

Rearranjo de Genomas

Zanoni Dias

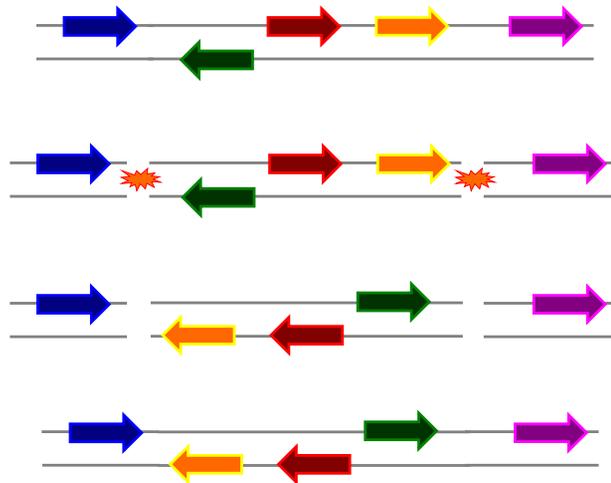
Instituto de Computação – Unicamp

16 de junho de 2009

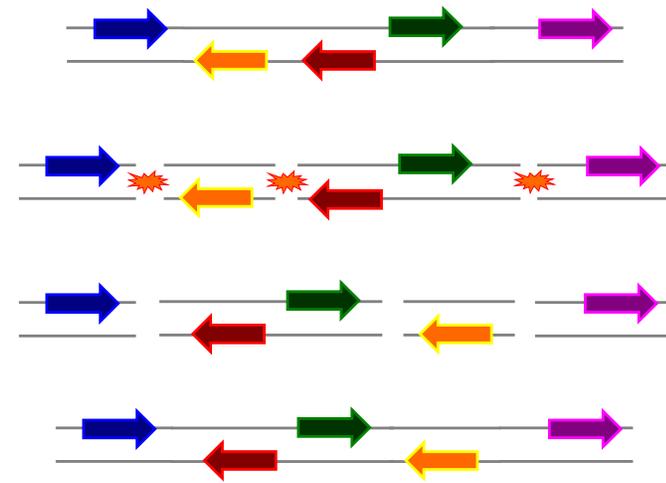
Rearranjo de Genomas

- Comparação de genomas considerando eventos de mutação que afetam grandes porções do genoma.
- Rearranjo de genomas é uma forma mais adequada de comparar genomas completos.
- A comparação é realizada apenas com relação ao conjunto dos blocos conservados (um ou mais genes inteiros).
- Principais eventos:
 - ▶ Conservativos:
 - ★ Reversão
 - ★ Transposição
 - ★ Transposição Reversa
 - ★ Fissão
 - ★ Fusão
 - ★ Translocação
 - ▶ Não Conservativos:
 - ★ Inserção
 - ★ Remoção
 - ★ Duplicação

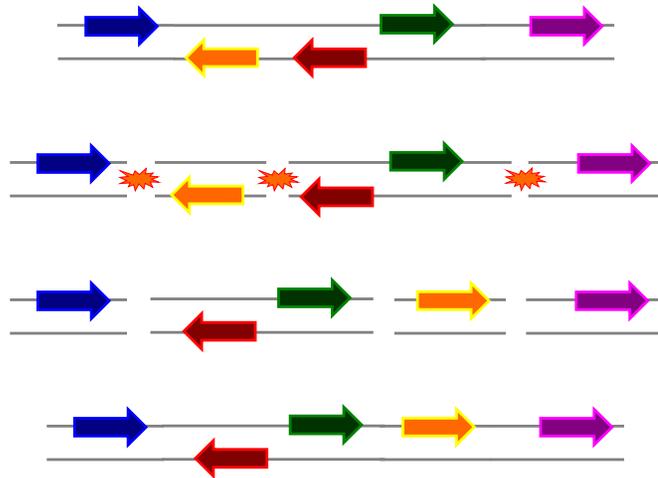
Reversão



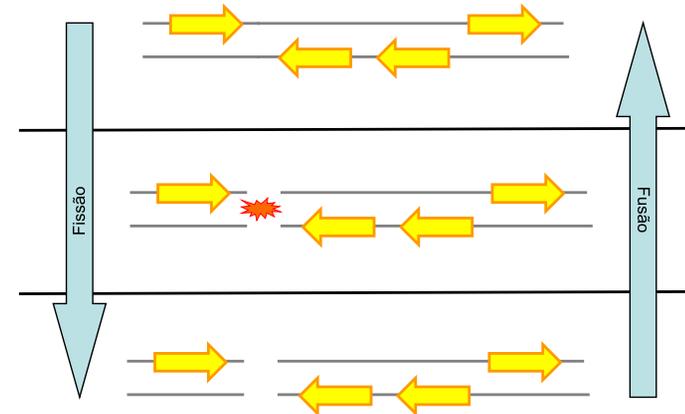
Transposição



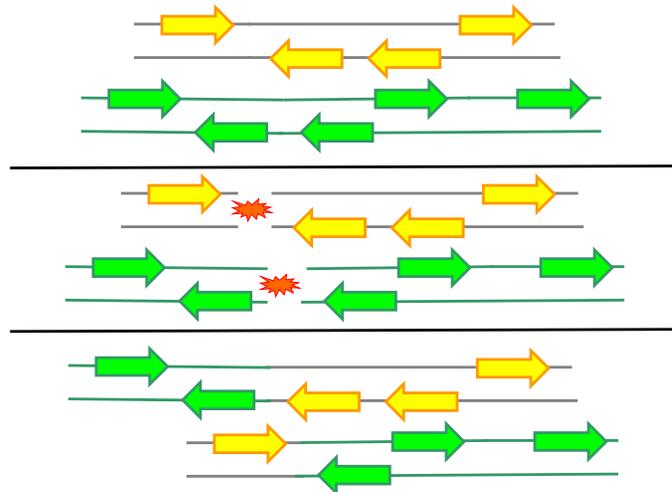
Transposição Reversa



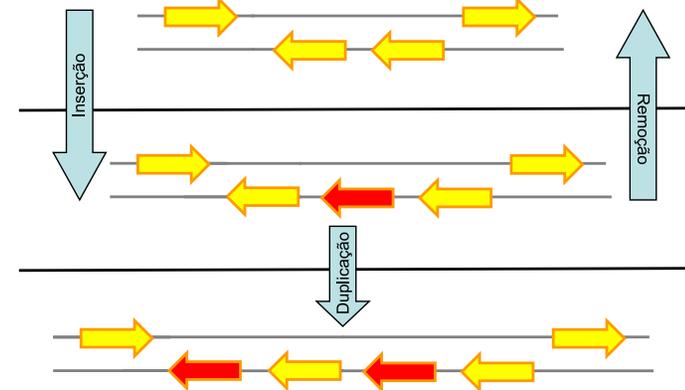
Fissão e Fusão



Translocação



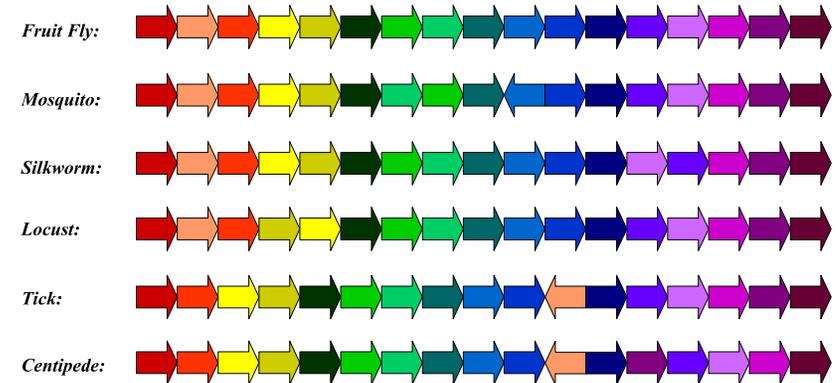
Eventos Não Conservativos: Inserção, Remoção e Duplicação



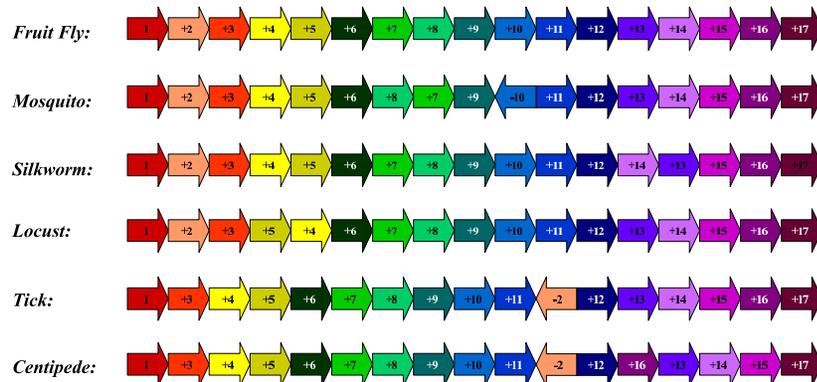
Genoma Mitocondrial

- Mitocôndria é uma organela envolvida no processo de respiração celular presente na maioria dos eucariotos.
- Possui um genoma circular com aproximadamente 16kbp, com 37 genes, sendo que 13 codificam proteínas, 22 codificam RNAs transportadores e 2 codificam RNAs ribossomais.
- O genoma mitocondrial é altamente conservado em animais, mas a ordem dos genes varia bastante de espécie para espécie.

Blocos Conservados em Mitocôndrias de Artrópodes



Blocos Conservados em Mitocôndrias de Artrópodes



Problema da Distância de Reversão sem Orientação de Genes

- Nem sempre é possível conhecer os blocos conservados e a orientações dos genes de dois genomas.
- Podemos representar um genoma com n blocos conservados como uma permutação, $\pi = \pi_1 \pi_2 \dots \pi_n$, dos números de 1 a n .
- A reversão $\rho(i, j)$, com $1 \leq i < j \leq n$, reverte a ordem de $\pi[i..j]$, ou seja, $\pi \cdot \rho(i, j) = \pi_1 \pi_2 \dots \pi_{i-1} \pi_j \pi_{j-1} \dots \pi_{i+1} \pi_i \pi_{j+1} \dots \pi_{n-1} \pi_n$.
- **Distância de Reversão:** dados dois genomas compostos por n blocos conservados, representados pelas permutações π e σ , calcular a distância de reversão ($d(\pi, \sigma)$) entre π e σ , ou seja, obter uma série de reversões $\rho_1, \rho_2, \dots, \rho_t$, de tamanho mínimo, tal que $d(\pi, \sigma) = t$ e $\pi \cdot \rho_1 \cdot \rho_2 \dots \rho_t = \sigma$.
- **Ordenação por Reversões:** dado um genoma composto por n blocos conservados, representado pela permutação π , calcular a distância de reversão ($d(\pi)$) entre π e a permutação identidade $\iota = (1, 2, \dots, n)$, ou seja, $d(\pi) = d(\pi, \iota)$.

Distância de Reversão × Ordenação por Reversões

$$\pi = 5 \quad 3 \quad 1 \quad 6 \quad 7 \quad 4 \quad 2$$

$$\sigma = 6 \quad 3 \quad 2 \quad 7 \quad 5 \quad 1 \quad 4$$

Distância de Reversão × Ordenação por Reversões

$$\pi = 5 \quad 3 \quad 1 \quad 6 \quad 7 \quad \boxed{4 \quad 2}$$

$$\sigma = 6 \quad 3 \quad 2 \quad 7 \quad 5 \quad 1 \quad 4$$

Distância de Reversão × Ordenação por Reversões

$$\pi = 5 \quad 3 \quad 1 \quad 6 \quad 7 \quad \boxed{4 \quad 2}$$

$$5 \quad 3 \quad 1 \quad 6 \quad 7 \quad 2 \quad 4$$

$$\sigma = 6 \quad 3 \quad 2 \quad 7 \quad 5 \quad 1 \quad 4$$

Distância de Reversão × Ordenação por Reversões

$$\pi = 5 \quad 3 \quad 1 \quad 6 \quad 7 \quad \boxed{4 \quad 2}$$

$$5 \quad 3 \quad \boxed{1 \quad 6 \quad 7 \quad 2} \quad 4$$

$$\sigma = 6 \quad 3 \quad 2 \quad 7 \quad 5 \quad 1 \quad 4$$

Distância de Reversão × Ordenação por Reversões

$$\pi = 5 \quad 3 \quad 1 \quad 6 \quad 7 \quad \boxed{4 \quad 2}$$

$$5 \quad 3 \quad \boxed{1 \quad 6 \quad 7 \quad 2} \quad 4$$

$$5 \quad 3 \quad 2 \quad 7 \quad 6 \quad 1 \quad 4$$

$$\sigma = 6 \quad 3 \quad 2 \quad 7 \quad 5 \quad 1 \quad 4$$

Distância de Reversão × Ordenação por Reversões

$$\pi = 5 \quad 3 \quad 1 \quad 6 \quad 7 \quad \boxed{4 \quad 2}$$

$$5 \quad 3 \quad \boxed{1 \quad 6 \quad 7 \quad 2} \quad 4$$

$$\boxed{5 \quad 3 \quad 2 \quad 7 \quad 6} \quad 1 \quad 4$$

$$\sigma = 6 \quad 3 \quad 2 \quad 7 \quad 5 \quad 1 \quad 4$$

Distância de Reversão × Ordenação por Reversões

$$\pi = 5 \quad 3 \quad 1 \quad 6 \quad 7 \quad \boxed{4 \quad 2}$$

$$5 \quad 3 \quad \boxed{1 \quad 6 \quad 7 \quad 2} \quad 4$$

$$\boxed{5 \quad 3 \quad 2 \quad 7 \quad 6} \quad 1 \quad 4$$

$$6 \quad 7 \quad 2 \quad 3 \quad 5 \quad 1 \quad 4$$

$$\sigma = 6 \quad 3 \quad 2 \quad 7 \quad 5 \quad 1 \quad 4$$

Distância de Reversão × Ordenação por Reversões

$$\pi = 5 \quad 3 \quad 1 \quad 6 \quad 7 \quad \boxed{4 \quad 2}$$

$$5 \quad 3 \quad \boxed{1 \quad 6 \quad 7 \quad 2} \quad 4$$

$$\boxed{5 \quad 3 \quad 2 \quad 7 \quad 6} \quad 1 \quad 4$$

$$6 \quad \boxed{7 \quad 2 \quad 3} \quad 5 \quad 1 \quad 4$$

$$\sigma = 6 \quad 3 \quad 2 \quad 7 \quad 5 \quad 1 \quad 4$$

Distância de Reversão × Ordenação por Reversões

$$\pi = \begin{matrix} 5 & 3 & 1 & 6 & 7 & 4 & 2 \\ & & & & & 4 & 2 \end{matrix}$$

