Partial Permutations Comparison, Maintenance and Applications

Avivit Levy*, Ely Porat# and B. Riva Shalom*

***Shenkar College, # Bar Ilan University**

CPM 2022

Outline

- Definitions
- Applications
- Negative Result – Reduction to OV
- Positive Results

Definition: Permutation

A permutation **is a bijection** $\Pi : \Sigma \rightarrow \Sigma$ mapping every symbol $\sigma_i \in \Sigma$ to a distinct symbol $\sigma_j \in \Sigma$.

$$\pi = \begin{pmatrix} a & b & c & d & e & f & g \\ a & c & e & g & b & d & f \end{pmatrix}$$

Definition: Partial Permutation

A partial permutation **is a bijection**

$\pi_{\mathrm{par}} : \Sigma_1 \rightarrow \Sigma_2$ mapping a subset $\Sigma_1 \subset \Sigma$ to a subset $\Sigma_2 \subset \Sigma$, where $|\Sigma_1| = |\Sigma_2|$.

$$\pi_{\mathrm{par}} = \begin{pmatrix} a & b & c & d & e & f & g \\ a & - & e & - & b & - & - \end{pmatrix}$$

dimension =7

Partial Permutations Comparison and Maintenance

Input: Sets A_1, A_2 of d-dimensional partial permutations over alphabet Σ.

Goals: **-** Be able to compare π_1 and π_2 , $\pi_1 \in A_1$ $\pi_2 \in A_2$.

- Maintain a dynamic set of partial permutations, supporting search operations.

Applications

- Gene Sequence Comparisons.
- Color Transformations as Data Augmentation Technique.
- Strict Parameterized Dictionary Matching with One Gap.

Comparison of Partial Permutations

$$\pi_1 = \begin{pmatrix} a & b & c & d & e & f & g \\ a & - & c & d & - & - & - \end{pmatrix} \quad \pi_2 = \begin{pmatrix} a & b & c & d & e & f & g \\ a & - & c & d & - & - & - \end{pmatrix}$$

$$\pi_1 = \begin{pmatrix} a & b & c & d & e & f & g \\ a & - & c & d & - & - & - \end{pmatrix} \quad \pi_2 = \begin{pmatrix} a & b & c & d & e & f & g \\ a & e & c & - & g & - & b \end{pmatrix}$$

Definition:
Conflict of Partial Permutations

Partial permutations $\pi_{1,}\ \pi_2$ have a ***conflict*** if either:

- $\exists$ i,j, i$\neq$j, s.t. $\exists\ \sigma_k$ where $\pi_1(\sigma_k) = \sigma_i$ and $\pi_2(\sigma_k) = \sigma_j$
- $\exists$ i,j, i$\neq$j, s.t. $\exists\ \sigma_k$ where $\pi_1(\sigma_i) = \pi_2(\sigma_j) = \sigma_k$

Example of Conflicting Partial Permutations

$$\pi_1 = \begin{pmatrix} a & \mathbf{b} & c & d & e & f & g \\ a & \mathbf{b} & c & f & - & - & - \end{pmatrix} \quad \pi_2 = \begin{pmatrix} a & \mathbf{b} & c & d & e & f & g \\ - & \mathbf{e} & - & - & g & - & b \end{pmatrix}$$

$$\pi_1 = \begin{pmatrix} a & b & c & d & e & f & g \\ a & - & \mathbf{c} & d & - & - & - \end{pmatrix} \quad \pi_2 = \begin{pmatrix} a & b & c & d & e & f & g \\ - & \mathbf{c} & - & - & g & - & b \end{pmatrix}$$

Universal Condition of Partial Permutation Agreement

Lemma :

Let $D(\pi)$ be the domain of π, and $C(\pi)$ be the codomain of π.

