Combinatória e Rearranjo de Genoma

VI Semana da Matemática da UFF

Luis Antonio Brasil Kowada - *GGM, UFF*
Luís Felipe Ignácio Cunha - *PESC, UFRJ*

11 de maio de 2012

"Durante a evolução, moléculas de hereditariedade são costuradas, modificadas, cortadas, alongadas, encurtadas e revertidas." (François Jacob)

“Durante a evolução, moléculas de hereditariedade são costuradas, modificadas, cortadas, alongadas, encurtadas e revertidas.” (François Jacob)

Alguns poucos conceitos Biológicos

- Moléculas de DNA são responsáveis por toda informação genética dos seres.
- De uma célula para outra, uma proteína para outra, conteúdos de DNA são quase similares, porém suas organizações se diferem drasticamente.
- Mutações afetam a organização do DNA. Mutações são chamadas de **Rearranjo de Genomas**.

Genoma?

- Segmentos de genes que contém informações necessárias para construção de moléculas numa célula.
- Gene é uma sequência de nucleotídeos (A, T, C, G) que codifica proteína.
- Cromossomo é uma ordenação do conjunto de genes.
- Cromossomos são feitos de DNA.
- Genoma é um conjunto de cromossomos.

E daí?

Relação com a Matemática

- Entender a **diferença (ou similaridade) estrutural** entre genomas motiva o estudo combinatório.
- Eventos mutacionais são poucos em relação ao tempo de evolução.
- Um estudo de **otimização combinatória** é minimizar o número de mutações que um genoma sofreu para gerar outro.

E daí?

Relação com a Matemática

- Entender a **diferença (ou similaridade) estrutural** entre genomas motiva o estudo combinatório.
- Eventos mutacionais são poucos em relação ao tempo de evolução.
- Um estudo de **otimização combinatória** é minimizar o número de mutações que um genoma sofreu para gerar outro.

E daí?

Relação com a Matemática

- Entender a **diferença (ou similaridade) estrutural** entre genomas motiva o estudo combinatório.
- Eventos mutacionais são poucos em relação ao tempo de evolução.
- Um estudo de **otimização combinatória** é minimizar o número de mutações que um genoma sofreu para gerar outro.

Comparação entre Sequências
Alinhamento de Sequências

Alinhamento de Sequências = Encontrar similaridades entre sequências.

Contextualização:

- Dois genomas são sequenciados por dois laboratórios e o objetivo é comparar os resultados;
- Encontrar diferenças entre duas sequências.

Comparação entre Sequências

Alinhamento de Sequências ou Distância de Edição

- Operações entre nucleotídeos: Emparelhamento, Substituição, Inserção e Deleção.
- Pesos são associados a cada operação: Emparelhamentos $= +1$, Substituição $= -1$, Inserção e Deleção $= -2$. (Parâmetros usados na prática)
- Alinhamento ótimo $= \max\{\sum \text{pesos}\}$. Construído por um algoritmo de "Programação Dinâmica".

Example

GA - CGGA**T**TAG
GA**T**CGGA**A**TAG

$9x1 + 1x(-1) + 1x(-2) = 6.$

Rearranjo entre genes

Eventos Globais

- O "grau de desorientação" entre a organização de genes de diferentes genomas indica a distância evolucionária entre organismos. (Dobzhansky e Sturtevant – 1936)
- Representaçao das posições relativas dos genes em diferentes genomas como **permutações**. (Watterson *et al.* – década 80)
- Transformar uma permutação em outra com o número mínimo de **operações**. (Matemáticos já utilizavam este modelo com outras motivações)

Rearranjo entre genes

Eventos Globais

- O "grau de desorientação" entre a organização de genes de diferentes genomas indica a distância evolucionária entre organismos. (Dobzhansky e Sturtevant – 1936)
- Representaçao das posições relativas dos genes em diferentes genomas como **permutações**. (Watterson *et al.* – década 80)
- Transformar uma permutação em outra com o número mínimo de **operações**. (Matemáticos já utilizavam este modelo com outras motivações)

Rearranjo entre genes

Eventos Globais

- O "grau de desorientação" entre a organização de genes de diferentes genomas indica a distância evolucionária entre organismos. (Dobzhansky e Sturtevant – 1936)
- Representaçao das posições relativas dos genes em diferentes genomas como **permutações**. (Watterson *et al.* – década 80)
- Transformar uma permutação em outra com o número mínimo de **operações**. (Matemáticos já utilizavam este modelo com outras motivações)

Rearranjo entre genes

Modelo

1. É conhecida a ordem dos genes em cada genoma;
2. Todos genomas compartilham o mesmo conjunto de genes;
3. Todos os genomas contém uma única cópia de cada gene;
4. Todos genomas possuem um único cromossomo.

Genomas são modelados por permutações. Mas o que são permutações?

Rearranjo entre genes

Modelo

1. É conhecida a ordem dos genes em cada genoma;
2. Todos genomas compartilham o mesmo conjunto de genes;
3. Todos os genomas contém uma única cópia de cada gene;
4. Todos genomas possuem um único cromossomo.