$$\begin{matrix} 5 & 3 & 1 & 6 & 7 & 2 & 4 \\ & & & & & 1 & 6 & 7 & 2 & 4 \end{matrix}$$

$$\begin{matrix} 5 & 3 & 2 & 7 & 6 & 1 & 4 \\ & & & & & 5 & 3 & 2 & 7 & 6 & 1 & 4 \end{matrix}$$

$$\begin{matrix} 6 & 7 & 2 & 3 & 5 & 1 & 4 \\ & & & & & 6 & 7 & 2 & 3 & 5 & 1 & 4 \end{matrix}$$

$$\sigma = \begin{matrix} 6 & 3 & 2 & 7 & 5 & 1 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 = \iota \end{matrix}$$

Distância de Reversão × Ordenação por Reversões

$$\pi = \begin{matrix} 5 & 3 & 1 & 6 & 7 & 4 & 2 \\ 5 & 2 & 6 & 1 & 4 & 7 & 3 = \pi' \end{matrix}$$

$$\begin{matrix} 5 & 3 & 1 & 6 & 7 & 2 & 4 \\ & & & & & 1 & 6 & 7 & 2 & 4 \end{matrix}$$

$$\begin{matrix} 5 & 3 & 2 & 7 & 6 & 1 & 4 \\ & & & & & 5 & 3 & 2 & 7 & 6 & 1 & 4 \end{matrix}$$

$$\begin{matrix} 6 & 7 & 2 & 3 & 5 & 1 & 4 \\ & & & & & 6 & 7 & 2 & 3 & 5 & 1 & 4 \end{matrix}$$

$$\sigma = \begin{matrix} 6 & 3 & 2 & 7 & 5 & 1 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 = \iota \end{matrix}$$

Distância de Reversão × Ordenação por Reversões

$$\pi = \begin{matrix} 5 & 3 & 1 & 6 & 7 & 4 & 2 \\ 5 & 2 & 6 & 1 & 4 & 7 & 3 = \pi' \end{matrix}$$

$$\begin{matrix} 5 & 3 & 1 & 6 & 7 & 2 & 4 \\ 5 & 2 & 6 & 1 & 4 & 3 & 7 \end{matrix}$$

$$\begin{matrix} 5 & 3 & 2 & 7 & 6 & 1 & 4 \\ & & & & & 5 & 3 & 2 & 7 & 6 & 1 & 4 \end{matrix}$$

$$\begin{matrix} 6 & 7 & 2 & 3 & 5 & 1 & 4 \\ & & & & & 6 & 7 & 2 & 3 & 5 & 1 & 4 \end{matrix}$$

$$\sigma = \begin{matrix} 6 & 3 & 2 & 7 & 5 & 1 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 = \iota \end{matrix}$$

Distância de Reversão × Ordenação por Reversões

$$\pi = \begin{matrix} 5 & 3 & 1 & 6 & 7 & 4 & 2 \\ 5 & 2 & 6 & 1 & 4 & 7 & 3 = \pi' \end{matrix}$$

$$\begin{matrix} 5 & 3 & 1 & 6 & 7 & 2 & 4 \\ 5 & 2 & 6 & 1 & 4 & 3 & 7 \end{matrix}$$

$$\begin{matrix} 5 & 3 & 2 & 7 & 6 & 1 & 4 \\ 5 & 2 & 3 & 4 & 1 & 6 & 7 \end{matrix}$$

$$\begin{matrix} 6 & 7 & 2 & 3 & 5 & 1 & 4 \\ & & & & & 6 & 7 & 2 & 3 & 5 & 1 & 4 \end{matrix}$$

$$\sigma = \begin{matrix} 6 & 3 & 2 & 7 & 5 & 1 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 = \iota \end{matrix}$$

Distância de Reversão × Ordenação por Reversões

$$\pi = \begin{matrix} 5 & 3 & 1 & 6 & 7 & 4 & 2 \\ 5 & 2 & 6 & 1 & 4 & 7 & 3 \end{matrix} = \pi'$$

$$\begin{matrix} 5 & 3 & 1 & 6 & 7 & 2 & 4 \\ 5 & 2 & 6 & 1 & 4 & 3 & 7 \end{matrix}$$

$$\begin{matrix} 5 & 3 & 2 & 7 & 6 & 1 & 4 \\ 5 & 2 & 3 & 4 & 1 & 6 & 7 \end{matrix}$$

$$\begin{matrix} 6 & 7 & 2 & 3 & 5 & 1 & 4 \\ 1 & 4 & 3 & 2 & 5 & 6 & 7 \end{matrix}$$

$$\sigma = \begin{matrix} 6 & 3 & 2 & 7 & 5 & 1 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix} = \iota$$

Ordenação por Reversões sem Orientação dos Genes

- Podemos adaptar algoritmos de ordenação para usarem apenas reversões para ordenação (sem necessariamente minimizar o número de reversões utilizadas).
- A complexidade de algoritmos de ordenação geralmente é calculada em termos do número de comparações efetuadas.
- No caso do problema da ordenação por reversões, seria interessante adaptar o algoritmo de ordenação que fizesse o menor número possível de trocas, já que as trocas de elementos devem ser transformadas em reversões.
- Entre os algoritmos de ordenação mais comumente utilizados, o Selection Sort é o único que faz no máximo $O(n)$ trocas.

Algoritmo Ingênuo Baseado no Selection Sort para Ordenação por Reversões sem Orientação dos Genes

Algoritmo 1: Selection Sorting by Reversal

Input: π, n

$t \leftarrow 0$

for all $i \in [1..n-1]$ **do**

$j \leftarrow 1$

while $\pi_j \neq i$ **do**

$j \leftarrow j+1$

end

if $j \neq i$ **then**

$t \leftarrow t+1$

$\pi \leftarrow \pi \cdot \rho(i, j)$

end

end

return t

Algoritmo Ingênuo Baseado no Selection Sort para Ordenação por Reversões sem Orientação dos Genes

- Complexidade: $O(n^2)$.
- Aproximação:
 - Considere a permutação $\pi = (n, 1, 2, \dots, n-2, n-1)$.
 - O algoritmo ingênuo usa $n-1$ reversões para ordenar π :
 - $\pi \leftarrow \pi \cdot \rho(1, 2) = (1, n, 2, \dots, n-2, n-1)$
 - $\pi \leftarrow \pi \cdot \rho(2, 3) = (1, 2, n, \dots, n-2, n-1)$
 - ...
 - $\pi \leftarrow \pi \cdot \rho(n-2, n-1) = (1, 2, \dots, n, n-1)$
 - $\pi \leftarrow \pi \cdot \rho(n-1, n) = (1, 2, \dots, n-1, n) = \iota$
 - É possível ordenar π com apenas duas reversões:
 - $\pi \leftarrow \pi \cdot \rho(2, n) = (n, n-1, n-2, \dots, 2, 1)$
 - $\pi \leftarrow \pi \cdot \rho(1, n) = (1, 2, \dots, n-1, n) = \iota$
 - Logo, o algoritmo ingênuo não garante uma aproximação melhor do que $(n-1)/2$.

O Problema da Ordenação de Panquecas

- Dada uma pilha de panquecas circulares, ordená-las, deixando a panqueca de menor diâmetro no todo da pilha. O único movimento permitido para ordenar as panquecas é o de inserir uma espátula num ponto qualquer da pilha e inverter a ordem de todas as panquecas acima da espátula.
- Qual o número mínimo de movimentos suficientes para ordenar qualquer pilha de n panquecas?
- O Problema da Ordenação de Panquecas é equivalente o problema da Ordenação por Reversões de Prefixos, ou seja, o problema da Ordenação por Reversões onde só são permitidas reversões do tipo $\rho(1, i)$, para $2 \leq i \leq n$.

Algoritmo Guloso para o Problema de Ordenação de Panquecas

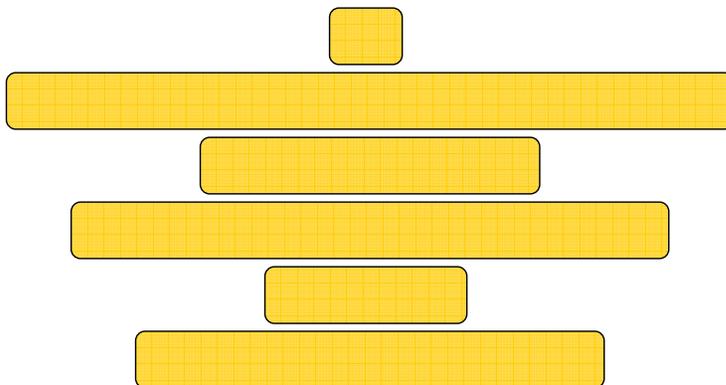
Algoritmo 2: Greedy Pancake Flipping Problem

```

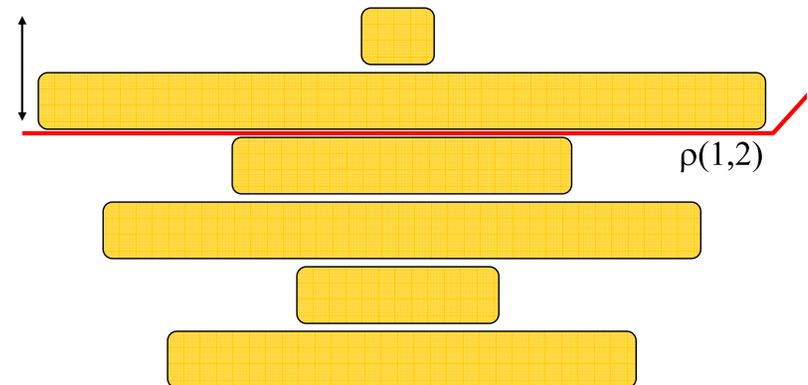
Input:  $\pi, n$ 
 $t \leftarrow 0$ 
for  $i \leftarrow n$  to 2 do
   $j \leftarrow 1$ 
  while  $\pi_j \neq i$  do
     $j \leftarrow j + 1$ 
  end
  end
  if  $j \neq i$  then
    if  $j \neq 1$  then
       $t \leftarrow t + 1$ 
       $\pi \leftarrow \pi \cdot \rho(1, j)$ 
    end
     $t \leftarrow t + 1$ 
     $\pi \leftarrow \pi \cdot \rho(1, i)$ 
  end
end
end
return  $t$ 