Partial permutations $\pi_{1,}\, \pi_2$ ***agree*** iff :

- $\forall\ \sigma_k \in D(\pi_1) \cap D(\pi_2)$ if $\pi_1(\sigma_k) = \sigma_i$ then $\pi_2(\sigma_k) = \sigma_i$
- $\forall\ \sigma_k \in C(\pi_1) \cap C(\pi_2)$ if $\pi_1(\sigma_i) = \sigma_k$ then $\pi_2(\sigma_i) = \sigma_k$

Example of Partial Permutations Agreement

$$\pi_1 = \begin{pmatrix} a & b & c & d & e & f & g \\ a & - & c & d & - & - & - \end{pmatrix} \quad \pi_2 = \begin{pmatrix} a & b & c & d & e & f & g \\ - & e & c & - & g & - & b \end{pmatrix}$$

$\pi_{1,}\, \pi_2$ agree

Representing Partial Permutations

A partial permutation π can be represented as a string S_π of length $|\Sigma|$. $\underline{\mathbf{S_\pi = \pi(\sigma_1)\ \pi(\sigma_2) \ldots\ \pi(\sigma_{|\Sigma|})}}$**,**

where $\pi(\sigma_i) = *$ (a don't care symbol) if $\sigma_i \notin D(\pi)$

Example: Let $\pi = \begin{pmatrix} a & b & c & d & e & f & g \\ - & e & - & - & g & - & b \end{pmatrix}$

We have that $S_\pi = * \, e * * \, g * b$

Representing Partial Permutations

Let S_{π_1} S_{π_2} be strings representing partial permutations. It is possible that S_{π_1} matches S_{π_2} yet the permutations they represent have a conflict.

$$\pi_1 = \begin{pmatrix} a & b & c & d & e & f \\ e & c & \mathbf{d} & - & - & - \end{pmatrix} \quad \pi_2 = \begin{pmatrix} a & b & c & d & e & f \\ e & c & - & - & \mathbf{d} & - \end{pmatrix}$$

$$S_{\pi_1} = e\ c\ d * * * \quad \text{matches}$$
$$S_{\pi_2} = e\ c * * d *$$

Definition: **The Inverse of a Partial Permutation**

Let π be a partial permutation over Σ, where $\pi:\Sigma_1 \rightarrow \Sigma_2$, $\Sigma_1 \subset \Sigma$, $\Sigma_2 \subset \Sigma$ and $|\Sigma_1| = |\Sigma_2|$.

The inverse of a partial permutation is the bijection π: $\Sigma_2 \rightarrow \Sigma_1$ s.t. $\forall \sigma_i \in \Sigma_2$, $\pi^{-1}(\sigma_i) = \sigma_j$ iff $\pi(\sigma_j) = \sigma_i$

$$\Pi = \begin{pmatrix} a & b & c & d & e & f & g \\ - & e & - & - & g & - & b \end{pmatrix} \qquad \Pi^{-1} = \begin{pmatrix} a & b & c & d & e & f & g \\ - & g & - & - & b & - & e \end{pmatrix}$$

Representing Partial Permutation Agreement

Lemma : Given partial permutations π_1, π_2 then,

$\mathbf{S_{\pi_1} \circ S_{\pi_1}^{-1}}$ **matches** $\mathbf{S_{\pi_2} \circ S_{\pi_2}^{-1}}$ iff π_1 and π_2 agree.

[$\circ$ denotes concatenation]

$$\pi_1 = \begin{pmatrix} a & b & c & d & e & f \\ e & c & \mathbf{d} & - & - & - \end{pmatrix} \qquad \pi_2 = \begin{pmatrix} a & b & c & d & e & f \\ e & c & - & - & \mathbf{d} & - \end{pmatrix}$$

$$S_{\pi_1} \circ S_{\pi_1}^{-1} = e\ c\ d * * * \circ * * b\ c\ a\ *$$

$$S_{\pi_2} \circ S_{\pi_2}^{-1} = e\ c * * d * \circ * * b\ e\ a\ *$$

The Partial Permutations Agreement Problem (PPA)

Input: Sets A_1, A_2 of d-dimensional partial permutations over alphabet Σ, where $|A_1| = |A_2| = n$.

Output: All pairs (π_1, π_2), $\pi_1 \in A_1$ $\pi_2 \in A_2$ where π_1 and π_2 agree.