Genomas são modelados por permutações. Mas o que são permutações?

Rearranjo entre genes

Permutações

Definition

Uma Permutação $\pi_{[n]}$ é uma função bijetiva com domínio no conjunto $\{1, 2, \ldots, n\}$ e imagem em $\{1, 2, \ldots, n\}$. $\pi_{[n]} \in S_n$.

$$\pi_{[n]} = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \pi_1 & \pi_2 & \pi_3 & \ldots & \pi_n \end{pmatrix} = [\pi_1\, \pi_2\, \pi_3 \ldots \pi_n].$$

Obs.: $\iota_{[n]} = [1\,2 \ldots n]$ (Permutação Identidade).
$\rho_{[n]} = [n\; n{-}1 \ldots 1]$ (Permutação Reversa).

Example

$$\pi_{[5]} = [3\,2\,5\,4\,1].$$

Rearranjo entre genes

Permutações

Definition

Uma Permutação com sinal $\pi_{[n]}$ é uma permutação no conjunto $\{-n, \dots, -2, -1, 1, 2, \dots, n\}$ e satisfaz $\pi_{-i} = -\pi_i$. $\pi_{[n]} \in S_n^{\pm}$.

$$\pi_{[n]} = \begin{pmatrix} -n & \dots & -2 & -1 & 1 & 2 & \dots & n \\ -\pi_n & \dots & -\pi_2 & -\pi_1 & \pi_1 & \pi_2 & \dots & \pi_n \end{pmatrix} \equiv [\pi_1\, \pi_2\, \pi_3 \dots \pi_n].$$

Example

$$\pi_{[5]} = [-3\ \ -2\ 5\ 4\ \ -1].$$

Reversão

Definition

Uma reversão $\rho(i,j)$, com $1 \leq i < j \leq n$, aplicada numa permutação $\pi_{[n]} \in S_n$, inverte o intervalo fechado delimitado por i e j. Transformando $\pi_{[n]}$ em $\pi_{[n]}\rho(i,j)$.

Uma reversão $\rho(i,j)$, com $1 \leq i \leq j \leq n$, aplicada numa permutação $\pi_{[n]} \in S_n^{\pm}$, inverte o intervalo fechado delimitado por i e j alterando o sinal de cada elemento. Transformando $\pi_{[n]}$ em $\pi_{[n]}\rho(i,j)$.

Example

$$\pi_{[5]} = [3\,2\,5\,4\,1].$$
$$\pi_{[5]}\rho(2,4) = [3\,4\,5\,2\,1].$$

$$\pi_{[5]} = [-3\;-2\,5\,4\;-1].$$
$$\pi_{[5]}\rho(2,4) = [-3\;-4\;-5\,2\;-1].$$

Distância de Reversão

Problema

- Distância de Reversão = Dadas duas permutações $\pi, \sigma \in S_n$, determinar o menor número de reversões r, tais que, $\pi\rho_1\rho_2...\rho_r = \sigma$. Assim, $d_r(\pi, \sigma) = r$.
- Equivalente a determinarmos a menor sequência de π à ι. Denotamos $d_r(\pi, \iota)$ por $d_r(\pi)$.

Distância de Reversão

Exemplo

Example

$$\begin{aligned}
\pi &= [4\ 5\ \underline{8\ 7\ 1\ 2\ 3\ 6}],\\
\pi\rho(3,8) &= [\underline{4\ 5\ 6\ 3\ 2\ 1}\ 7\ 8],\\
\pi\rho(3,8)\rho(1,6) &= [1\ 2\ 3\ \underline{6\ 5\ 4}\ 7\ 8],\\
\pi\rho(3,8)\rho(1,6)\rho(4,6) &= [1\ 2\ 3\ 4\ 5\ 6\ 7\ 8].
\end{aligned}$$

$d_r(\pi) \leq 3$.

Distância de Reversão

Como conseguir a solução do problema de maneira eficiente?

PERGUNTA: Dada uma permutação π e um inteiro k, consigo dizer de maneira eficiente se $d_r(\pi) \leq k$?

Distância de Reversão

Como conseguir a solução do problema de maneira eficiente?

PERGUNTA: Dada uma permutação π e um inteiro k, consigo dizer de maneira eficiente se $d_r(\pi) \leq k$?

RESPOSTA: **Não Sei!!!!!**

Distância de Reversão

Como conseguir a solução do problema de maneira eficiente?

Eu não sei mas nenhuma dessas pessoas sabem.

Distância de Reversão

Limites para Distância

- Dada uma permutação π, sei apresentar limites para $d_r(\pi)$.
- COMO?

Distância de Reversão

Limites para Distância

- Dada uma permutação π, sei apresentar limites para $d_r(\pi)$.
- COMO?