```

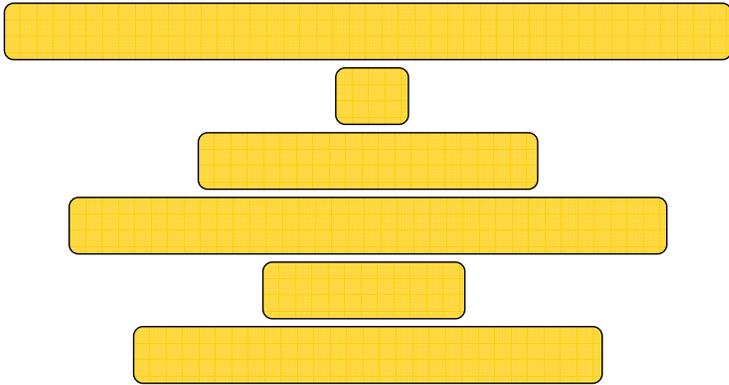
O Problema da Ordenação de Panquecas



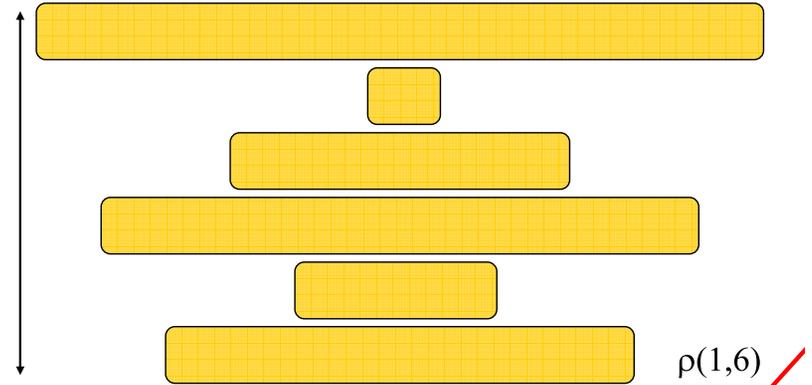
O Problema da Ordenação de Panquecas



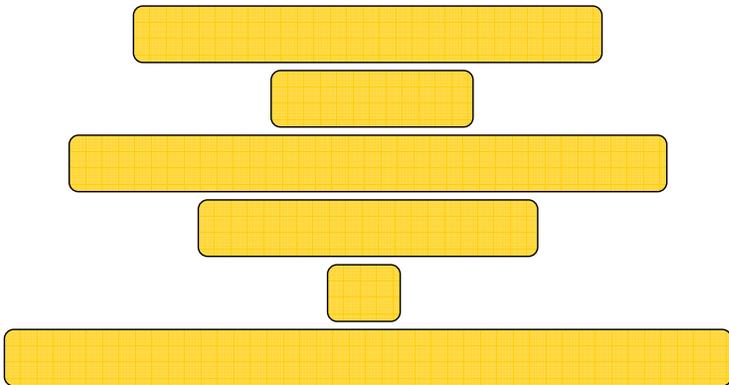
O Problema da Ordenação de Panquecas



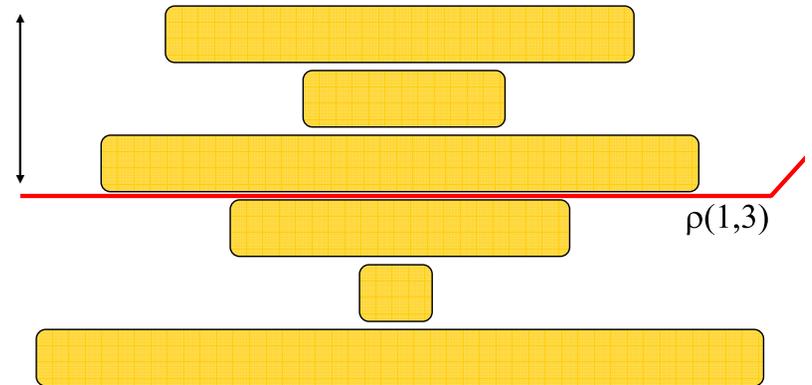
O Problema da Ordenação de Panquecas



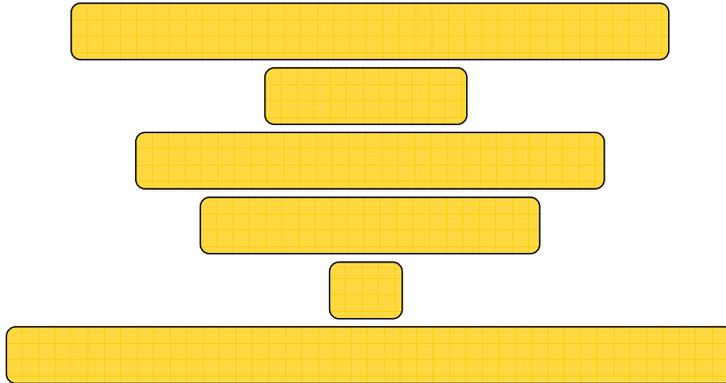
O Problema da Ordenação de Panquecas



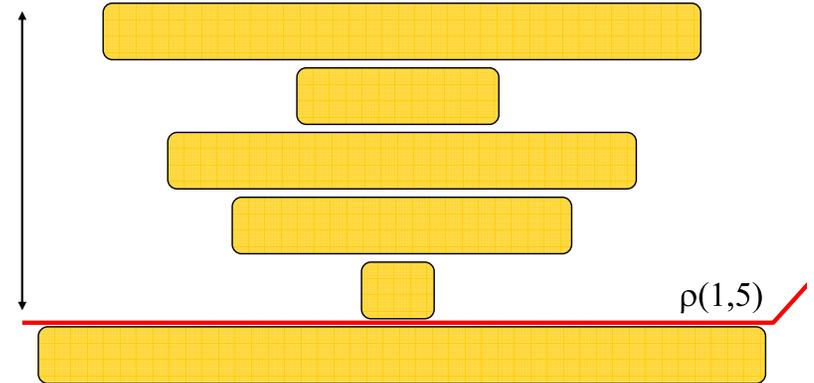
O Problema da Ordenação de Panquecas



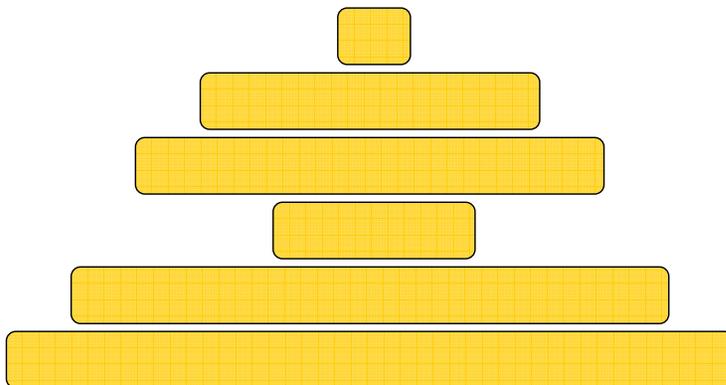
O Problema da Ordenação de Panquecas



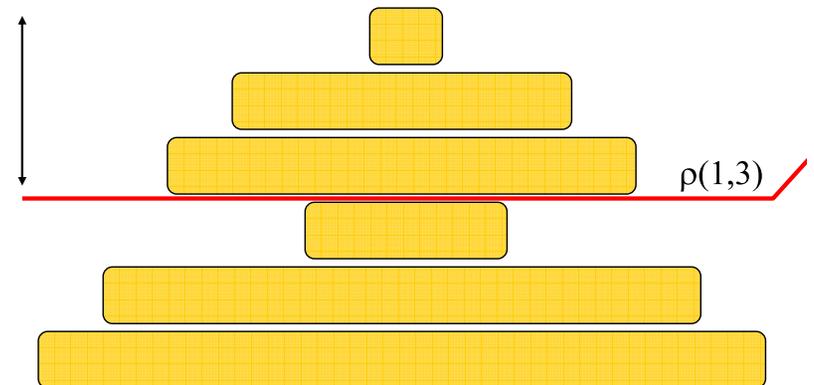
O Problema da Ordenação de Panquecas



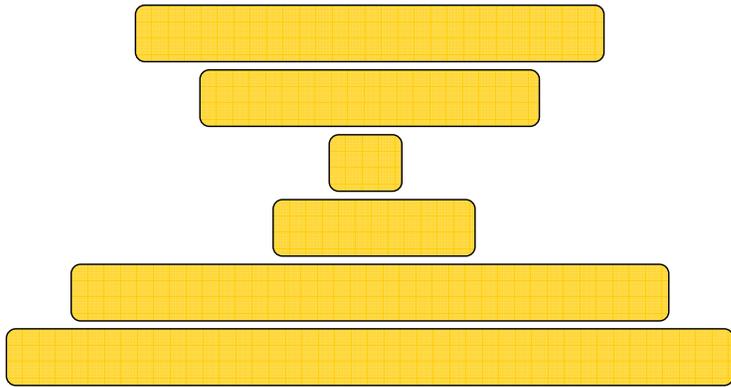
O Problema da Ordenação de Panquecas



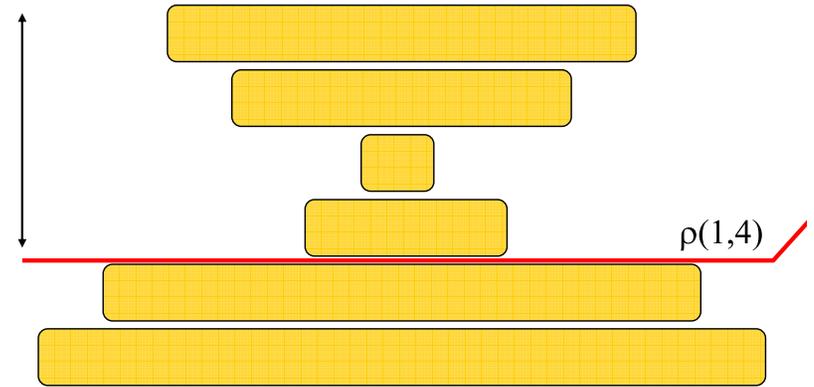
O Problema da Ordenação de Panquecas



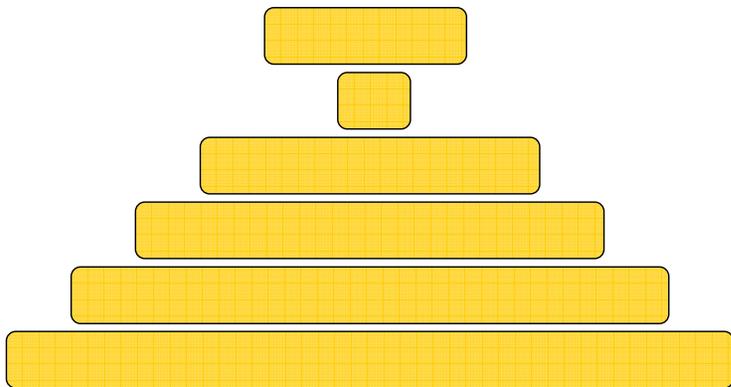
O Problema da Ordenação de Panquecas



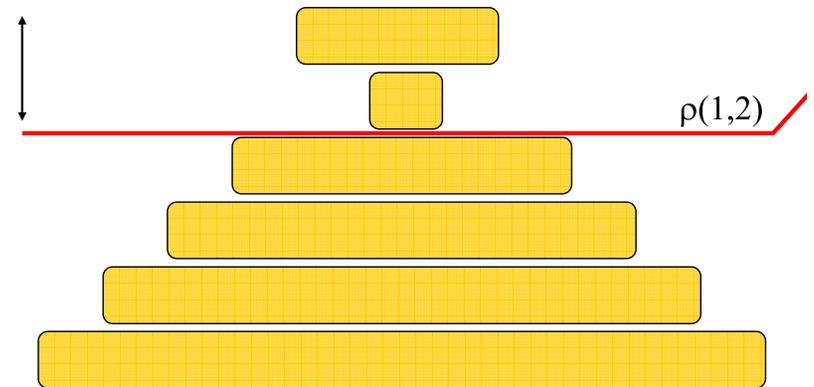
O Problema da Ordenação de Panquecas



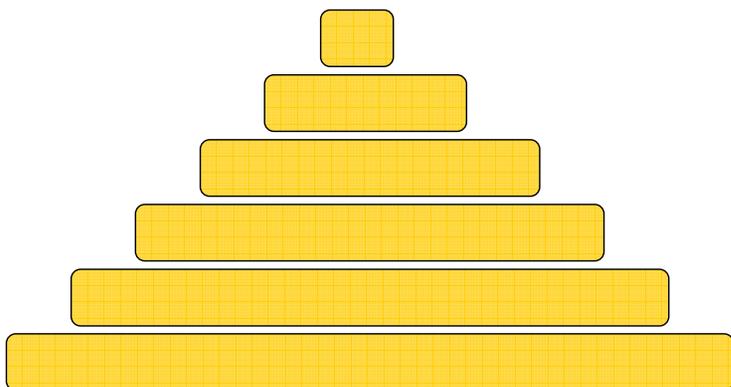
O Problema da Ordenação de Panquecas



O Problema da Ordenação de Panquecas



O Problema da Ordenação de Panquecas



O Problema da Ordenação de Panquecas

- Complexidade: $O(n^2)$.
- O algoritmo guloso ordena qualquer pilha de panquecas em no máximo $2n - 3$ movimentos.
- William Gates e Christos Papadimitriou provaram, em 1979, que $(5n + 5)/3$ movimentos são suficientes e $17n/16$ movimentos podem ser necessários para qualquer pilha de n panquecas
- Em 1997, Mohammad Heydari e Ivan Sudborough mostraram que podem ser necessários $15n/14$ movimentos para ordenar uma pilha de n panquecas.
- Em 2009, Chalam Chitturi, Bill Fahle, Zhaobing Meng, Linda Morales, Charles Shields, Ivan Sudborough, e Walter Voit, pela primeira vez em 30 anos, obtiveram um limite superior melhor do que o provado por Gates e Papadimitriou. A prova que são suficientes $(18/11)n$ movimentos ainda não foi publicada.

