Réarrangements génomiques

Informatique Génomique - Master 1

Guillaume Blin

IGM-LabInfo UMR 8049,
Bureau 4B066
Université de Marne La Vallée
gblin@univ-mlv.fr
http://igm.univ-mlv.fr/∼gblin

2007-08

Plan

Evolution des génomes

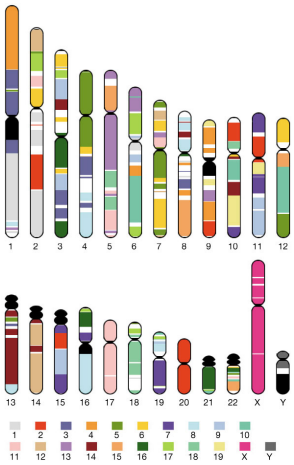
Les opérations évolutives

- Mutations ponctuelles
  - substitutions, insertions, délétions de nucléotides
- Réarrangements génomiques
  - Modifications dans l'organisation des génomes
- Mesure de l'évolution et comparaison entre espèces en analysant les mutations ponctuelles.

Evolution des génomes

Comparaison Souris-Humain

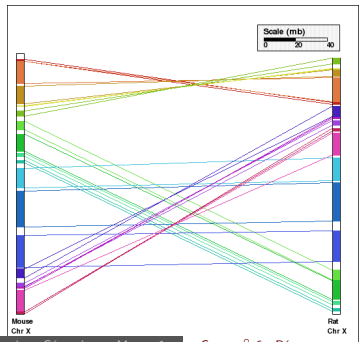
- Chromosomes Humain pourvus de segments communs avec la souris
- Chaque couleur représente un chromosome spécifique du génome de la souris



Evolution des génomes

Modélisation des génomes

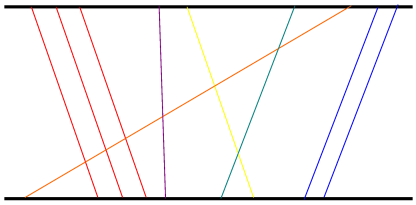
- Etape 1 : Recherche de synténies (portions communes de chromosomes)



Evolution des génomes

Modélisation des génomes

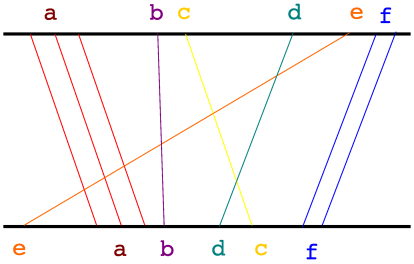
- Etape 2 : Réencodage



Evolution des génomes

Modélisation des génomes

- Etape 2 : Réencodage



- Homme : (a b c d e f)
- Souris : (e a b d c f)

Evolution des génomes

Opérations de réarrangements génomiques

- Opérations sur un seul chromosome
  - Délétion : $abc \rightarrow ac$
  - Insertion : $abc \rightarrow abdc$
  - Duplication (ou non) : $abcd \rightarrow abcbdb$
  - Inversion (Reversal) : $ab_1b_2b_3c \rightarrow ab_3b_2b_1c$
  - Transposition : $abcde \rightarrow adbce$
  - Transposition inverse : $abcde \rightarrow adcbe$
- Opérations sur deux chromosomes
  - Translocations : $abcd;\alpha\beta\gamma \rightarrow ab\gamma;\alpha\beta cd$
  - Fusion : $abcd;\alpha\beta\gamma \rightarrow abcd\alpha\beta\gamma$
  - Fission : $abcde \rightarrow abc;de$

Evolution des génomes

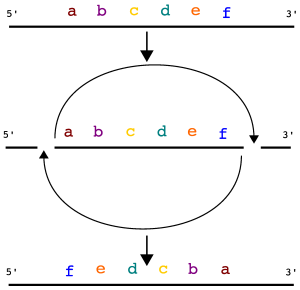
Problématique biologique

- Étant donnés 2 génomes, trouver un scénario évolutif permettant de passer de l'un à l'autre
- Intérêt pour
  - la longueur de ce scénario (= nb d'évènements évolutifs) comme une mesure de distance évolutive
  - les réarrangements impliqués dans ce scénario
- On se restreint à l'inversion, un réarrangement très fréquent