Distância de Reversão

Limites para Distância

Definition

Dada uma permutação $\pi_{[n]}$, um par (π_i, π_{i+1}) com $0 \leq i \leq n$ é chamado de strong breakpoint se,

$$\begin{aligned} \pi_{i+1} &\neq \pi_i + 1, \text{ e} \\ \pi_{i+1} &\neq \pi_i - 1. \end{aligned}$$

$sb(\pi)$ é o número de strong breakpoints que há em π. Indicamos o par (π_i, π_{i+1}) de strong breakpoints por um sinal de $\bullet$ entre π_i e π_{i+1} em π.

Example

$\pi = [4\ 5\ 8\ 7\ 1\ 2\ 3\ 6]$.

$[0 \bullet 4\ 5 \bullet 8\ 7 \bullet 1\ 2\ 3 \bullet 6 \bullet 9]$.

$\{(0,4),\ (5,8),\ (7,1),\ (3,6),\ (6,9)\} \Rightarrow sb(\pi) = 5$.

Distância de Reversão
Limites para Distância

$$\left\lceil \frac{sb(\pi)}{2} \right\rceil \leq d_r(\pi) \leq sb(\pi) - 1.$$

Distância de Reversão

Limites para Distância

Example

$\pi = [0 \ \bullet \ 4 \ 5 \ \bullet 8 \ 7 \ \bullet \ 1 \ 2 \ 3 \bullet \ 6 \ \bullet \ 9]$. $sb(\pi) = 5$, assim $3 \leq d_r(\pi) \leq 4$.

No exemplo anterior ordenamos π utilizando exatamente 3 reversões:

$$
\begin{aligned}
\pi &= [4\ 5\ \underline{8\ 7\ 1\ 2\ 3\ 6}],\\
\pi\rho(3,8) &= [\underline{4\ 5\ 6\ 3\ 2\ 1}\ 7\ 8],\\
\pi\rho(3,8)\rho(1,6) &= [1\ 2\ 3\ \underline{6\ 5\ 4}\ 7\ 8],\\
\pi\rho(3,8)\rho(1,6)\rho(4,6) &= [1\ 2\ 3\ 4\ 5\ 6\ 7\ 8].
\end{aligned}
$$

Assim $d_r(\pi) = 3$.

Como poderíamos ordenar π utilizando 4 reversões?

Distância de Reversão

Limites para Distância

Example

$\pi = [0 \bullet 4\ 5 \bullet 8\ 7 \bullet 1\ 2\ 3 \bullet 6 \bullet 9]$. $sb(\pi) = 5$, assim $3 \leq d_r(\pi) \leq 4$.

No exemplo anterior ordenamos π utilizando exatamente 3 reversões:

$$\begin{aligned}
\pi &= [4\ 5\ \underline{8\ 7\ 1\ 2\ 3\ 6}],\\
\pi\rho(3,8) &= [\underline{4\ 5\ 6\ 3\ 2\ 1}\ 7\ 8],\\
\pi\rho(3,8)\rho(1,6) &= [1\ 2\ 3\ \underline{6\ 5\ 4}\ 7\ 8],\\
\pi\rho(3,8)\rho(1,6)\rho(4,6) &= [1\ 2\ 3\ 4\ 5\ 6\ 7\ 8].
\end{aligned}$$

Assim $d_r(\pi) = 3$.

Como poderíamos ordenar π utilizando 4 reversões?

Distância de Reversão
Limites para Distância

$$\pi = [4\ 5\ 8\ 7\ \underline{1\ 2\ 3\ 6}],$$

$$\pi\rho(5,8) = [\underline{4\ 5\ 8\ 7\ 6\ 3\ 2\ 1}],$$

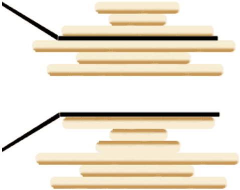
$$\pi\rho(5,8)\rho(1,8) = [1\ 2\ 3\ \underline{6\ 7\ 8\ 5\ 4}],$$

$$\pi\rho(5,8)\rho(1,8)\rho(4,8) = [1\ 2\ 3\ 4\ 5\ \underline{8\ 7\ 6}],$$

$$\pi\rho(5,8)\rho(1,8)\rho(4,8)\rho(6,8) = [1\ 2\ 3\ 4\ 5\ 6\ 7\ 8].$$

Pancake-Flipping

Reversão Pré-fixada

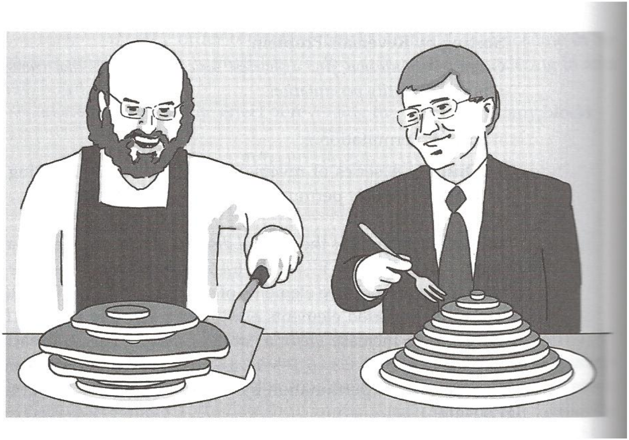


Problem

Um cozinheiro distraído, quando prepara panquecas de diferentes tamanhos as deixa numa ordem não decrescente da base ao topo. Se existem n panquecas de tamanhos distintos, qual o número máximo de rotações necessárias para deixar as panquecas em ordem decrescente da base ao topo?