Breakpoints e Strips

- Vamos considerar a permutação estendida que pode ser obtida a partir de π inserindo-se dois novos elementos: $\pi_0 = 0$ e $\pi_{n+1} = n + 1$.
- Um par de elementos π_i e π_{i+1} , para $0 \leq i \leq n$, é uma *adjacência* se $|\pi_i - \pi_{i+1}| = 1$. Caso contrário, o par de elementos é chamado de *breakpoint*.
- Uma *strip* $\pi[i..j]$ é uma trecho maximal em π tal que todos os pares (π_k, π_{k+1}) são adjacências, para $i \leq k < j$.
- O número de *breakpoints* numa permutação π é denotado por $b(\pi)$.
- A única permutação sem *breakpoints* é a permutação identidade ($b(\iota) = 0$). Logo, ordenar por reversões é equivalente a remover todos o *breakpoints* de π .
- Seja $\Delta_b(\pi, \rho) = b(\pi \cdot \rho) - b(\pi)$. Logo, $\Delta_b(\pi, \rho) = \{-2, -1, 0, 1, 2\}$.
- O número de *breakpoints* de uma permutação π estabelece o seguinte limite inferior para a distância de ordenação:

$$d(\pi) \geq \frac{b(\pi)}{2}$$

Breakpoints e Strips

	0 • 4	3 • 2 • 5	6	7 • 1 • 8	9	
$\Delta_b(\pi, \rho) = +1$						
	0 • 4	5 • 2	3 • 6	7 • 1 • 8	9	
$\Delta_b(\pi, \rho) = +2$						
	0 • 4	5 • 2	6 • 3 • 7 • 1	8	9	
$\Delta_b(\pi, \rho) = -1$						
	0 • 4	5 • 2	1	7 • 3 • 6 • 8	9	
$\Delta_b(\pi, \rho) = 0$						
	0 • 4	5 • 1	2	7 • 3	6 • 8	9
$\Delta_b(\pi, \rho) = -2$						
	0 • 4	5 • 1	2	3 • 7	6 • 8	9

Breakpoints e Strips

Definição

Uma strip $\pi[i..j]$ é chamada decrescente se $i < j$ e $\pi_k > \pi_{k+1}$, para $i \leq k < j$. Caso contrário, a strip é chamada de crescente.

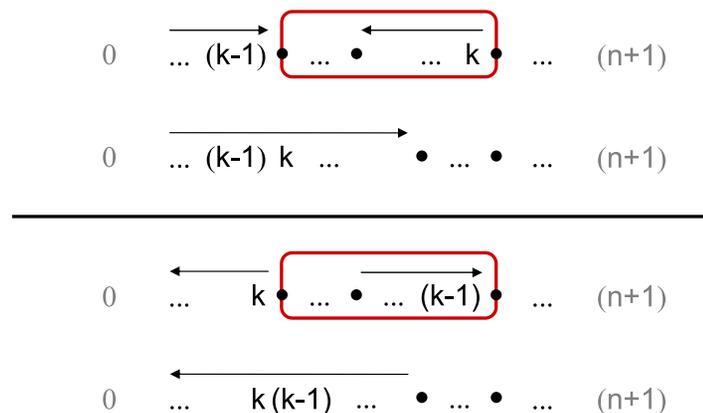
Teorema

Se o elemento k pertence a uma strip decrescente e o elemento $k - 1$ pertence a uma strip crescente, então existe uma reversão ρ tal que $\Delta_b(\pi, \rho) < 0$.

Lema

Seja π uma permutação com pelo menos uma strip decrescente. Então, existe uma reversão ρ tal que $\Delta_b(\pi, \rho) < 0$.

Remoção de Pelo Menos Um Breakpoint com Auxílio de uma Strip Decrescente



Breakpoints e Strips

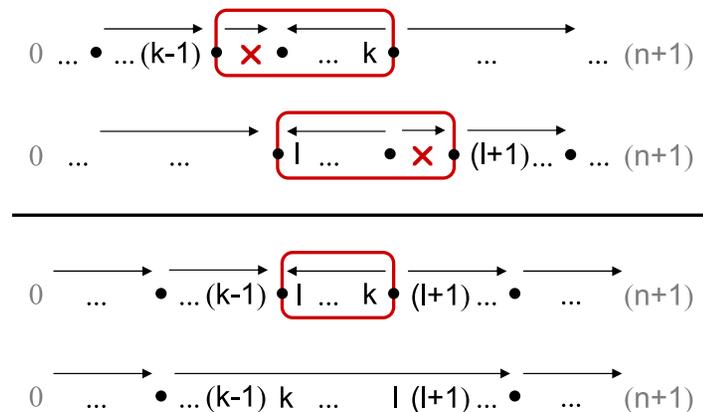
Teorema

Seja π uma permutação que possui uma única strip decrescente. Se todas as reversões ρ que removem breakpoints de π não deixam nenhuma strip decrescente em $\pi \cdot \rho$, então existe uma reversão ρ tal que $\Delta_b(\pi, \rho) = -2$.

Lema

Seja π uma permutação com pelo menos uma strip decrescente. Seja k o menor elemento entre todas as strips decrescentes de π e seja l o maior elemento entre todas as strips decrescentes de π . Seja ρ_k a reversão que posiciona k ao lado de $k - 1$, e seja ρ_l a reversão que posiciona l ao lado de $l + 1$. Se tanto $\pi \cdot \rho_k$ quanto $\pi \cdot \rho_l$ não possuírem nenhuma strip decrescente, então $\rho_k = \rho_l$ e $\Delta_b(\pi, \rho) = -2$.

Remoção de Dois Breakpoints ao Destruir a Última Strip Decrescente



Algoritmo Guloso para Distância de Reversão sem Orientação dos Genes

Algoritmo 3: Greedy Sorting by Reversal

```

Input:  $\pi, n$ 
 $t \leftarrow 0$ 
while  $\pi \neq \iota$  do
  if  $\pi$  has a decreasing strip then
     $k \leftarrow$  the smallest element in all decreasing strips
     $\rho \leftarrow$  the reversal that cuts after  $k$  and after  $k - 1$ 
    if  $\pi \cdot \rho$  has no decreasing strip then
       $l \leftarrow$  the largest element in all decreasing strips
       $\rho \leftarrow$  the reversal that cuts before  $l$  and before  $l + 1$ 
    end
  end
else
   $\rho \leftarrow$  the reversal that cuts the first two breakpoints
end
 $\pi \leftarrow \pi \cdot \rho$ 
 $t \leftarrow t + 1$ 
end
return  $t$ 

```

Algoritmo Guloso para Distância de Reversão sem Orientação dos Genes

$$\pi = \overrightarrow{0} \bullet \overrightarrow{3 \ 4} \bullet \overrightarrow{1 \ 2} \bullet \overrightarrow{7 \ 8} \bullet \overrightarrow{5 \ 6} \bullet \overrightarrow{9}$$

$$\iota = \overrightarrow{0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9}$$

Algoritmo Guloso para Distância de Reversão sem Orientação dos Genes

$$\pi = \overrightarrow{0} \bullet \overrightarrow{3 \ 4} \bullet \overrightarrow{1 \ 2} \bullet \overrightarrow{7 \ 8} \bullet \overrightarrow{5 \ 6} \bullet \overrightarrow{9}$$

$$\iota = \overrightarrow{0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9}$$

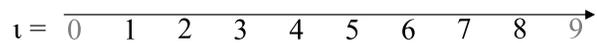
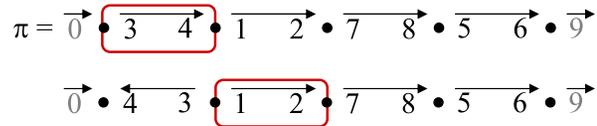
Algoritmo Guloso para Distância de Reversão sem Orientação dos Genes

$$\pi = \overrightarrow{0} \bullet \overrightarrow{3 \ 4} \bullet \overrightarrow{1 \ 2} \bullet \overrightarrow{7 \ 8} \bullet \overrightarrow{5 \ 6} \bullet \overrightarrow{9}$$

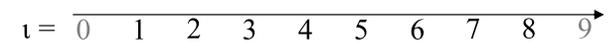
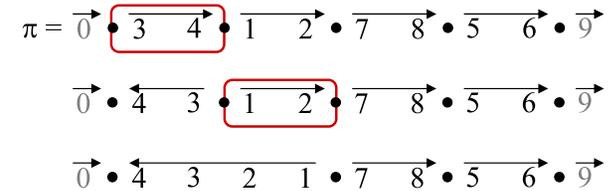
$$\overrightarrow{0} \bullet \overleftarrow{4 \ 3} \bullet \overrightarrow{1 \ 2} \bullet \overrightarrow{7 \ 8} \bullet \overrightarrow{5 \ 6} \bullet \overrightarrow{9}$$

$$\iota = \overrightarrow{0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9}$$

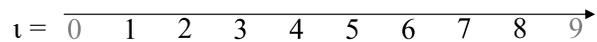
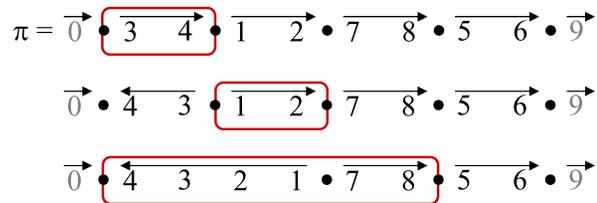
**Algoritmo Guloso para Distância de Reversão
sem Orientação dos Genes**



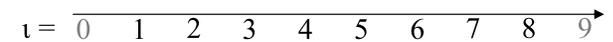
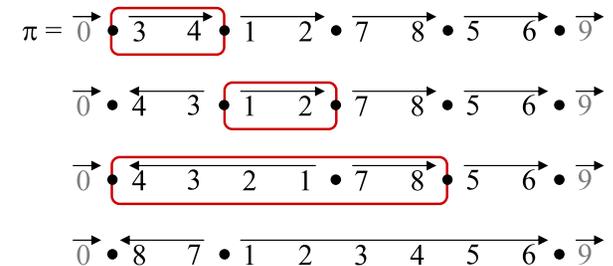
**Algoritmo Guloso para Distância de Reversão
sem Orientação dos Genes**



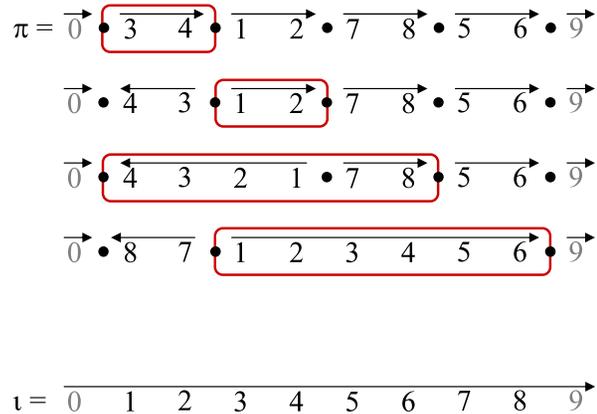
**Algoritmo Guloso para Distância de Reversão
sem Orientação dos Genes**



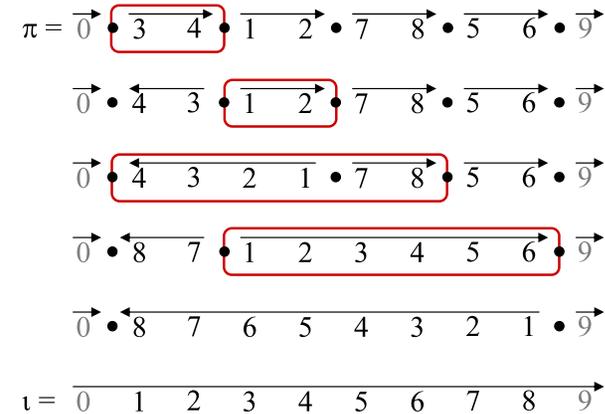
**Algoritmo Guloso para Distância de Reversão
sem Orientação dos Genes**



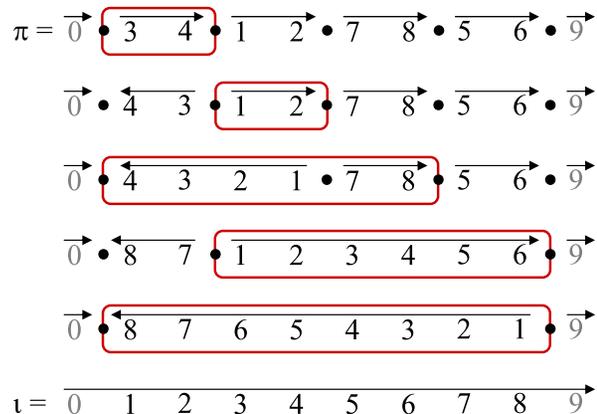
Algoritmo Guloso para Distância de Reversão sem Orientação dos Genes



Algoritmo Guloso para Distância de Reversão sem Orientação dos Genes



Algoritmo Guloso para Distância de Reversão sem Orientação dos Genes



Algoritmo Guloso para Distância de Reversão sem Orientação dos Genes

- Complexidade: $O(n^2)$.
- O algoritmo ordena qualquer permutação usando, no máximo, $b(\pi)$ reversões.
- Sendo assim, temos que:

$$\frac{b(\pi)}{2} \leq d(\pi) \leq b(\pi)$$

logo, o algoritmo guloso é um algoritmo de aproximação com fator:

$$\frac{b(\pi)}{\frac{b(\pi)}{2}} = 2.$$

Ordenação Ótima por Reversões

$$\pi = \overrightarrow{0} \bullet \overrightarrow{3 \ 4} \bullet \overrightarrow{1 \ 2} \bullet \overrightarrow{7 \ 8} \bullet \overrightarrow{5 \ 6} \bullet \overrightarrow{9}$$

$$\iota = \overrightarrow{0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9}$$

