $a \in \Sigma$, for ordered Σ, if uav is ambiguous, then $uaav$ is also ambiguous. True for the ternary, but counterexamples exist for quaternary alphabets [Teh et al., 2018] (see also [Poovanandran and Teh, 2020]).

Conclusions (how come 3?

It was conjectured in [Şerbănuţă and Şerbănuţă, 2006], that for any $u, v \in \Sigma^*$ and $a \in \Sigma$, for ordered Σ, if uav is ambiguous, then $uaav$ is also ambiguous. True for the ternary, but counterexamples exist for quaternary alphabets [Teh et al., 2018] (see also [Poovanandran and Teh, 2020]).

Injectivity settled for ternary alphbets [Baek et al., 2024], [Salomaa, 2010]

Conclusions (how come 3?

It was conjectured in [Şerbănuţă and Şerbănuţă, 2006], that for any $u, v \in \Sigma^*$ and $a \in \Sigma$, for ordered Σ, if uav is ambiguous, then $uaav$ is also ambiguous.
True for the ternary, but counterexamples exist for quaternary alphabets [Teh et al., 2018] (see also [Poovanandran and Teh, 2020]).

Injectivity settled for ternary alphbets [Baek et al., 2024], [Salomaa, 2010], but only with rational exponents for higher alphabets [Hahn et al., 2023].

Conclusions (how come 3?

It was conjectured in [Şerbănuță and Şerbănuță, 2006], that for any $u, v \in \Sigma^*$ and $a \in \Sigma$, for ordered Σ, if uav is ambiguous, then $uaav$ is also ambiguous. True for the ternary, but counterexamples exist for quaternary alphabets [Teh et al., 2018] (see also [Poovanandran and Teh, 2020]).

Injectivity settled for ternary alphbets [Baek et al., 2024], [Salomaa, 2010], but only with rational exponents for higher alphabets [Hahn et al., 2023].

Theorem (for $w \in \Sigma_3^*$)

$[w]$ is $\mathbb{P}$-distinguishable if and only if $[w]$ is trivially $\mathbb{P}$-distinguishable.

Future work

- Figure a way to translate the injectivity to integer coefficients.
- Can we do upward projections (and how to define these)?
- Find the minimal Hamming distances for alphabets of size 4 or higher.
- Is there a fixed upper bound on the minimal Hamming distance of any Parikh equivalence class, for any ordered alphabet?

Future work

- Figure a way to translate the injectivity to integer coefficients.
- Can we do upward projections (and how to define these)?
- Find the minimal Hamming distances for alphabets of size 4 or higher.
- Is there a fixed upper bound on the minimal Hamming distance of any Parikh equivalence class, for any ordered alphabet?

Thanks, questions?