Pancake-Flipping

Reversão Pré-fixada



Pancake-Flipping

Reversão Pré-fixada

- Uma reversão pré-fixada é uma reversão $\rho(i, j)$ com $i = 1$.
- Distância de Reversão Pré-fixada = Dadas duas permutações $\pi, \sigma \in S_n$. Distância de reversão préfixada é determinar a menor sequência de reversões pré-fixadas, tais que, $\pi\rho_1\rho_2 ... \rho_r = \sigma$. Assim, $d_{rp}(\pi, \sigma) = r$.
- Equivalente a determinarmos a menor sequência de π à ι. Denotamos $d_{rp}(\pi, \iota)$ por $d_{rp}(\pi)$.

Obs.: Nenhum limite que existe para distância de Reversão se aplica para o caso pré-fixado.

Reversão Pré-fixada

Como conseguir a solução do problema de maneira eficiente?

PERGUNTA: Dada uma permutação π e um inteiro k, consigo dizer de maneira eficiente se $d_{rp}(\pi) \leq k$?

Reversão Pré-fixada

Como conseguir a solução do problema de maneira eficiente?

PERGUNTA: Dada uma permutação π e um inteiro k, consigo dizer de maneira eficiente se $d_{rp}(\pi) \leq k$?

RESPOSTA: **Também não sei!!!!**

Distância de Reversão Sinalizada

Problema

- Distância de Reversão Sinalizada = Dadas duas permutações $\pi, \sigma \in S_n^{\pm}$, determinar o menor número de reversões r, tais que, $\pi \rho_1 \rho_2 ... \rho_r = \sigma$. Assim, $d_{rs}(\pi, \sigma) = r$.
- Equivalente a determinarmos a menor sequência de π à ι. Denotamos $d_{rs}(\pi, \iota)$ por $d_{rs}(\pi)$.

Distância de Reversão Sinalizada

Exemplo

Example

$$
\begin{aligned}
\pi &= [-4\ \underline{-5\ 8\ 7\ 1\ -2\ 3}\ 6],\\
\pi\rho(2,7) &= [-4\ -3\ 2\ -1\ -7\ \underline{-8\ 5\ 6}],\\
\pi\rho(2,7)\rho(6,8) &= [\underline{-4\ -3\ 2\ -1}\ -7\ -6\ -5\ 8],\\
\pi\rho(2,7)\rho(6,8)\rho(1,4) &= [1\ -2\ 3\ 4\ \underline{-7\ -6\ -5}\ 8],\\
\pi\rho(2,7)\rho(6,8)\rho(1,4)\rho(5,7) &= [1\ \underline{-2}\ 3\ 4\ 5\ 6\ 7\ 8],\\
\pi\rho(2,7)\rho(6,8)\rho(1,4)\rho(5,7)\rho(2,2) &= [1\ 2\ 3\ 4\ 5\ 6\ 7\ 8].
\end{aligned}
$$

$d_{rs}(\pi) \leq 5$.

Reversão Sinalizada

Como conseguir a solução do problema de maneira eficiente?

PERGUNTA: Dada uma permutação π e um inteiro k, consigo dizer de maneira eficiente se $d_{rs}(\pi) \leq k$?

Reversão Sinalizada

Como conseguir a solução do problema de maneira eficiente?

PERGUNTA: Dada uma permutação π e um inteiro k, consigo dizer de maneira eficiente se $d_{rs}(\pi) \leq k$?

RESPOSTA: **Essa eu sei!!!!**

Para isso, necessitamos de uma *poderosa ferramenta*.

Reversão Sinalizada

Como conseguir a solução do problema de maneira eficiente?

PERGUNTA: Dada uma permutação π e um inteiro k, consigo dizer de maneira eficiente se $d_{rs}(\pi) \leq k$?

RESPOSTA: **Essa eu sei!!!!**

Para isso, necessitamos de uma *poderosa ferramenta*.

Diagrama Realidade e Desejo

Definição

Definition

Diagrama Realidade e Desejo de uma permutação $\pi_{[n]}$, $RD(\pi)$ = Multigrafo $RD(\pi) = (V, R \cup D)$ onde:

$$
\begin{array}{lll}
V & := & \{0, -\pi_1, +\pi_1, -\pi_2, +\pi_2, \cdots, -\pi_n, +\pi_n, -(n+1)\}; \\
 & & \textit{Arestas de Realidade:} \\
R & := & (0, -\pi_1),\ \text{se}\ \pi_1 > 0, ((0, +\pi_1),\ \text{se}\ \pi_1 < 0) \cup \\
 & & (+\pi_i, -\pi_{i+1}),\ \text{se}\ \pi_i, \pi_{i+1} > 0, (-\pi_i, +\pi_{i+1}),\ \text{se}\ \pi_i, \pi_{i+1} < 0 \cup \\
 & & (-\pi_i, -\pi_{i+1}),\ \text{se}\ \pi_i < 0\ \text{e}\ \pi_{i+1} > 0, (+\pi_i, +\pi_{i+1}),\ \text{se}\ \pi_i > 0\ \text{e}\ \pi_{i+1} < 0; \\
 & & \textit{Arestas de Desejo:} \\
D & := & \{(+i, -(i+1)) | i = 1, ..., n-1\} \cup \{(0, -1), (+n, -(n+1))\}.
\end{array}
$$

- Tamanho de um ciclo = número de arestas de realidade do ciclo.
- $c(\pi)$ = número de ciclos que existem em $RD(\pi)$.