Ordenação Ótima por Reversões

$$\pi = \overrightarrow{0} \bullet \overrightarrow{3 \ 4} \bullet \overrightarrow{1 \ 2} \bullet \overrightarrow{7} \bullet \boxed{\overrightarrow{8 \ 5 \ 6}} \bullet \overrightarrow{9}$$

$$\iota = \overrightarrow{0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9}$$

Ordenação Ótima por Reversões

$$\pi = \overrightarrow{0} \bullet \overrightarrow{3 \ 4} \bullet \overrightarrow{1 \ 2} \bullet \overrightarrow{7} \bullet \boxed{\overrightarrow{8 \ 5 \ 6}} \bullet \overrightarrow{9}$$

$$\overrightarrow{0} \bullet \overrightarrow{3 \ 4} \bullet \overrightarrow{1 \ 2} \bullet \overleftarrow{6 \ 5 \ 8} \bullet \overrightarrow{9}$$

$$\iota = \overrightarrow{0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9}$$

Ordenação Ótima por Reversões

$$\pi = \overrightarrow{0} \bullet \overrightarrow{3 \ 4} \bullet \overrightarrow{1 \ 2} \bullet \overrightarrow{7} \bullet \boxed{\overrightarrow{8 \ 5 \ 6}} \bullet \overrightarrow{9}$$

$$\overrightarrow{0} \bullet \boxed{\overrightarrow{3 \ 4 \ 1}} \bullet \overrightarrow{2} \bullet \overleftarrow{6 \ 5 \ 8} \bullet \overrightarrow{9}$$

$$\iota = \overrightarrow{0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9}$$

Principais Resultados para Distância de Reversão sem Orientação dos Genes

- John Kececioglu e David Sankoff, em 1995, apresentaram o algoritmo guloso com fator de aproximação 2 e conjecturaram que o problema de distância de reversão sem orientação é \mathcal{NP} -Completo.
- Vineet Bafna e Pavel Pevzner, em 1996, apresentaram um algoritmo com fator de aproximação 1.75.
- Alberto Caprara, em 1997, provou que o problema da distância de reversão sem orientação é \mathcal{NP} -Completo.
- David Christie, em 1998, apresentou um algoritmo com fator de aproximação 1.5.
- Piotr Berman e Marek Karpinski, em 1999, provaram que o problema da distância de reversão sem orientação é $\mathcal{MAX-SNP}$ -Difícil.
- Piotr Berman, Sridhar Hannenhalli e Marek Karpinski, em 2002, apresentaram um algoritmo com fator de aproximação 1.375.

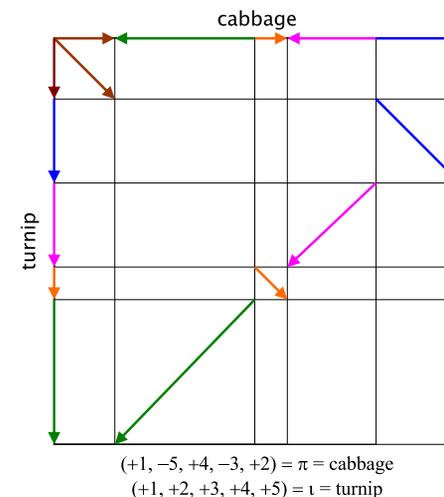
Ordenação por Reversões sem Orientação dos Genes com o Número Mínimo de Reversões Imposto pelo Limite Inferior de Breakpoints

- John Kececioglu e David Sankoff, em 1995, conjecturaram que o problema de decidir se uma permutação π pode ser ordenada usando exatamente $b(\pi)/2$ reversões é um problema \mathcal{NP} -Completo.
- Nicholas Tran, em 1997, provou que é possível decidir se uma permutação π pode ser ordenada usando exatamente $b(\pi)/2$ reversões em tempo $O(n^2 \log n)$.

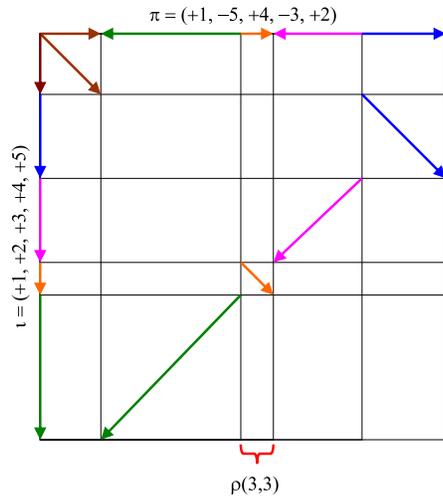
Problema da Distância de Reversão com Orientação de Genes

- Podemos representar um genoma com n blocos conservados com orientação dos genes conhecida como uma permutação sinalizada, $\pi = \pi_1 \pi_2 \dots \pi_n$, com $\pi_i \in \{-1, -2, \dots, -n, +1, +2, \dots, +n\}$, de tal forma que $|\pi_i| - |\pi_j| \neq 0$, para $1 \leq i < j \leq n$.
- A reversão $\rho(i, j)$, com $1 \leq i < j \leq n$, reverte a ordem de $\pi[i..j]$ e os sinais de todos os elementos pertencentes a este intervalo.
- *Distância de Reversão*: dados dois genomas compostos por n blocos conservados, representados pelas permutações sinalizadas π e σ , calcular a distância de reversão ($d(\pi, \sigma)$) entre π e σ , ou seja, obter uma série de reversões $\rho_1, \rho_2, \dots, \rho_t$, de tamanho mínimo, tal que $d(\pi, \sigma) = t$ e $\pi \cdot \rho_1 \cdot \rho_2 \dots \rho_t = \sigma$.
- *Ordenação por Reversões*: dado um genoma composto por n blocos conservados, representado pela permutação π , calcular a distância de reversão ($d(\pi)$) entre π e a permutação identidade $\iota = (+1, +2, \dots, +n)$, ou seja, $d(\pi) = d(\pi, \iota)$.

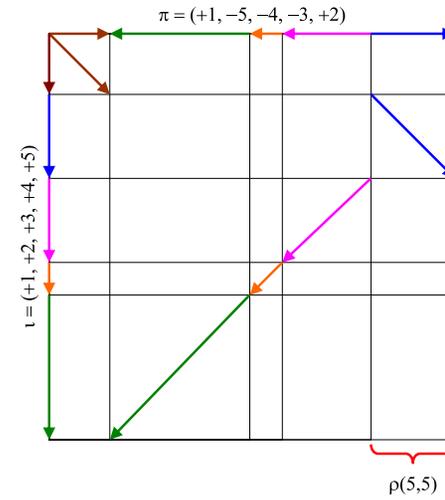
Transforming Cabbage into Turnip



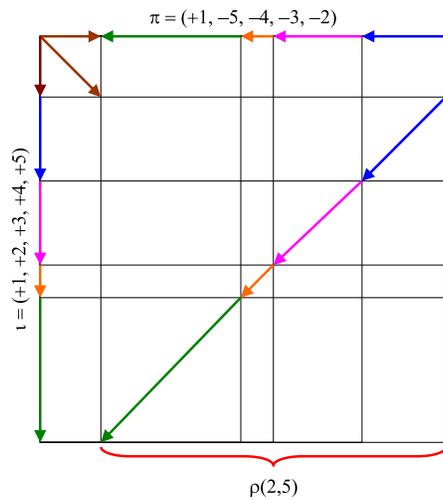
Transforming Cabbage into Turnip



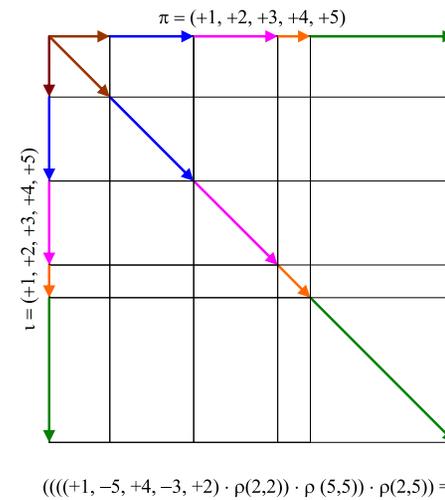
Transforming Cabbage into Turnip



Transforming Cabbage into Turnip



Transforming Cabbage into Turnip



$$(((+1, -5, +4, -3, +2) \cdot \rho(2,2)) \cdot \rho(5,5)) \cdot \rho(2,5) = \tau$$

Permutação Reduzida

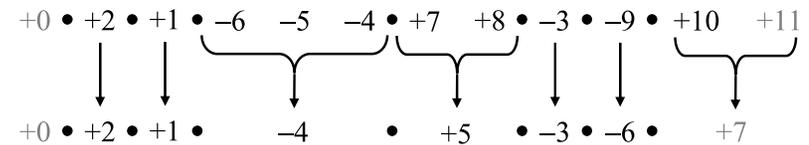
Breakpoints e Permutação Reduzida

Definição

Um par de elementos π_i e π_{i+1} , para $0 \leq i \leq n$, é uma adjacência se $\pi_{i+1} - \pi_i = 1$. Caso contrário, o par de elementos é chamado de breakpoint.

Definição

Uma permutação π é chamada reduzida se ela não contém adjacências.



Reversão Orientada

Definição

Um par orientado (π_i, π_j) é um par de elementos de π , tal que $i < j$, $||\pi_i| - |\pi_j|| \leq 1$ e $\pi_i \times \pi_j < 0$.

Definição

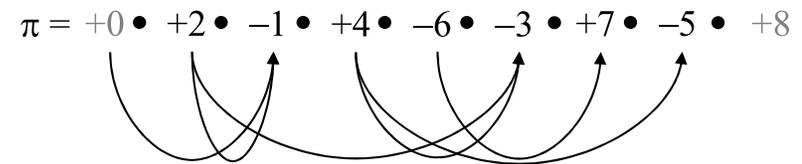
A pontuação de uma reversão ρ em relação a π , representada por $\text{score}(\pi, \rho)$, é o número de pares orientados em $\pi \cdot \rho$.

Definição

Seja (π_i, π_j) um par orientado. Logo, as seguintes reversões são chamadas orientadas:

- $\rho(i, j - 1)$, if $\pi_i + \pi_j = +1$.
- $\rho(i + 1, j)$, if $\pi_i + \pi_j = -1$.