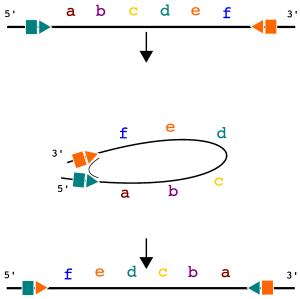
Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Evolution des génomes

Les mécanismes de l'inversion



- Cassure + mauvais réassemblage



- Crossing-over intrachromosomique

Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Evolution des génomes

Modélisation informatique

- Les génomes sont représentés par des permutations signées sur l'ensemble $\{1, 2, \ldots, n\}$ – les gènes sont signés : $+i$ ou $-i$, suivant le brin à partir duquel ils sont transcrits
- L'inversion d'un intervalle de la permutation inverse l'ordre et l'orientation des éléments de cet intervalle
- Le problème SBR :
  - Entrée : 2 permutations signées
  - Sortie : Trouver un scénario d'inversion de longueur minimale pour passer de l'une à l'autre
- Trouver la longueur de ce scénario correspond au problème de la distance d'inversion.

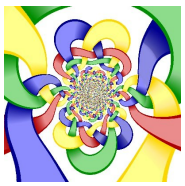
Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Evolution des génomes

Exemple : permutation et inversion

- Sans perte de généralité, on considère qu'une des 2 permutations est l'identité (renommage)
- D'où le nom de `Sorting By Reversals`
- Permutation signée sur $\{1, 2, \ldots, 8\}$

- π_1 = `-2 -1 4 3 5 -8 6 7`



Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Evolution des génomes

Exemple : permutation et inversion

- Sans perte de généralité, on considère qu'une des 2 permutations est l'identité (renommage)
- D'où le nom de `Sorting By Reversals`
- Permutation signée sur $\{1, 2, \ldots, 8\}$

- π_1 = `-2 -1 4 3 5 -8 6 7`

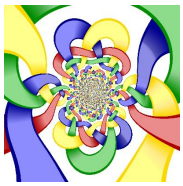
Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Evolution des génomes

Exemple : permutation et inversion

- Sans perte de généralité, on considère qu'une des 2 permutations est l'identité (renommage)
- D'où le nom de `Sorting By Reversals`
- Permutation signée sur $\{1, 2, \ldots, 8\}$

- π'_1 = `(0 -2 -1 4 3 5 -8 6 7 9)`



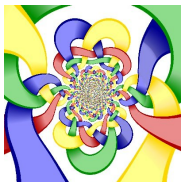
Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Evolution des génomes

Exemple : permutation et inversion

- Sans perte de généralité, on considère qu'une des 2 permutations est l'identité (renommage)
- D'où le nom de `Sorting By Reversals`
- Permutation signée sur $\{1, 2, \ldots, 8\}$

- π'_1 = `(0 -2 -1 4 3 5 -8 6 7 9)`
- π'_1 = `(0 -4 1 2 3 5 -8 6 7 9)`



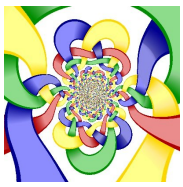
Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Evolution des génomes

Exemple : permutation et inversion

- Sans perte de généralité, on considère qu'une des 2 permutations est l'identité (renommage)
- D'où le nom de `Sorting By Reversals`
- Permutation signée sur $\{1, 2, \ldots, 8\}$

- π'_1 = `(0 -2 -1 4 3 5 -8 6 7 9)`
- π'_1 = `(0 -4 1 2 3 5 -8 6 7 9)`
- π'_1 = `(0 -3 -2 -1 4 5 -8 6 7 9)`



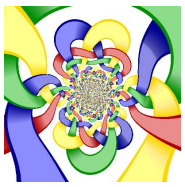
Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Evolution des génomes

Exemple : permutation et inversion

- Sans perte de généralité, on considère qu'une des 2 permutations est l'identité (renommage)
- D'où le nom de `Sorting By Reversals`
- Permutation signée sur $\{1, 2, \ldots, 8\}$

- π'_1 = `(0 -2 -1 4 3 5 -8 6 7 9)`
- π'_1 = `(0 -4 1 2 3 5 -8 6 7 9)`
- π'_1 = `(0 -3 -2 -1 4 5 -8 6 7 9)`
- π'_1 = `(0 1 2 3 4 5 -8 6 7 9)`