If $|A_1| = n$ and $|A_2| = 1$, the problem is called the Single query PPA problem (SPPA)

Naïve Solution for PPA Problem

For each pair (π_1, π_2), $\pi_1 \in A_1$, $\pi_2 \in A_2$

Match $S_{\pi_1} \circ S_{\pi_1}^{-1}$ with $S_{\pi_2} \circ S_{\pi_2}^{-1}$.

Time: $O(n^2|\Sigma|)$

Space: $O(n|\Sigma|)$

The Orthogonal Vectors Problem (OV)

Input: Sets S_1, S_2 of binary d-dimensional vectors where $|S_1| = |S_2| = n$, $d=\omega(\log n)$.

Output: Decide whether there exist $v_1 \in S_1$, $v_2 \in S_2$, where $v_1 \cdot v_2 = \sum_{i=1}^{d} v_1[i] \cdot v_2[i] = 0$.

Fine Grained Complexity

Theorem [Impagliazzo & Paturi, 2001]:

Given sets S_1, S_2 of binary d-dimensional vectors, where $|S_1|=|S_2|= n$ and $d = \Theta(\log n)$, there does not exist $\varepsilon>0$ s.t. OV is solvable in $O(n^{2-\varepsilon})$ time unless the Strong Exponential Time Hypothesis (SETH) is false.

Lemma

OV is reducible to Partial Permutations Agreement (PPA) in $O(nd)$ time and space.

Given $(S_1, S_2, n, d) \in OV$ we construct A_1, A_2, Σ where $v_1 \in S_1, v_2 \in S_2$ s. t. v_1, v_2 are orthogonal iff there are $\pi_i \in A_1, \pi_j \in A_2$ s. t. π_1 agree with π_2.

The Reduction

Given sets S_1, S_2 of n d-dimensional binary vectors, we construct sets A_1, A_2 of partial permutations over Σ.

We set $|\Sigma| = d+1$.

$\forall v_i = (b^i_1, b^i_2, \ldots, b^i_d) \in S_x$ we construct a partial permutation gadget including the mapping of $\sigma_k \in \Sigma$ iff $b^i_k = 1$ $\forall 1 \leq k \leq d$

$v_i = (1, 0, 1, 1, 0, 0)$

d =6

$$\pi_{v_i} = \begin{pmatrix} a & b & c & d & e & f & g \\ ? & - & ? & ? & - & - & - \end{pmatrix}$$

The Construction for S_1 Vectors

For $v_i = (b^i_1, b^i_2, \ldots, b^i_d) \in S_1$, their associated gadget π_{v_i} has the following mapping:

- $\forall \sigma_k \in \Sigma$ if $b^i_k = 1$ $\forall 1 \le k \le d$, then $\boldsymbol{\pi_{v_i}(\sigma_k) = \sigma_k}$
- Consider $b^i_{d+1} = 0$

$$v_i = (1, 0, 1, 1, 0, 0) \in S_1 \Rightarrow \pi_{v_i} = \begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{pmatrix} a & b & c & d & e & f & g \\ a & - & c & d & - & - & - \end{pmatrix} \end{array}$$

The Construction for S_2 Vectors

For $v_j = (b^j_1, b^j_2, \ldots, b^j_d) \in S_2$, their associated gadget π_{v_j} has the following mapping:

- $\forall \sigma_k \in \Sigma$ if $b^j_k = 1$ $\forall 1 \le k \le d$, then $\boldsymbol{\pi_{v_j}(\sigma_k) = \sigma_g}$ where σ_g is the symbol that is cyclically to the right of σ_k in the sorting of $D(\pi_{v_j})$
- Consider $b^j_{d+1} = 1$

$$v_j = (0, 1, 0, 0, 1, 0) \in S_2 \Rightarrow \pi_{v_j} = \begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{pmatrix} a & b & c & d & e & f & g \\ - & e & - & - & g & - & b \end{pmatrix} \end{array}$$