Pares Orientados × Reversões Orientadas



$$(+0, -1) \Rightarrow \rho(1,2)$$

$$(+4, -3) \Rightarrow \rho(3,4)$$

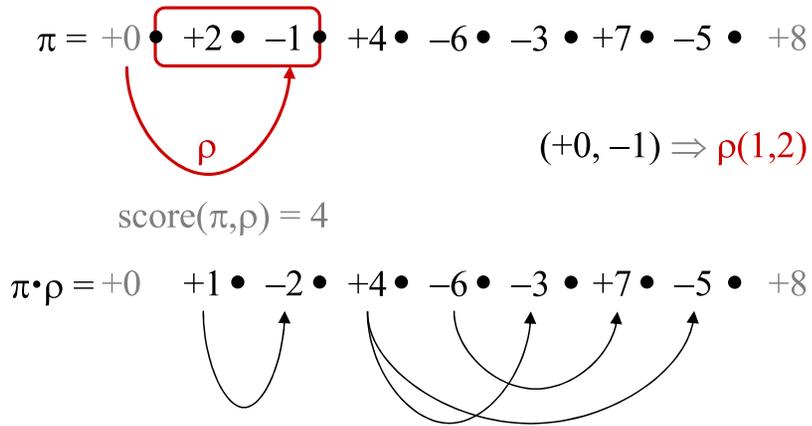
$$(+2, -1) \Rightarrow \rho(1,1)$$

$$(+4, -5) \Rightarrow \rho(4,7)$$

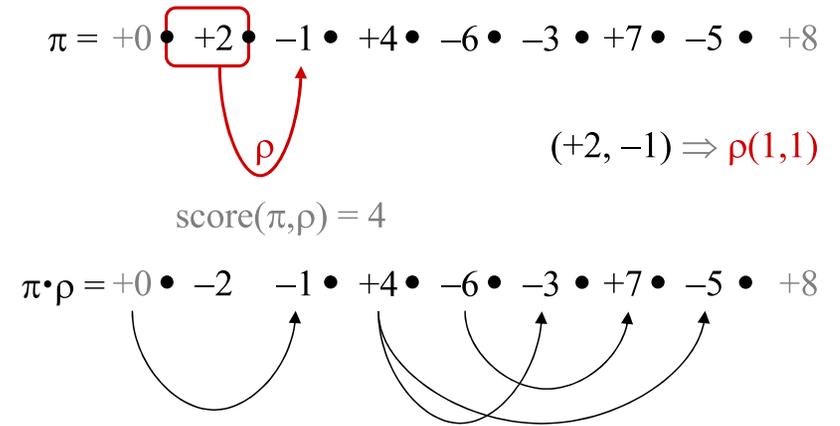
$$(+2, -3) \Rightarrow \rho(2,5)$$

$$(-6, +7) \Rightarrow \rho(4,5)$$

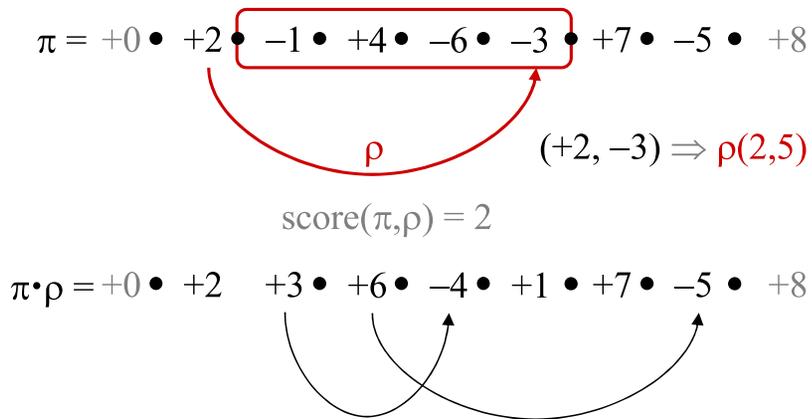
Pares Orientados × Reversões Orientadas



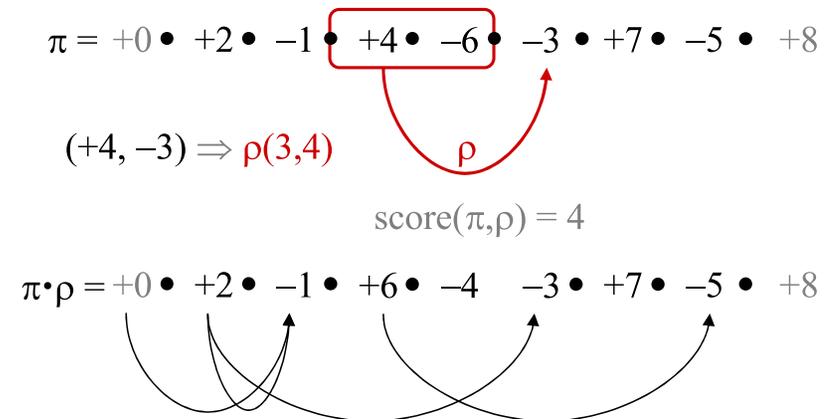
Pares Orientados × Reversões Orientadas



Pares Orientados × Reversões Orientadas



Pares Orientados × Reversões Orientadas



Pares Orientados × Reversões Orientadas

$$\pi = +0 \bullet +2 \bullet -1 \bullet +4 \bullet \boxed{-6 \bullet -3 \bullet +7 \bullet -5} \bullet +8$$

$$(+4, -5) \Rightarrow \rho(4,7)$$

ρ

$$\text{score}(\pi, \rho) = 4$$

$$\pi \cdot \rho = +0 \bullet +2 \bullet -1 \bullet +4 \bullet +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8$$



Pares Orientados × Reversões Orientadas

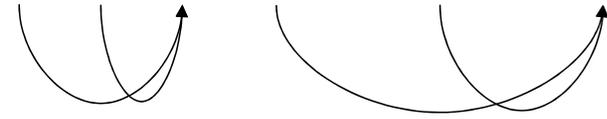
$$\pi = +0 \bullet +2 \bullet -1 \bullet +4 \bullet \boxed{-6 \bullet -3} \bullet +7 \bullet -5 \bullet +8$$

$$(-6, +7) \Rightarrow \rho(4,5)$$

ρ

$$\text{score}(\pi, \rho) = 4$$

$$\pi \cdot \rho = +0 \bullet +2 \bullet -1 \bullet +4 \bullet +3 \bullet +6 \bullet +7 \bullet -5 \bullet +8$$



Ordenação Usando Apenas Pares Ordenados

$$\pi^{(0)} = +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8$$

Pares Orientados × Reversões Orientadas

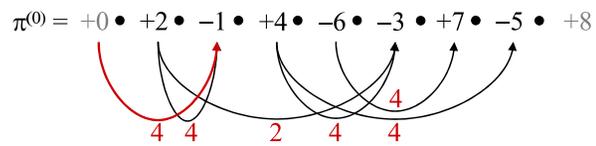
Lema

Uma reversão ρ é orientada em relação a π se e somente se $\Delta_b(\pi, \rho) < 0$.

Teorema

Seja $\pi^{(i)}$ uma permutação que contém pelo menos um par orientado. Seja ρ_i uma reversão de score máximo em relação a $\pi^{(i)}$. Defina $\pi^{(i+1)}$ como $\pi^{(i+1)} = (\pi^{(i)} \cdot \rho_i)$. Seja $\pi^{(1)}, \pi^{(2)}, \dots, \pi^{(k)}$ uma série maximal de permutações gerada a partir de $\pi^{(0)}$. Logo $\pi^{(k)}$ é formada apenas por elementos positivos e $d(\pi^{(0)}, \pi^{(k)}) = k$.

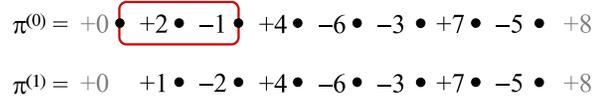
Ordenação Usando Apenas Pares Ordenados



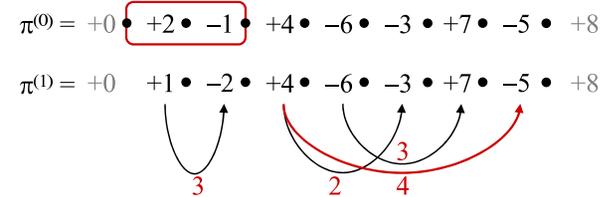
Ordenação Usando Apenas Pares Ordenados



Ordenação Usando Apenas Pares Ordenados



Ordenação Usando Apenas Pares Ordenados



Ordenação Usando Apenas Pares Ordenados

$$\pi^{(0)} = +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8$$

$$\pi^{(1)} = +0 \bullet +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8$$

Ordenação Usando Apenas Pares Ordenados

$$\pi^{(0)} = +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8$$

$$\pi^{(1)} = +0 \bullet +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8$$

$$\pi^{(2)} = +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8$$

Ordenação Usando Apenas Pares Ordenados

$$\pi^{(0)} = +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8$$

$$\pi^{(1)} = +0 \bullet +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8$$

$$\pi^{(2)} = +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8$$

Ordenação Usando Apenas Pares Ordenados

$$\pi^{(0)} = +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8$$

$$\pi^{(1)} = +0 \bullet +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8$$

$$\pi^{(2)} = +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8$$

Ordenação Usando Apenas Pares Ordenados

$$\begin{aligned} \pi^{(0)} &= +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(1)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(2)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8 \\ \pi^{(3)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -6 \bullet -3 \bullet +7 \bullet +8 \end{aligned}$$

Ordenação Usando Apenas Pares Ordenados

$$\begin{aligned} \pi^{(0)} &= +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(1)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(2)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8 \\ \pi^{(3)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -6 \bullet -3 \bullet +7 \bullet +8 \end{aligned}$$

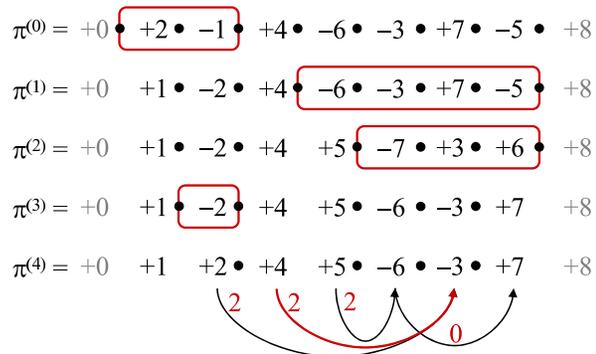
Ordenação Usando Apenas Pares Ordenados

$$\begin{aligned} \pi^{(0)} &= +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(1)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(2)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8 \\ \pi^{(3)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -6 \bullet -3 \bullet +7 \bullet +8 \end{aligned}$$

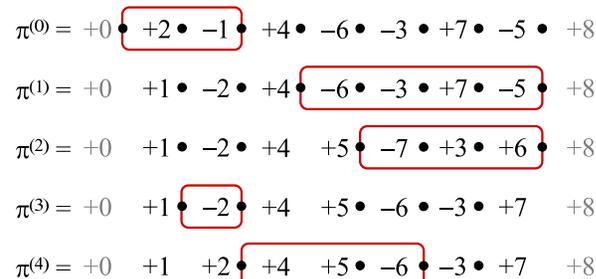
Ordenação Usando Apenas Pares Ordenados

$$\begin{aligned} \pi^{(0)} &= +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(1)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(2)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8 \\ \pi^{(3)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -6 \bullet -3 \bullet +7 \bullet +8 \\ \pi^{(4)} &= +0 \bullet +1 \bullet +2 \bullet +4 \bullet +5 \bullet -6 \bullet -3 \bullet +7 \bullet +8 \end{aligned}$$

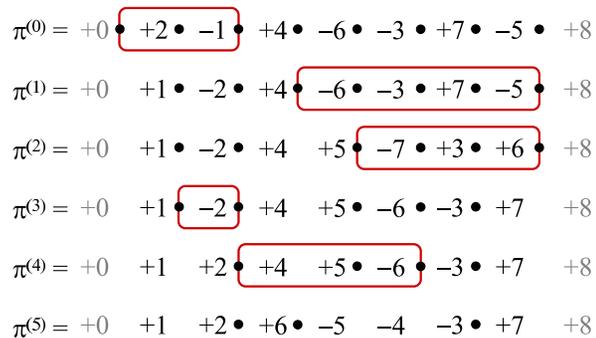
Ordenação Usando Apenas Pares Ordenados



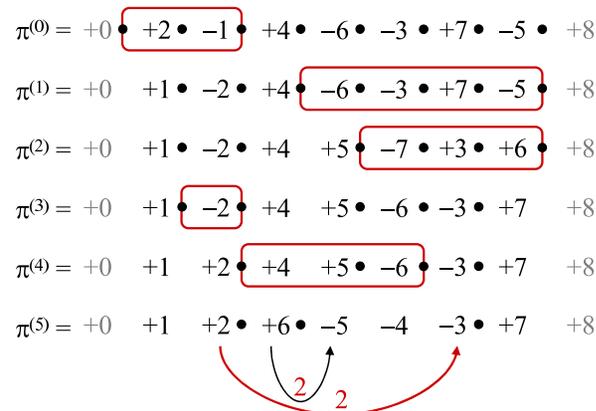
Ordenação Usando Apenas Pares Ordenados



Ordenação Usando Apenas Pares Ordenados



Ordenação Usando Apenas Pares Ordenados



Ordenação Usando Apenas Pares Ordenados

$$\begin{aligned} \pi^{(0)} &= +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(1)} &= +0 \quad +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(2)} &= +0 \quad +1 \bullet -2 \bullet +4 \quad +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8 \\ \pi^{(3)} &= +0 \quad +1 \bullet -2 \bullet +4 \quad +5 \bullet -6 \bullet -3 \bullet +7 \quad +8 \\ \pi^{(4)} &= +0 \quad +1 \quad +2 \bullet +4 \quad +5 \bullet -6 \bullet -3 \bullet +7 \quad +8 \\ \pi^{(5)} &= +0 \quad +1 \quad +2 \bullet +6 \bullet -5 \quad -4 \quad -3 \bullet +7 \quad +8 \end{aligned}$$

Ordenação Usando Apenas Pares Ordenados

$$\begin{aligned} \pi^{(0)} &= +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(1)} &= +0 \quad +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(2)} &= +0 \quad +1 \bullet -2 \bullet +4 \quad +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8 \\ \pi^{(3)} &= +0 \quad +1 \bullet -2 \bullet +4 \quad +5 \bullet -6 \bullet -3 \bullet +7 \quad +8 \\ \pi^{(4)} &= +0 \quad +1 \quad +2 \bullet +4 \quad +5 \bullet -6 \bullet -3 \bullet +7 \quad +8 \\ \pi^{(5)} &= +0 \quad +1 \quad +2 \bullet +6 \bullet -5 \quad -4 \quad -3 \bullet +7 \quad +8 \\ \pi^{(6)} &= +0 \quad +1 \quad +2 \quad +3 \quad +4 \quad +5 \bullet -6 \bullet +7 \quad +8 \end{aligned}$$

Ordenação Usando Apenas Pares Ordenados

$$\begin{aligned} \pi^{(0)} &= +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(1)} &= +0 \quad +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(2)} &= +0 \quad +1 \bullet -2 \bullet +4 \quad +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8 \\ \pi^{(3)} &= +0 \quad +1 \bullet -2 \bullet +4 \quad +5 \bullet -6 \bullet -3 \bullet +7 \quad +8 \\ \pi^{(4)} &= +0 \quad +1 \quad +2 \bullet +4 \quad +5 \bullet -6 \bullet -3 \bullet +7 \quad +8 \\ \pi^{(5)} &= +0 \quad +1 \quad +2 \bullet +6 \bullet -5 \quad -4 \quad -3 \bullet +7 \quad +8 \\ \pi^{(6)} &= +0 \quad +1 \quad +2 \quad +3 \quad +4 \quad +5 \bullet -6 \bullet +7 \quad +8 \end{aligned}$$


Ordenação Usando Apenas Pares Ordenados

$$\begin{aligned} \pi^{(0)} &= +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(1)} &= +0 \quad +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\ \pi^{(2)} &= +0 \quad +1 \bullet -2 \bullet +4 \quad +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8 \\ \pi^{(3)} &= +0 \quad +1 \bullet -2 \bullet +4 \quad +5 \bullet -6 \bullet -3 \bullet +7 \quad +8 \\ \pi^{(4)} &= +0 \quad +1 \quad +2 \bullet +4 \quad +5 \bullet -6 \bullet -3 \bullet +7 \quad +8 \\ \pi^{(5)} &= +0 \quad +1 \quad +2 \bullet +6 \bullet -5 \quad -4 \quad -3 \bullet +7 \quad +8 \\ \pi^{(6)} &= +0 \quad +1 \quad +2 \quad +3 \quad +4 \quad +5 \bullet -6 \bullet +7 \quad +8 \end{aligned}$$