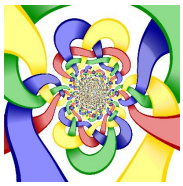


Evolution des génomes

Exemple : permutation et inversion

- Sans perte de généralité, on considère qu'une des 2 permutations est l'identité (renommage)
- D'où le nom de `Sorting By Reversals`
- Permutation signée sur $\{1, 2, \ldots, 8\}$

- π'_1 = (0 −2 −1 4 3 5 −8 6 7 9)
- π'_1 = (0 −4 1 2 3 5 −8 6 7 9)
- π'_1 = (0 −3 −2 −1 4 5 −8 6 7 9)
- π'_1 = (0 1 2 3 4 5 −8 6 7 9)
- π'_1 = (0 1 2 3 4 5 −7 −6 8 9)

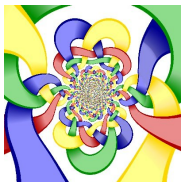


Evolution des génomes

Exemple : permutation et inversion

- Sans perte de généralité, on considère qu'une des 2 permutations est l'identité (renommage)
- D'où le nom de `Sorting By Reversals`
- Permutation signée sur $\{1, 2, \ldots, 8\}$

- π'_1 = (0 −2 −1 4 3 5 −8 6 7 9)
- π'_1 = (0 −4 1 2 3 5 −8 6 7 9)
- π'_1 = (0 −3 −2 −1 4 5 −8 6 7 9)
- π'_1 = (0 1 2 3 4 5 −8 6 7 9)
- π'_1 = (0 1 2 3 4 5 −7 −6 8 9)
- π'_1 = (0 1 2 3 4 5 6 7 8 9)

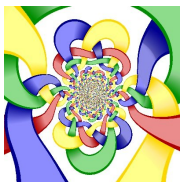


Evolution des génomes

Exemple : permutation et inversion

- Sans perte de généralité, on considère qu'une des 2 permutations est l'identité (renommage)
- D'où le nom de `Sorting By Reversals`
- Permutation signée sur $\{1, 2, \ldots, 8\}$

- π'_1 = (0 −2 −1 4 3 5 −8 6 7 9)
- π'_1 = (0 −4 1 2 3 5 −8 6 7 9)
- π'_1 = (0 −3 −2 −1 4 5 −8 6 7 9)
- π'_1 = (0 1 2 3 4 5 9)
- π'_1 = (0 1 2 3 4 5 −7 −6 8 9)
- π'_1 = (0 1 2 3 4 5 6 7 8 9) – $d(\pi_1) = 5$



Historique

- 1992 – Sankoff 1ère formulation du problème
- 1995 – Hannenhalli et Pevzner "Transforming Cabbage into Turnip"
  1er algorithme polynomial et LA théorie d'H.-P.
- 1996 – Berman et Hannenhalli Amélioration de la complexité
- 1999 – Kaplan, Shamir et Tarjan "A Faster and Simpler Algorithm for SBR"
- 2001 – Bergeron "A Very Elementary Presentation of the H-P Theory"
- 2001 – Bader, Moret et Yan Algorithme linéaire pour le problème de la distance
- 2004 – Tannier et Sagot "Sorting by reversals in subquadratic time"

Plan

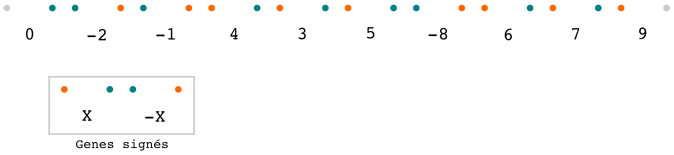
Les permutations signées

Les points : adjacences et breakpoints

- π_1 = (0 −2 −1 4 3 5 −8 6 7 9)
- Un point $p \cdot q$ est défini par une paire d'éléments consécutifs dans la permutation
  Exemple : $0 \cdot -2$ et $-2 \cdot -1$ sont les deux premiers points de π_1
- Un point de la forme $i \cdot i+1$, ou $-(i+1) \cdot -i$ est une adjacence, autrement c'est un breakpoint
  Exemple : $-2 \cdot -1$ et $6 \cdot 7$ sont des adjacences ; tous les autres points de π_1 sont des breakpoints