Diagrama Realidade e Desejo
Exemplo

Example

$$\pi = [-4\ \ -5\ 8\ 7\ 1\ \ -2\ 3\ 6],$$

Diagrama Realidade e Desejo

Exemplo

Example

$$\pi = [-4\ \ -5\ 8\ 7\ 1\ \ -2\ 3\ 6],$$

0 +4 -4 +5 -5 -8 +8 -7 +7 -1 +1 +2 -2 -3 +3 -6 +6 -9

Diagrama Realidade e Desejo

Exemplo

Example

$$\pi = [-4\ \ -5\ 8\ 7\ 1\ \ -2\ 3\ 6],$$

0 +4 -4 +5 -5 -8 +8 -7 +7 -1 +1 +2 -2 -3 +3 -6 +6 -9

Diagrama Realidade e Desejo
Exemplo

Example

$$\pi = [-4\ \ -5\ 8\ 7\ 1\ \ -2\ 3\ 6],$$

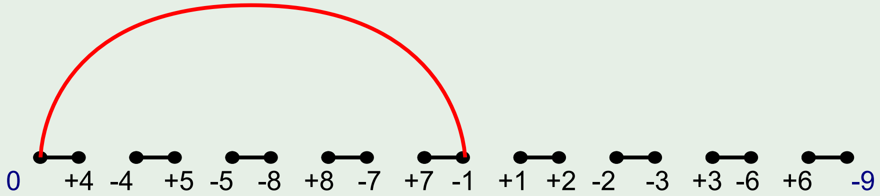


Diagrama Realidade e Desejo

Exemplo

Example

$$\pi = [-4\ \ -5\ 8\ 7\ 1\ \ -2\ 3\ 6],$$

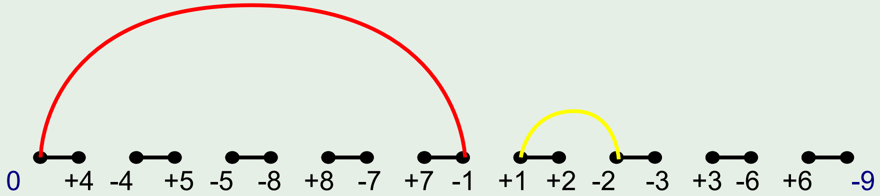
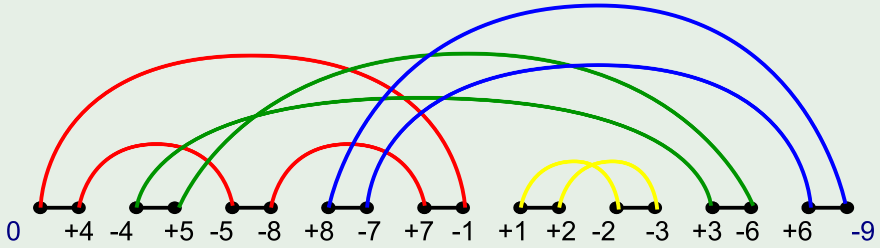


Diagrama Realidade e Desejo

Exemplo

Example

$$\pi = [-4 \;\; -5\; 8\; 7\; 1 \;\; -2\; 3\; 6],$$



- *Arestas de Realidade* = Arestas **pretas**; *Arestas de Desejo* = Arestas vermelhas, verdes, azuis e amarelas.

Diagrama Realidade e Desejo

Orientação

- Um ciclo C é orientado quando um par de arestas de realidade possuem *setas divergentes*;
- Quando não existe um par de arestas em C divergentes, dizemos que C é não-orientado. Todo par de arestas de realidade possuem *setas convergentes*.
- Um ciclo C é fortemente não-orientado quando for *não-orientado* e entre todos os pares de arestas de realidade de C não existe uma aresta de realidade de um ciclo D orientado.