Ordenação Usando Apenas Pares Ordenados

$$\begin{aligned}
 \pi^{(0)} &= +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\
 \pi^{(1)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\
 \pi^{(2)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8 \\
 \pi^{(3)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -6 \bullet -3 \bullet +7 \bullet +8 \\
 \pi^{(4)} &= +0 \bullet +1 \bullet +2 \bullet +4 \bullet +5 \bullet -6 \bullet -3 \bullet +7 \bullet +8 \\
 \pi^{(5)} &= +0 \bullet +1 \bullet +2 \bullet +6 \bullet -5 \bullet -4 \bullet -3 \bullet +7 \bullet +8 \\
 \pi^{(6)} &= +0 \bullet +1 \bullet +2 \bullet +3 \bullet +4 \bullet +5 \bullet -6 \bullet +7 \bullet +8 \\
 \pi^{(7)} &= +0 \bullet +1 \bullet +2 \bullet +3 \bullet +4 \bullet +5 \bullet +6 \bullet +7 \bullet +8
 \end{aligned}$$

Ordenação Usando Apenas Pares Ordenados

$$\begin{aligned}
 \pi^{(0)} &= +0 \bullet +2 \bullet -1 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\
 \pi^{(1)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet -6 \bullet -3 \bullet +7 \bullet -5 \bullet +8 \\
 \pi^{(2)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -7 \bullet +3 \bullet +6 \bullet +8 \\
 \pi^{(3)} &= +0 \bullet +1 \bullet -2 \bullet +4 \bullet +5 \bullet -6 \bullet -3 \bullet +7 \bullet +8 \\
 \pi^{(4)} &= +0 \bullet +1 \bullet +2 \bullet +4 \bullet +5 \bullet -6 \bullet -3 \bullet +7 \bullet +8 \\
 \pi^{(5)} &= +0 \bullet +1 \bullet +2 \bullet +6 \bullet -5 \bullet -4 \bullet -3 \bullet +7 \bullet +8 \\
 \pi^{(6)} &= +0 \bullet +1 \bullet +2 \bullet +3 \bullet +4 \bullet +5 \bullet -6 \bullet +7 \bullet +8 \\
 \iota &= +0 \bullet +1 \bullet +2 \bullet +3 \bullet +4 \bullet +5 \bullet +6 \bullet +7 \bullet +8
 \end{aligned}$$

Hurdles

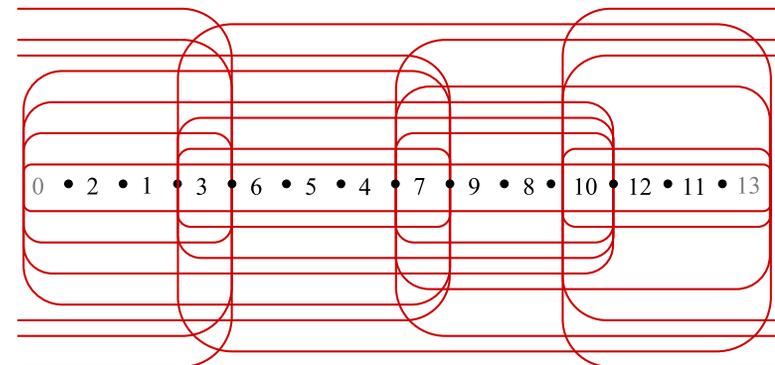
Definição

Seja π uma permutação reduzida formada apenas por elementos positivos (logo, sem pares orientados). Suponha que π foi estendida, com $\pi_0 = 0$ e $\pi_{n+1} = n + 1$, e circularizada, considerando que o elemento 0 é consecutivo ao elemento $n + 1$. Um framed interval em π é um intervalo da forma $i \pi_{j+1} \pi_{j+2} \dots \pi_{j+k-1} i + k$, tal que todos inteiros entre i e $i + k$ pertencem ao intervalo $[i..i + k]$ (considerado de forma circular).

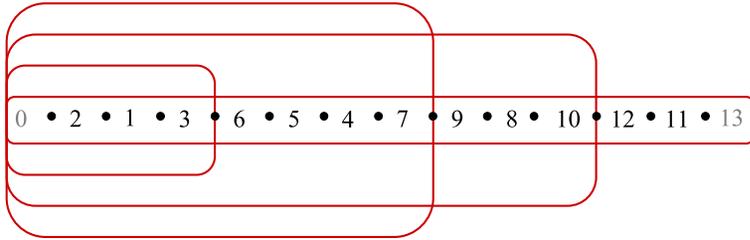
Definição

Seja π uma permutação reduzida formada apenas por elementos positivos. Um hurdle em π é um framed interval que não contém outros framed intervals.

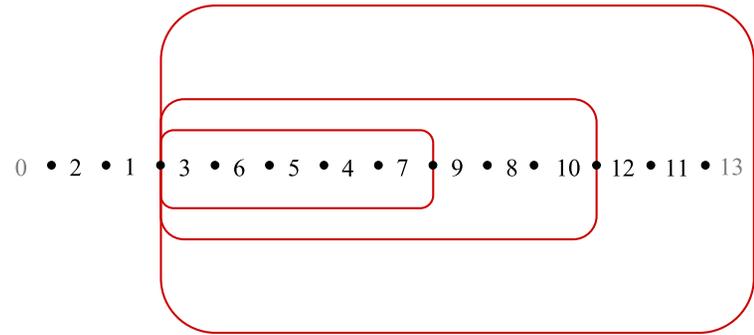
Framed Intervals



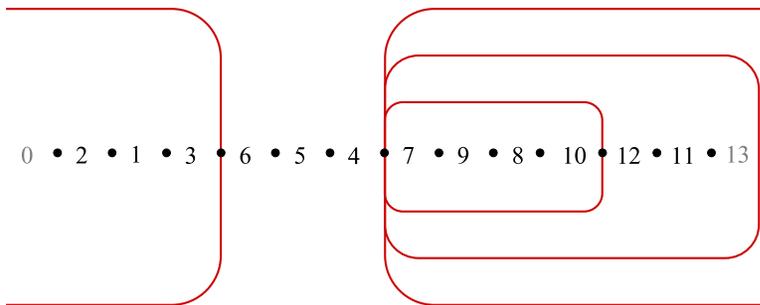
Framed Intervals



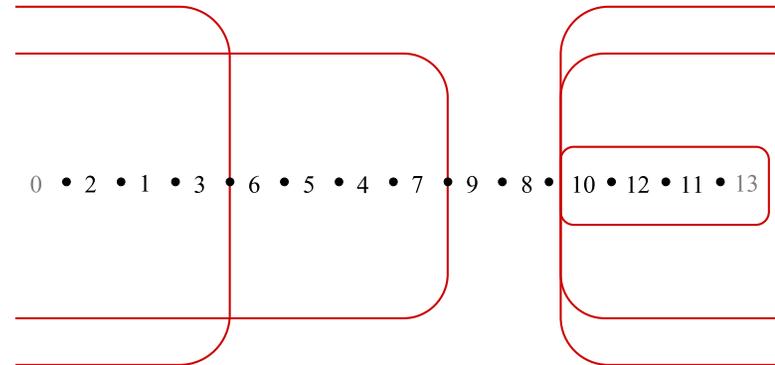
Framed Intervals



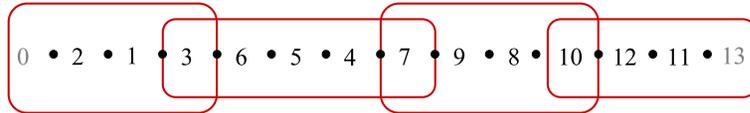
Framed Intervals



Framed Intervals



Hurdles



Hurdles

Definição

A posição do elemento j na permutação π é indicada por π_j^{-1} .

Definição

Uma reversão ρ corta um hurdle i $\pi_{j+1} \pi_{j+2} \dots i+1 \dots \pi_{j+k-1} i+k$ se $\rho = \rho(\pi_i^{-1} + 1, \pi_{i+1}^{-1} - 1)$, ou seja, se reverte os elementos entre i e $i+1$.

Definição

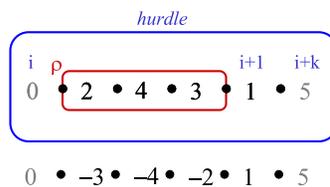
Uma reversão ρ une dois hurdles $i \dots i+k \dots i' \dots i'+k'$ da permutação π se $\rho = \rho(\pi_{i+k}^{-1}, \pi_{i'}^{-1})$, ou seja, se reverte os elementos entre $i+k$ e i' (inclusive ambos).

Definição

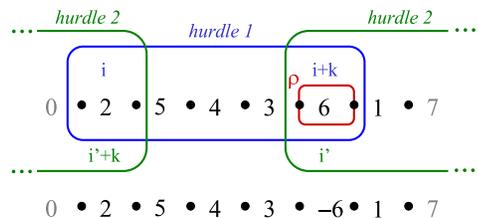
Um hurdle é chamado de simples se quando cortado o número de hurdles diminui. Caso contrário, o hurdle é chamado de super.