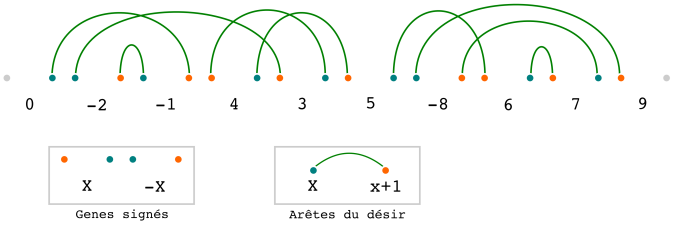
Les permutations signées

Le graphe de breakpoint



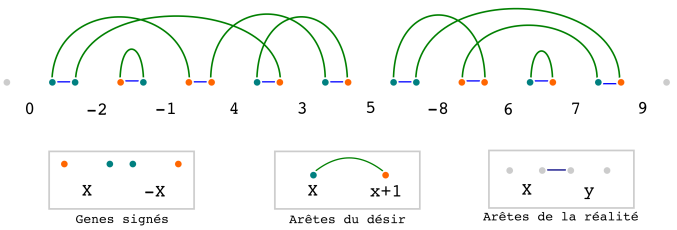
Les permutations signées

Le graphe de breakpoint



Les permutations signées

Le graphe de breakpoint



Les permutations signées

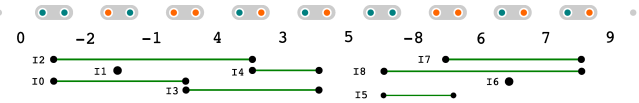
La formule d'Hannenhalli et Pevzner

- $d(\pi) = n - c(\pi) + h(\pi) + f(\pi)$ où
  - $c(\pi)$: nombre de cycles de la permutation π
  - $h(\pi)$: nombre de hurdles de la permutation π
  - $f(\pi) = 1$ si π est une forteresse, $f(\pi) = 0$ sinon
- Hurdles : un certain type de composants non orientés pouvant être simple si son élimination fait diminuer le nombre total de hurdles ; ou super hurdle sinon
- Forteresse : une permutation qui a un nombre impair de hurdles et qui sont toutes des super hurdles

Les permutations signées

Les intervalles élémentaires

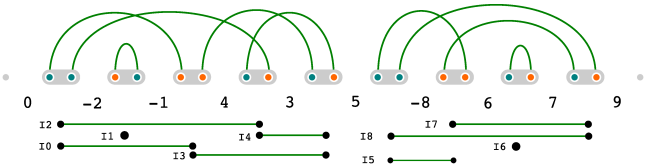
- À chaque paire d'éléments non signés $(k, k+1)$, $0 \leq k < n$, on associe l'intervalle élémentaire I_k dont les extrémités sont :
  - Le point droit de k, si k est positif, son point gauche sinon.
  - Le point gauche de $k+1$, si $k+1$ est positif, son point droit sinon.
  - Les éléments k et $k+1$ sont appelés les extrémités de l'intervalle élémentaire



Les permutations signées

Les intervalles élémentaires

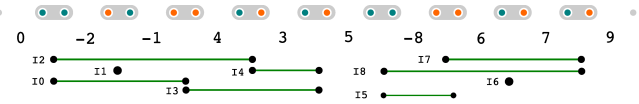
- Intervalles élémentaires = arêtes du désir



Les permutations signées

Les intervalles élémentaires

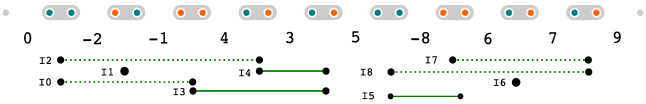
- Un intervalle élémentaire peut contenir 0, 1 ou 2 de ses extrémités. (i.e. k ou $k+1$)
  Exemple :
  - I_0 contient une de ses extrémités,
  - I_3 contient les deux,
  - I_5 n'en contient aucune.
- Les intervalles élémentaires vides, tels que I_1 et I_6 , correspondent à des adjacences de la permutation.



Les permutations signées

Intervalles orientés et non orientés

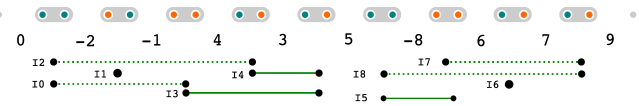
- Un intervalle dont les extrémités ont des signes différents est dit orienté, sinon il est non orienté.
- Les intervalles orientés sont exactement ceux qui contiennent une seule de leurs extrémités
- Les intervalles orientés jouent un rôle crucial puisqu'ils peuvent être utilisés pour créer des adjacences.