Example

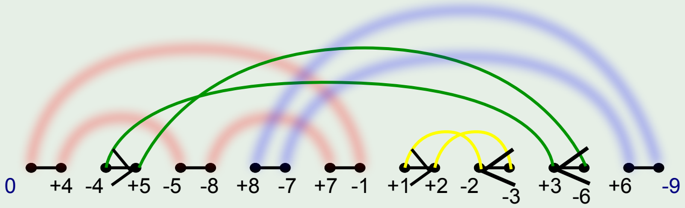


Figura: ciclos orientados

Example

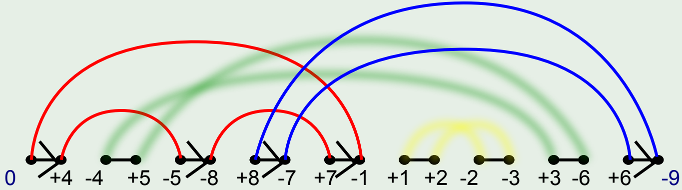


Figura: ciclos não orientados

Fórmula para distância

$$d_{rs}(\pi) = n + 1 - c(\pi) + t(\pi).$$

$$\begin{aligned}
c(\pi\rho_1) - c(\pi) &\leq 1, \\
c(\pi\rho_1\rho_2) - c(\pi\rho_1) &\leq 1, \\
&\vdots \\
c(\pi\rho_1 \ldots \rho_r) - c(\pi\rho_1 \ldots \rho_{r-1}) &\leq 1. \\
n + 1 - c(\pi) &\leq r = d_{rs}(\pi).
\end{aligned}$$

Para o valor exato para distância somamos $t(\pi)$ que representa o número de ciclos *fortemente não-orientados*.

Fórmula para distância

$$d_{rs}(\pi) = n + 1 - c(\pi) + t(\pi).$$

$$\begin{aligned}
c(\pi\rho_1) - c(\pi) \quad & \leq 1, \\
c(\pi\rho_1\rho_2) - c(\pi\rho_1) \quad & \leq 1, \\
. \\
. \\
. \\
c(\pi\rho_1 ... \rho_r) - c(\pi\rho_1 ... \rho_{r-1}) \quad & \leq 1. \\
n + 1 - c(\pi) \quad & \leq r = d_{rs}(\pi).
\end{aligned}$$

Para o valor exato para distância somamos $t(\pi)$ que representa o número de ciclos *fortemente não-orientados*.

Fórmula para distância

$$d_{rs}(\pi) = n + 1 - c(\pi) + t(\pi).$$

$$\begin{aligned}
c(\pi\rho_1) - c(\pi) & \leq 1, \\
c(\pi\rho_1\rho_2) - c(\pi\rho_1) & \leq 1, \\
. & \\
. & \\
. & \\
c(\pi\rho_1 \ldots \rho_r) - c(\pi\rho_1 \ldots \rho_{r-1}) & \leq 1. \\
n + 1 - c(\pi) & \leq r = d_{rs}(\pi).
\end{aligned}$$

Para o valor exato para distância somamos $t(\pi)$ que representa o número de ciclos *fortemente não-orientados*.

Transposição

Definition

Uma transposição $t(i, j, k)$, com $1 \leq i < j < k \leq n+1$, aplicada em $\pi_{[n]} \in S_n$, troca o intervalo fechado delimitado por i e $j-1$ com o intervalo fechado delimitado por j e $k-1$. Transformando $\pi_{[n]}$ em $\pi_{[n]}t(i, j, k)$.

Example

$$\pi_{[5]} = [3\,4\,5\,2\,1].$$
$$\pi_{[5]}t(2, 4, 6) = [3\,2\,1\,4\,5].$$

Distância de Transposição

Problema

- Distância de Transposição = Dadas duas permutações $\pi, \sigma \in S_n$, determinar o menor número de transposições q, tais que, $\pi t_1 t_2 \ldots t_q = \sigma$. Assim, $d_t(\pi, \sigma) = q$.
- Equivalente a determinarmos a menor sequência de π à ι. Denotamos $d_t(\pi, \iota)$ por $d_t(\pi)$.

Distância de Transposição

Como conseguir a solução do problema de maneira eficiente?

PERGUNTA: Dada uma permutação π e um inteiro k, consigo dizer de maneira eficiente se $d_t(\pi) \leq k$?

Distância de Transposição

Como conseguir a solução do problema de maneira eficiente?

PERGUNTA: Dada uma permutação π e um inteiro k, consigo dizer de maneira eficiente se $d_t(\pi) \leq k$?

RESPOSTA: **Não Sei!!!!!**

Temos porém alguns limites.

Distância de Transposição

Como conseguir a solução do problema de maneira eficiente?

PERGUNTA: Dada uma permutação π e um inteiro k, consigo dizer de maneira eficiente se $d_t(\pi) \leq k$?

RESPOSTA: **Não Sei!!!!!**

Temos porém alguns limites.

Distância de Transposição
Exemplo

Example

$$
\begin{aligned}
\pi &= [4\ 5\ \underline{8\ 7\ 1\ 2\ 3}\ 6],\\
\pi t(1,3,8) &= [8\ \underline{7}\ 1\ 2\ 3\ 4\ 5\ 6],\\
\pi t(1,3,8)t(1,2,3) &= [7\ 8\ \underline{1\ 2\ 3\ 4\ 5\ 6}],\\
\pi t(1,3,8)t(1,2,3)t(1,3,9) &= [1\ 2\ 3\ 4\ 5\ 6\ 7\ 8].
\end{aligned}
$$

$d_t(\pi) \leq 3$.