Cutting Hurdles × Merging Hurdles

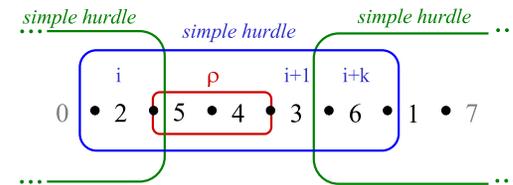
Cutting Hurdles



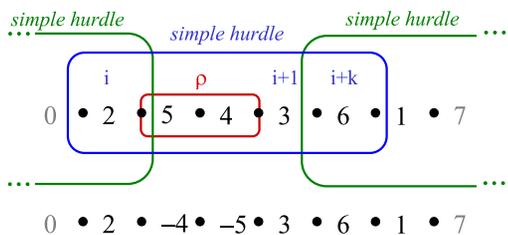
Merging Hurdles



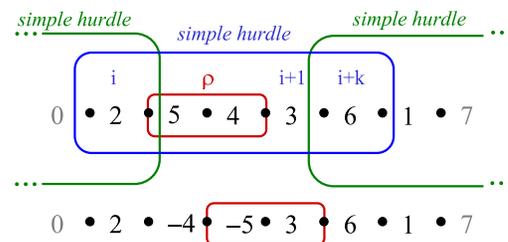
Simple Hurdle



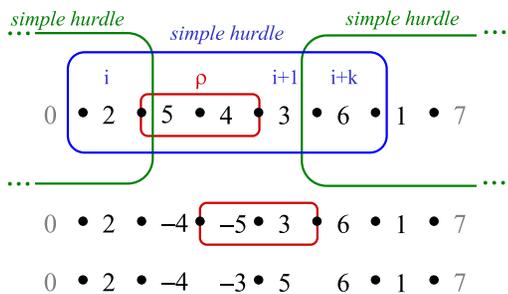
Simple Hurdle



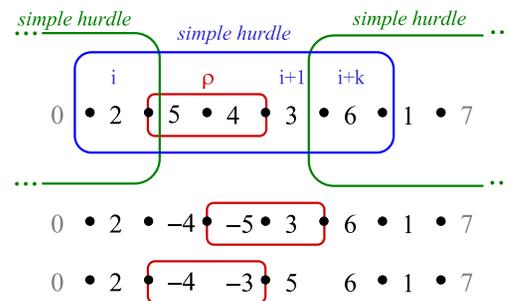
Simple Hurdle



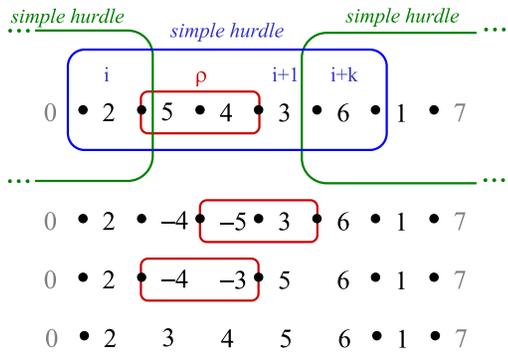
Simple Hurdle



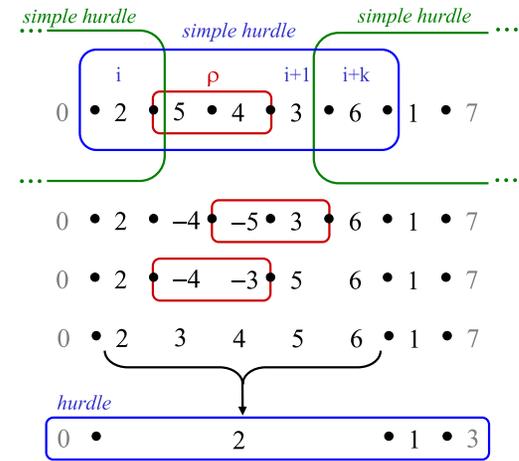
Simple Hurdle



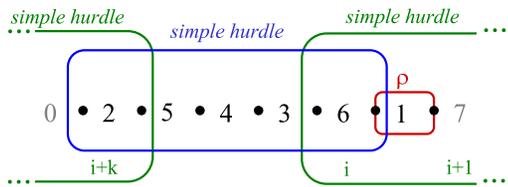
Simple Hurdle



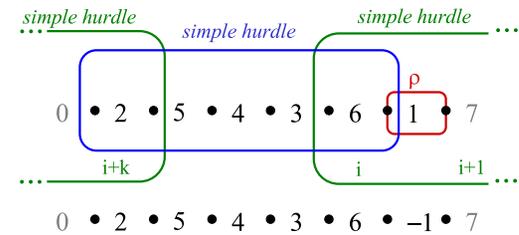
Simple Hurdle



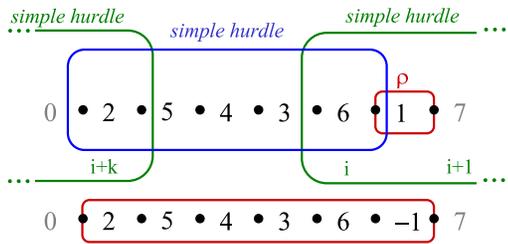
Simple Hurdle



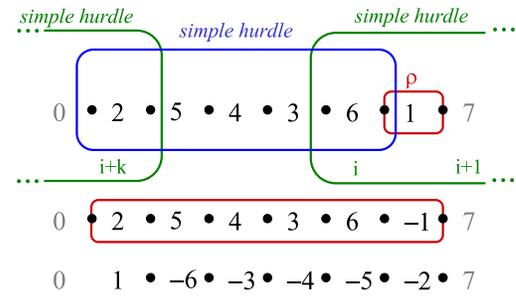
Simple Hurdle



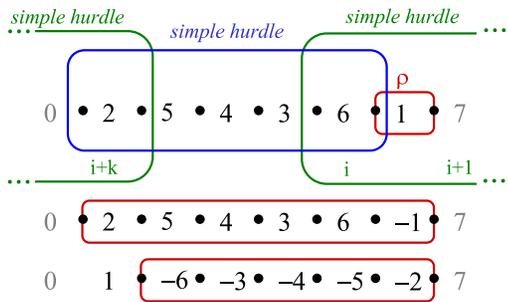
Simple Hurdle



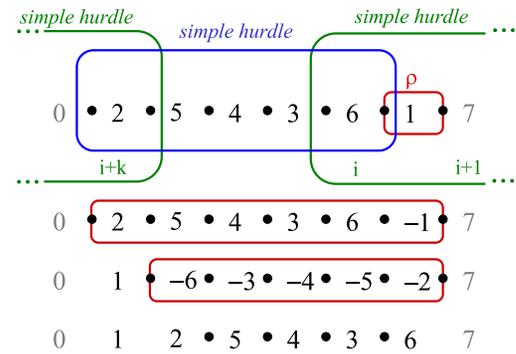
Simple Hurdle



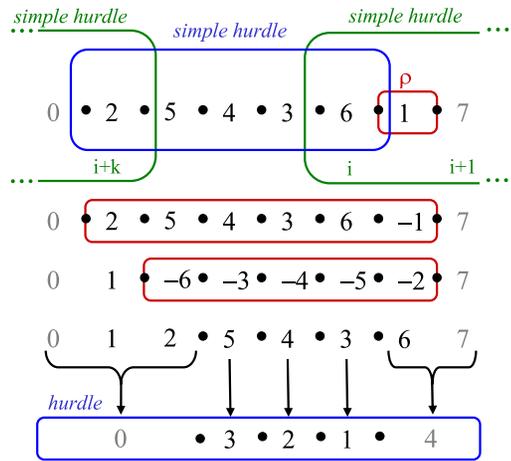
Simple Hurdle



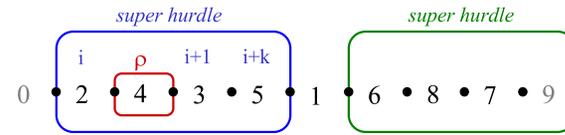
Simple Hurdle



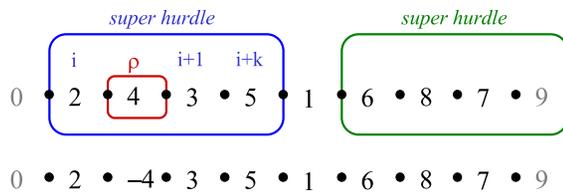
Simple Hurdle



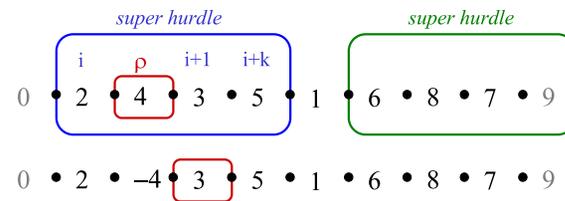
Super Hurdle



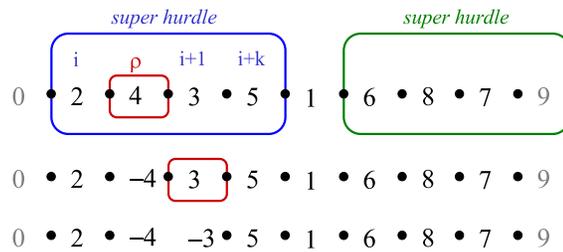
Super Hurdle



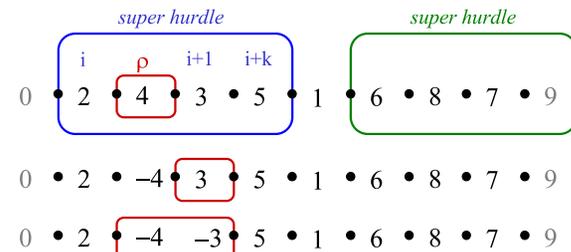
Super Hurdle



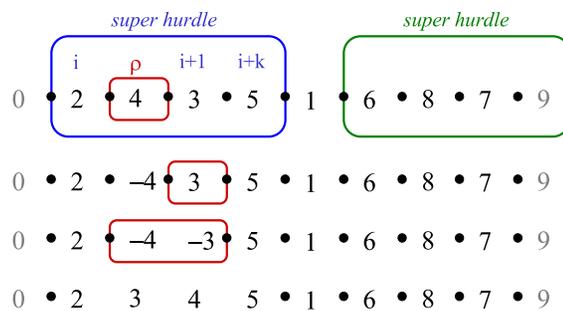
Super Hurdle



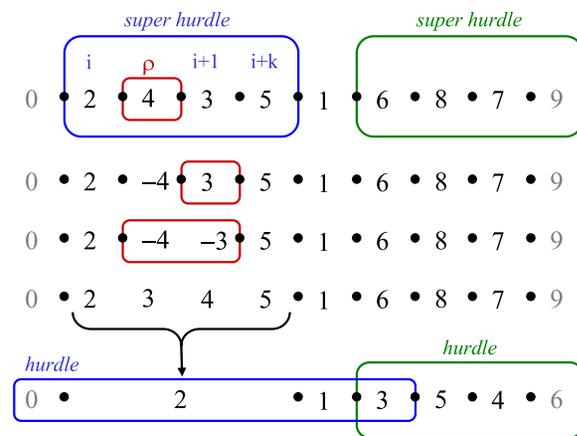
Super Hurdle



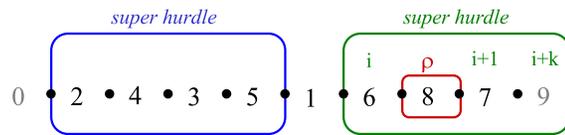
Super Hurdle



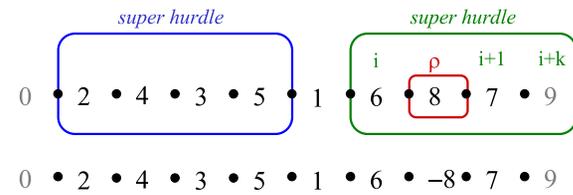
Super Hurdle



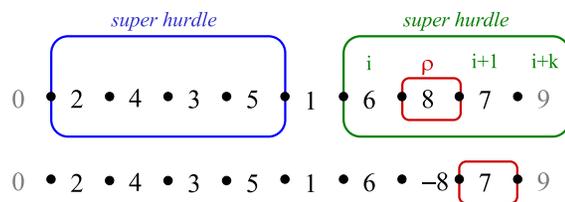
Super Hurdle



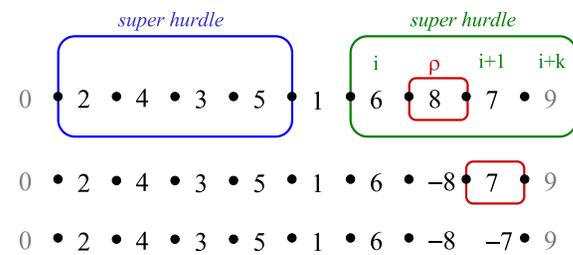
Super Hurdle



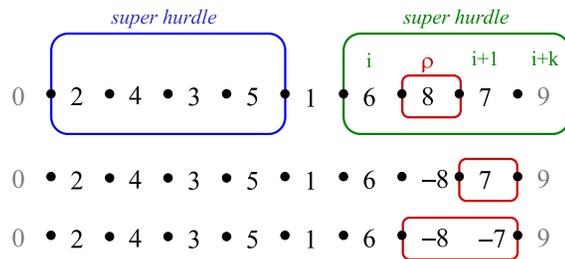
Super Hurdle



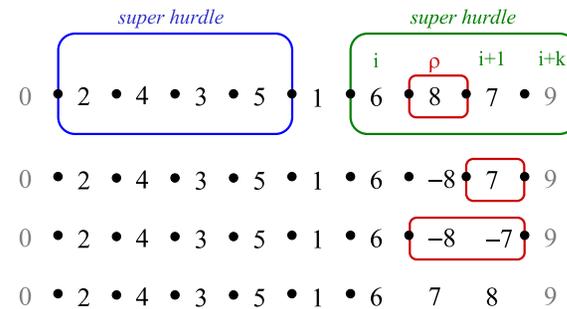
Super Hurdle



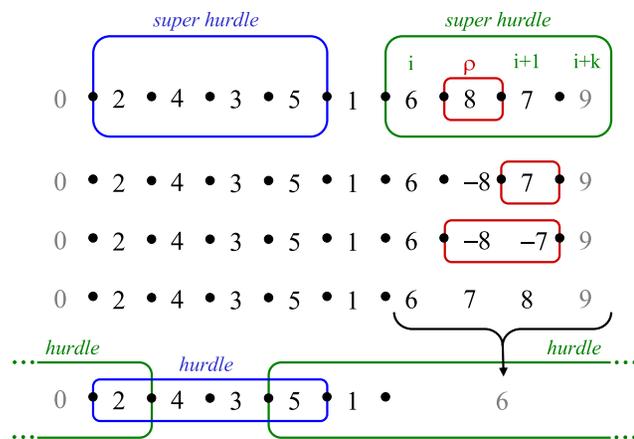
Super Hurdle



Super Hurdle



Super Hurdle



Algoritmo Ótimo para Distância de Reversão com Orientação dos Genes

Algoritmo 4: Optimal Sorting by Reversal

```

Input:  $\pi, n$ 
 $t \leftarrow 0$ 
while  $\pi \neq \iota$  do
  if  $\pi$  has a oriented pair then  $\rho \leftarrow$  the reversal that has maximal score else
    if  $\pi$  has 2k hurdles then
      if  $\pi$  has 2 hurdles then  $\rho \leftarrow$  the reversal that merges the two hurdles
      else  $\rho \leftarrow$  any reversal that merges two non-consecutives hurdles
    end
    if  $\pi$  has only one hurdle then  $\rho \leftarrow$  the reversal that cuts the hurdle else
      if  $\pi$  has a simple hurdle then  $\rho \leftarrow$  any reversal that cuts a simple hurdle else
        if  $\pi$  has 3 hurdles then  $\rho \leftarrow$  any reversal that merges two hurdles
        else  $\rho \leftarrow$  any reversal that merges two non-consecutives hurdles
      end
    end
  end
   $\pi \leftarrow \pi \cdot \rho$ 
   $t \leftarrow t + 1$ 
end
return t

```

Algoritmo Ótimo para Distância de Reversão com Orientação dos Genes

- Complexidade:
 - ▶ Determinar todos os pares orientados de uma permutação: $O(n)$.
 - ▶ Determinar a reversão com maior *score*: $O(n) \times O(n) = O(n^2)$.
 - ▶ Determinar todos os *hurdles*: $O(n^2)$.
 - ▶ Total: $O(n) \times O(n^2) = O(n^3)$.
- Algoritmo proposto por Anne Bergeron, em 2004.
- O problema da distância de reversão com orientação de genes foi originalmente resolvido em tempo polinomial por Sridhar Hannenhalli e Pavel Pevzner, em 1999.
- David Bader, Bernard Moret e Mi Yan, em 2001, mostraram que é possível calcular a distância de reversão com orientação de genes conhecida (sem listar as reversões utilizadas), em $O(n)$.
- Krister Swenson, Yu Lin, Vaibhav Rajan e Bernard Moret, em 2008, provaram que a chance de uma permutação aleatória (com sinal) possuir um *hurdle* é de $\Theta(n^{-2})$.