Les permutations signées

Intervalles orientés et non orientés

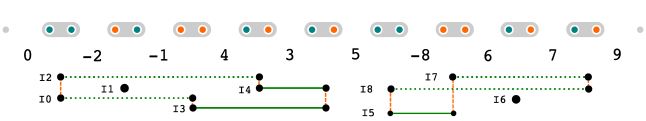
- Inverser un intervalle orienté I_k crée, dans la permutation résultante, soit l'adjacence $k \cdot k+1$ soit l'adjacence $-(k+1) \cdot -k$
- Exactement 2 intervalles élémentaires se rencontrent à chaque breakpoint de la permutation



Les permutations signées

Les cycles

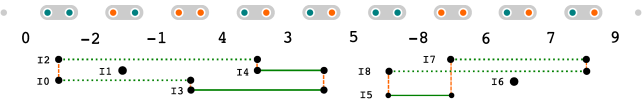
- Un cycle est une séquence $b_1, b_2, \ldots, b_k$ de points tels que 2 points successifs soient les extrémités d'un intervalle élémentaire, y compris b_k et b_1
- Les adjacences sont des cycles triviaux consistant en un seul point
- Un cycle contient toujours un nombre pair d'intervalles orientés



Les permutations signées

La relation de chevauchement

- Deux intervalles élémentaires I et J se chevauchent si chacun contient exactement une des extrémités de l'autre.
- Les intervalles qui se rencontrent à un breakpoint peuvent se chevaucher ou non.
  Exemples :
  - I_0 et I_2 se chevauchent (I_0 contient -2, et I_2 contient 1)
  - I_0 et I_3 ne se chevauchent pas
  - I_3 et I_4 se chevauchent (I_3 contient 4, et I_4 contient 3)

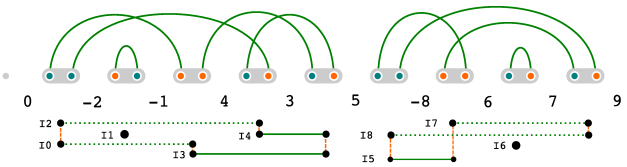


Les permutations signées

La relation de chevauchement

Exemples :

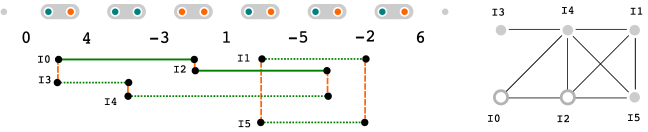
- I_0 et I_2 se chevauchent (I_0 contient -2, et I_2 contient 1)
- I_0 et I_3 ne se chevauchent pas
- I_3 et I_4 se chevauchent (I_3 contient 4, et I_4 contient 3)



Les permutations signées

Le graphe de chevauchement

- Une permutation et son graphe de chevauchement O
- Deux sommets sont connectés dans O ssi les intervalles correspondants se chevauchent.
  - Sommets pleins pour les intervalles orientés
  - Sommets vides pour les intervalles non orientés.



Plan

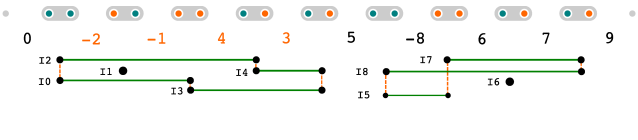
Les inversions

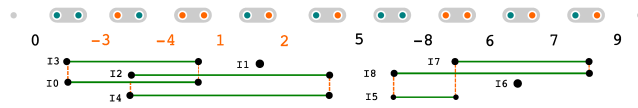
Inversions et nombre de cycles

- Une inversion peut seulement modifier le nombre de cycles par $+1$, 0, ou -1
- Preuve :
  - Une inversion échange les éléments de deux points d'une permutation
  - Si ces deux points appartiennent au même cycle, alors
    - soit le cycle est coupé en deux, $(+1)$
    - soit il est conservé mais avec des breakpoints différents. (0)
  - Si les deux points appartiennent à des cycles différents, alors
    - ces deux cycles sont fusionnés. (-1)
  - (Kececioglu et Sankoff, 1994)