Será que esse limite é justo?

Distância de Transposição

Exemplo

Example

$$
\begin{array}{lcl}
\pi & = & [4\ 5\ \underline{8\ 7\ 1\ 2\ 3}\ 6], \\
\pi t(1,3,8) & = & [8\ \underline{7}\ 1\ 2\ 3\ 4\ 5\ 6], \\
\pi t(1,3,8)t(1,2,3) & = & [7\ 8\ \underline{1\ 2\ 3\ 4\ 5\ 6}], \\
\pi t(1,3,8)t(1,2,3)t(1,3,9) & = & [1\ 2\ 3\ 4\ 5\ 6\ 7\ 8].
\end{array}
$$

$d_t(\pi) \leq 3$.

Será que esse limite é justo?

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

Assim como para o caso sinalizado, $RD(\pi)$ é definido do seguinte modo:

Definition

Diagrama Realidade e Desejo de uma permutação $\pi_{[n]}$, $RD(\pi) =$ Multigrafo $RD(\pi) = (V, R \cup D)$ onde:

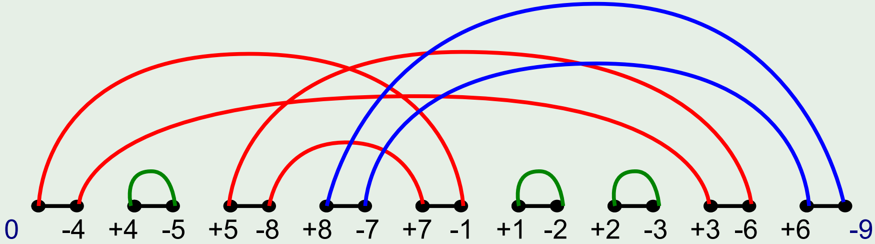
$$
\begin{aligned}
V &:= \{0, -\pi_1, +\pi_1, -\pi_2, +\pi_2, \cdots, -\pi_n, +\pi_n, -(n+1)\}; \\
&\quad \textit{Arestas de Realidade:} \\
R &:= \{(+\pi_i, -\pi_{i+1} | i = 1, ..., n-1)\} \cup \{(0, -\pi_1), (+\pi_n, -(n+1))\}; \\
&\quad \textit{Arestas de Desejo:} \\
D &:= \{(+i, -(i+1)) | i = 1, ..., n-1\} \cup \{(0, -1), (+n, -(n+1))\}.
\end{aligned}
$$

Diagrama Realidade e Desejo

Exemplo

Example

$$\pi = [4\ 5\ 8\ 7\ 1\ 2\ 3\ 6],$$



Limites para distância de transposição
Novamente o Diagrama Realidade e Desejo

Cada transposição segue uma das seguintes formas:

- 2-move = Aumenta em 2 o número de ciclos ímpares;
- 0-move = Permanece igual o número de ciclos ímpares;
- -2-move = Diminui em 2 o número de ciclos ímpares.

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

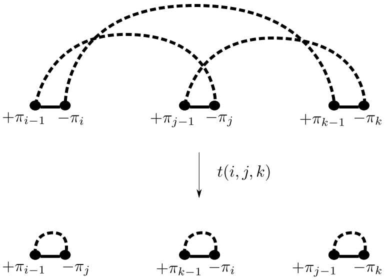


Figura: 2-move

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

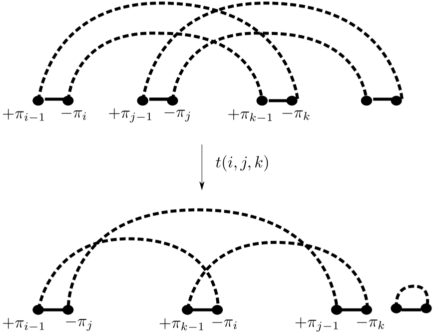


Figura: 2-move

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

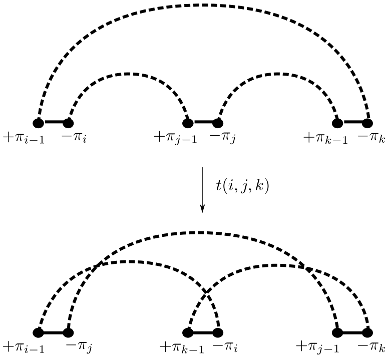


Figura: 0-move

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

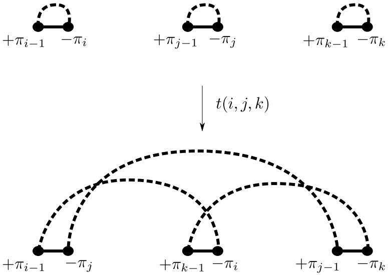


Figura: -2-move

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

Example

$\pi = [4\ 5\ 8\ 7\ 1\ 2\ 3\ 6]$.

Figura: π

Limites para distância de transposição
Novamente o Diagrama Realidade e Desejo

Example

$\pi t(1,3,8) = [8\ 7\ 1\ 2\ 3\ 4\ 5\ 6]$.