The Reduction Proof

- Let $v_1 \in S_1$, $v_2 \in S_2$ be **non-orthogonal**. Thus, $v_1 \cdot v_2 \geq 1$ implying there exists at least one index k, s.t. $b^1_k = 1$ and $b^2_k = 1$.
- According to the construction $\boldsymbol{\pi_{v_1}(\sigma_k) = \sigma_k}$
- Yet, $\pi_{v_2}(\sigma_k) = \sigma_g$ where σ_g is the symbol that is cyclically to the right of σ_k in the sorting of $D(\pi_{v_2})$
- Due to the addition of σ_{d+1} only to π_{v_2}, $\pi_{v_2}(\sigma_k) \neq \sigma_k$
- Hence, $\boldsymbol{\pi_{v_1}(\sigma_k) \neq \pi_{v_2}(\sigma_k)}$ and there is **a conflict.**

Example of the Reduction Proof

- Let $v_1 \in S_1$, $v_2 \in S_2$ **be non-orthogonal**. Thus, $v_1 \cdot v_2 \geq 1$ implying there exist at least one index k, s.t. $b^1{}_k = 1$ and $b^2{}_k = 1$.

$$v_1 = (1, \boxed{1}, 1, 1, 0, 0) \in S_1 \qquad v_2 = (0, \boxed{1}, 0, 0, 0, 0) \in S_2$$

$$\Downarrow \qquad\qquad\qquad\qquad \Downarrow$$

$$\pi_{v_1} = \begin{pmatrix} a & \boxed{\mathbf{b}} & c & d & e & f & g \\ a & \boxed{\mathbf{b}} & c & d & - & - & - \end{pmatrix} \qquad \pi_{v_2} = \begin{pmatrix} a & \boxed{\mathbf{b}} & c & d & e & f & g \\ - & \boxed{g} & - & - & - & - & b \end{pmatrix}$$

The Reduction Proof

- Let $v_1 \in S_1$, $v_2 \in S_2$ **be orthogonal**. Thus, $v_1 \cdot v_2 = 0$ implying at most one of $\{b^1_k, b^2_k\}$ equals 1, $\forall\ 1 \leq k \leq d$.

According to the construction: If $b^1_k = 1$ then $\boldsymbol{\pi_{v_1}(\sigma_k) = \sigma_k}$

Yet, $b^2_k = 0$, thus $\sigma_k \notin D(\pi_{v_2})$.

It is not possible that $\boldsymbol{\pi_{v_2}(\sigma_g) = \sigma_k}$, for $\sigma_g \neq \sigma_k$, as $\boldsymbol{\pi_{v_2}}$ maps $\boldsymbol{\sigma_g}$ to a symbol that is in the sorting of $D(\pi_{v_j})$, and $D(\pi_{v_2}) = C(\pi_{v_2})$

Hence, there is no conflict.

Example of the Reduction Proof

- Let $v_1 \in S_1$, $v_2 \in S_2$ **be orthogonal**. Thus, $v_1 \cdot v_2 = 0$ implying at most one of $\{ b^1_k , b^2_k \}$ equals 1, $\forall\ 1 \le k \le d$.
- We get that the gadget partial permutations agree.

$$v_1 = (1, 0, 1, 1, 0, 0) \in S_1 \qquad v_2 = (0, 1, 0, 0, 1, 0) \in S_2$$

$$\Downarrow \qquad\qquad\qquad \Downarrow$$

$$\pi_{v_1} = \begin{pmatrix} a & b & c & d & e & f & g \\ a & - & c & d & - & - & - \end{pmatrix} \qquad \pi_{v_2} = \begin{pmatrix} a & b & c & d & e & f & g \\ - & e & - & - & g & - & b \end{pmatrix}$$

Results

Positive:		Negative:
Almost Full Perm.	Partial Permutations	
PPA query: $O(n \cdot poly(\|\Sigma\|))$	PPA query: $O(n \cdot poly(\|\Sigma\|))$	If PPA is solvable in $O(n^{2-\varepsilon}\|\Sigma\|)$ time ($\varepsilon>0$)
Update $O(poly(\|\Sigma\|))$	Update $O(poly(\|\Sigma\|))$	and $O(n\|\Sigma\|)$ space,
Space: $O(n \cdot poly(\|\Sigma\|))$	Space: $O(n \cdot poly(\|\Sigma\|))$	then SETH is false.

Open Issues

- Is there an efficient solution for PPA for special cases?
- It seems that partial permutations play a hidden role in more applications. It may be interesting to explore their algorithmic usage.

Thank You!