Les inversions

Inversions et nombre de cycles

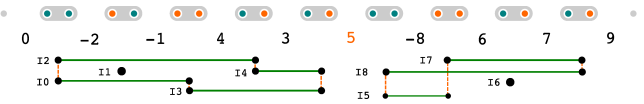


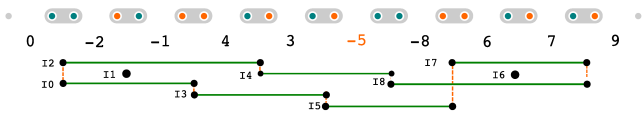


- Un cycle coupé en deux $\Leftrightarrow +1$

Les inversions

Inversions et nombre de cycles

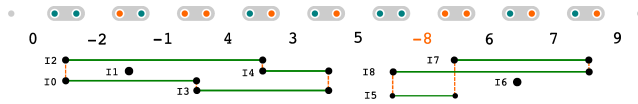


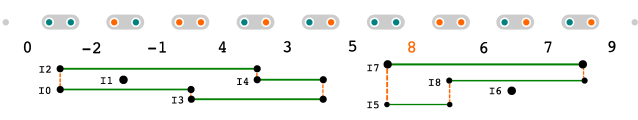


- Deux cycles fusionnés en un seul $\Leftrightarrow -1$

Les inversions

Inversions et nombre de cycles





- Aucun changement dans le nombre de cycles $\Leftrightarrow 0$

Les inversions

Inversion d'un intervalle orienté

- Inverser un intervalle orienté crée une adjacence
- Une adjacence est un cycle trivial
- Inverser un intervalle orienté augmente de $+1$ le nombre de cycles
- La permutation identité sur l'ensemble $\{0, 1, 2, \ldots, n\}$ est la seule avec n cycles, tous des adjacences.
- Puisque au plus un cycle peut être ajouté par une inversion, une première borne inférieure pour la distance d'inversion :

$$d(\pi) \geq n - c(\pi)$$

Les inversions

Inversion et graphe de chevauchement

- Si I et J se chevauchent, alors inverser l'intervalle I change l'orientation de J, puisqu'une seule des extrémités de J change de signe
- Quand deux intervalles J et K chevauchent un intervalle I, l'effet de l'inversion de I est de complémenter la relation de chevauchement entre J et K :
  - si J et K se chevauchaient avant l'inversion, ils ne se chevauchent plus après ;
  - si J et K ne se chevauchaient pas avant, ils se chevauchent après

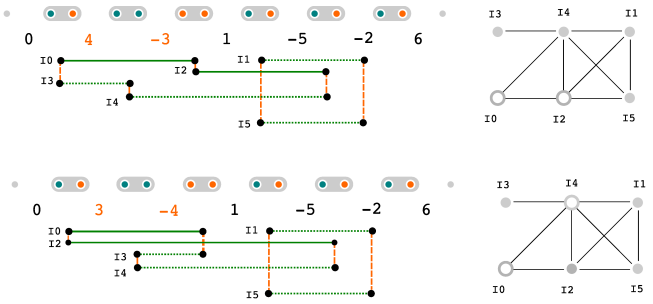
Les inversions

Inversion et graphe de chevauchement

- Formellement :
- Soit GI le sous-graphe du graphe de chevauchement formé par le sommet I et ses sommets voisins.
- Considérons l'inversion de l'intervalle élémentaire I
  1. Si I est non orienté, les effets sur le graphe de chevauchement sont de changer la couleur de tous les sommets dans $GI - \{I\}$, et de complémenter les arêtes de $GI - \{I\}$
  2. Si I est orienté, les effets sur le graphe de chevauchement sont de changer la couleur de tous les sommets de GI, et de complémenter les arêtes de GI.

Les inversions

Inversion d'un intervalle non orienté



Les inversions

Les composants

- Les points, les intervalles élémentaires et les cycles sont organisés en structures d'ordre supérieur appelés composants.
- Un composant de P est un intervalle de i à $(i + j)$ ou de $-(i + j)$ à $-i$, pour $j > 0$, dont l'ensemble des éléments non-signés est $\{i, \ldots, i + j\}$, et tel qu'il ne soit pas l'union de tels intervalles.
  Exemple :

  $\pi_2 = (0\ {-3}\ 1\ 2\ 4\ 6\ 5\ 7\ {-15}\ {-13}\ {-14}\ {-12}\ {-10}\ {-11}\ {-9}\ 8\ 16).$

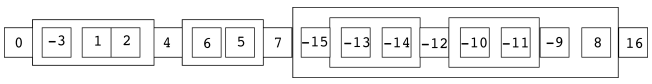
  - π_2 a six composants :
    $(0 \ldots 4), (4 \ldots 7), (7 \ldots 16), (1 \ldots 2), (-15 \ldots -12)$ et $(-12 \ldots -9)$