Figura: $\pi t(1,3,8)$

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

Example

$\pi t(1,3,8)t(1,2,3) = [7\ 8\ 1\ 2\ 3\ 4\ 5\ 6]$.

Figura: $\pi t(1,3,8)t(1,2,3)$

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

Example

$\pi t(1,3,8)t(1,2,3)t(1,3,9) = [1\ 2\ 3\ 4\ 5\ 6\ 7\ 8]$.

Figura: $\pi t(1,3,8)t(1,2,3)t(1,3,9)$

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

Como determinar limites para distância de transposição utilizando o Diagrama Realidade e Desejo?

- O melhor possível é a cada transposição ser do tipo $2-$*move*.

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

Como determinar limites para distância de transposição utilizando o Diagrama Realidade e Desejo?

- O melhor possível é a cada transposição ser do tipo $2-move$.

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

$$d_t(\pi) \geq \left\lceil \frac{n+1-c_{ímpar}(\pi)}{2} \right\rceil .$$

$$\begin{aligned}
c_{ímpar}(\pi t_1) - c_{ímpar}(\pi) &\leq 2,\\
c_{ímpar}(\pi t_1 t_2) - c_{ímpar}(\pi t_1) &\leq 2,\\
&\vdots\\
c_{ímpar}(\pi t_1 \dots t_q) - c_{ímpar}(\pi t_1 \dots t_{q-1}) &\leq 2.\\
n+1-c_{ímpar}(\pi) &\leq 2q = 2d_t(\pi).
\end{aligned}$$

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

$$d_t(\pi) \geq \left\lceil \frac{n+1-c_{ímpar}(\pi)}{2} \right\rceil .$$

$$\begin{array}{ll}
c_{ímpar}(\pi t_1) - c_{ímpar}(\pi) & \leq 2, \\
c_{ímpar}(\pi t_1 t_2) - c_{ímpar}(\pi t_1) & \leq 2, \\
. & \\
. & \\
. & \\
c_{ímpar}(\pi t_1 \ldots t_q) - c_{ímpar}(\pi t_1 \ldots t_{q-1}) & \leq 2. \\
n+1-c_{ímpar}(\pi) & \leq 2q = 2d_t(\pi).
\end{array}$$

Limites para distância de transposição

Novamente o Diagrama Realidade e Desejo

$$\left\lceil \frac{n+1-c_{ímpar}(\pi)}{2} \right\rceil \leq d_t(\pi) \leq \left\lceil \frac{11}{8}\frac{(n+1-c_{ímpar}(\pi))}{2} \right\rceil .$$

Voltando ao Exemplo

Example

$\pi = [4\ 5\ 8\ 7\ 1\ 2\ 3\ 6]$. Pelos limites anteriores obtemos $3 \leq d_t(\pi) \leq 5$, como temos uma sequência com 3 transposições – $t(1,3,8)$, $t(1,2,3)$, $t(1,3,9)$ – obtemos que $d_t(\pi) = 3$.

Árvores Filogenéticas

- Dado um conjunto de permutações como entradas e um parâmetro inteiro k, encontrar, caso exista, uma árvore onde as folhas as são permutações da entrada e os nós internos – ancestrais das folhas – possuem distância de rearranjo até k em relação as folhas, sendo igual a k na raíz da árvore.
- Entender a evolução das espécies.

Árvores Filogenéticas

Example

Rearranjo de *Reversão Sinalizada*. Conjunto de entrada $=\{[3\ 2\ 1\ 4],\ [3\ -1\ -2\ -4],\ [-3\ 2\ -4\ -1],\ [-3\ 1\ -4\ -2]\}$.
Parâmetro para distância: 2.

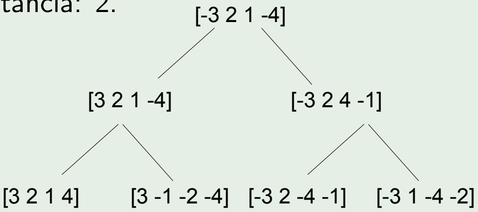


Figura: Existe uma árvore filogenética que satisfaça o conjunto de entrada e inteiro dados.

Alguns dos muitos Problemas em Aberto

Problemas que não sei se sei ou não sei

- Diâmetro de Transposição;
- Distância de Transposição Pré-fixada;
- Distância de Reversão e Transposição;

Alguns dos muitos Problemas em Aberto

Problemas que não sei se sei ou não sei

- Diâmetro de Transposição;
- Distância de Transposição Pré-fixada;
- Distância de Reversão e Transposição;

Alguns dos muitos Problemas em Aberto

Problemas que não sei se sei ou não sei

- Diâmetro de Transposição;
- Distância de Transposição Pré-fixada;
- Distância de Reversão e Transposição;

Obrigado!

luis@vm.uff.br
lfignacio@cos.ufrj.br / www.cos.ufrj.br/~lfignacio

Obrigado!

luis@vm.uff.br
lfignacio@cos.ufrj.br / www.cos.ufrj.br/∼lfignacio