Les inversions

Diagramme en boîtes

- Les composants d'une permutation peuvent être représentés par un diagramme de boîtes

$\pi_2 = (0\ {-3}\ 1\ 2\ 4\ 6\ 5\ 7\ {-15}\ {-13}\ {-14}\ {-12}\ {-10}\ {-11}\ {-9}\ 8\ 16).$

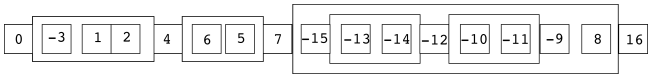


- Un point $p \cdot q$ appartient au plus petit composant qui contient à la fois p et q
- Les extrémités d'un intervalle élémentaire appartiennent au même composant, donc tous les points d'un cycle appartiennent au même composant

Les inversions

Diagramme en boîtes

$\pi_2 = (0\ {-3}\ 1\ 2\ 4\ 6\ 5\ 7\ {-15}\ {-13}\ {-14}\ {-12}\ {-10}\ {-11}\ {-9}\ 8\ 16).$

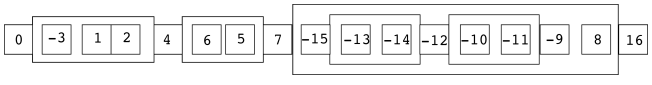


- Le signe d'un point $p \cdot q$ est positif si à la fois p et q sont positifs, il est négatif si à la fois p et q le sont
- Un composant est non orienté si il a un ou plusieurs breakpoints, et que tous ont le même signe, sinon le composant est orienté
  Exemple :
  - $(4 \ldots 7), (-15 \ldots -12)$ et $(-12 \ldots -9)$ sont non orientés
  - $(0 \ldots 4), (7 \ldots 16)$ sont orientés

Les inversions

Diagramme en boîtes

$\pi_2 = (0\ {-3}\ 1\ 2\ 4\ 6\ 5\ 7\ {-15}\ {-13}\ {-14}\ {-12}\ {-10}\ {-11}\ {-9}\ 8\ 16).$



- Deux composants distincts d'une permutation sont soit disjoints, soit emboîtés avec des extrémités différentes, soit ils se chevauchent d'un élément
- Quand deux composants se chevauchent d'un élément, on dit qu'ils sont liés
- Une suite de composants liés forme une chaîne
- Une chaîne qui ne peut pas être étendue par la gauche ou par la droite est dite maximale

Les inversions

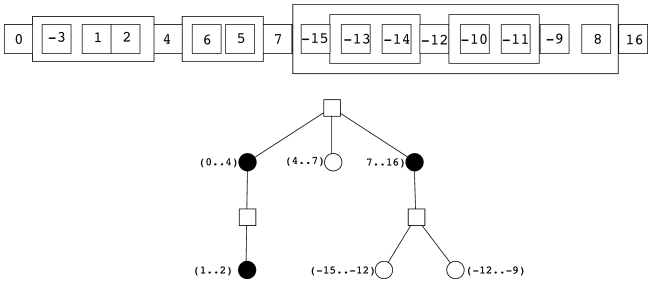
Représentation en arbre : T_π

- Soit une permutation π sur l'ensemble $\{0, 1, \ldots, n\}$ et ses composants, l'arbre T_π est défini par la construction suivante :
  1. Chaque composant est représenté par un noeud rond
  2. Chaque chaîne maximale est représentée par un noeud carré dont les fils (ordonnés) sont les noeuds ronds qui représentent les composants de la chaîne
  3. Un noeud carré est le fils du plus petit composant qui contient cette chaîne

Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Les inversions

Représentation en arbre : T_π



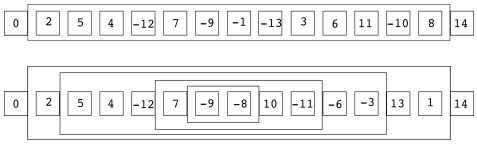
- L'arbre T_π associé à la permutation π

Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Les inversions

Effets d'une inversion sur les composants

- Une inversion peut créer ou détruire un nombre quelconque de composants. Exemple : Inversion de $(-1 \ldots 8) \Rightarrow$ 4 nouveaux composants



- Créer des composants orientés, ou des adjacences, est généralement considéré comme un pas vers le tri optimal d'une permutation.
- Cependant, la création de composants non orientés devrait être évité. Heureusement peu d'inversions ont cet effet.

Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Plan

Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Les algorithmes

Les différents algorithmes

- Le problème SBR
  - Berman et Hannenhalli 1996, $O(n^2\alpha(n))$
  - Kaplan, Shamir et Tarjan 1999, $O(n^2)$
  - Bergeron 2001, $O(n^2)$ ou $O(n^3)$
  - Tannier et Sagot 2004, $O(n\sqrt{n}\log n)$
- Le problème de la distance d'inversion
  - Bader, Moret et Yan 2001, $O(n)$

Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Les algorithmes

Principe de A. Bergeron, J. Mixtacki et J. Stoye

1. Élimination des composants non orientés
   - *Hurdles cutting*
   - *Hurdles merging*
2. Trier les composants orientés
   - Trouver les *safe inversions* ($d(\pi \cdot p) = d(\pi) - 1$)
3. Calcul de la distance d'inversion

Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Les algorithmes

Élimination des composants non orientés

- HurdleCutting
  - Si un composant C est non orienté, l'inversion d'un intervalle élémentaire dont les extrémités appartiennent à C oriente C, et laisse le nombre de cycles de la permutation inchangé
- HurdleMerging
  - Une inversion qui a ses deux extrémités dans des composants différents A et B détruit, ou oriente, tous les composants sur le chemin de A à B dans T_π , sans créer de nouveau composants non orientés.

Anatomie d'une permutation signée Les effets d'une inversion Les algorithmes

Les algorithmes

Trier les composants orientés

- Question : Comment choisir des inversions sûres ? (i.e. celles qui ne crée par de nouveaux composants non orientés)
- Le score d'une inversion est défini comme le nombre d'intervalles orientés dans la permutation résultante.
- Théorème [Bergeron, 2001] L'inversion d'un intervalle orienté de score maximum ne crée pas de nouveaux composants non orientés.
- Corollaire Si une permutation π n'est constituée que de composants orientés, $d(\pi) = n - c(\pi)$

Les algorithmes

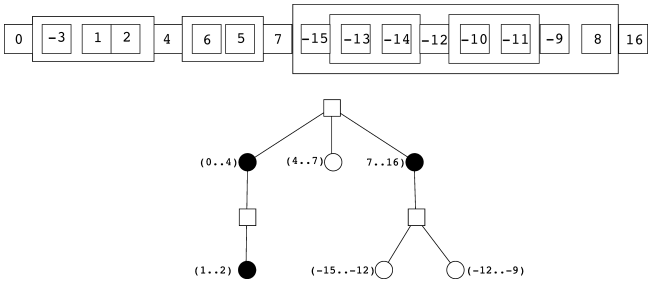
Calcul de la distance d'inversion

- On travaille sur l'arbre T_π
- Une couverture C de T_π est un ensemble de chemins joignant tous les composants non orientés de π, et tel que chaque noeud extrémité d'un chemin appartient à un seul chemin.
- *HurdleCutting* et *HurdleMerging* $\Rightarrow$ chaque couverture de T_π décrit une séquence d'inversions qui oriente π
- Un chemin est long s'il contient au moins 2 composants non orienté, et son coût est 2
- Un chemin est court et son coût est 1
- Le coût d'une couverture est la somme du coût de ses chemins ; une couverture est optimale si son coût est minimal (soit t ce coût)

$$d(\pi) = n - c(\pi) + t$$

Les algorithmes

Représentation en arbre : T_π

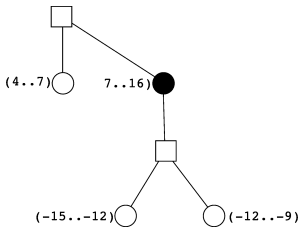


- Le long chemin joignant les composants $(4 \dots 7)$ et $(-12 \dots -9)$ détruirait ces composants, ainsi que $(7 \dots 16)$.

Les algorithmes

Une formule simple pour calculer t

- Soit T le plus petit sous-arbre non enraciné de T_π qui contient tous les composants non orientés
- Toutes les feuilles de T sont donc des composants non orientés, mais les noeuds ronds internes peuvent être des composants orientés
- Une branche est l'ensemble des noeuds sur le chemin montant d'une feuille jusqu'à un noeud de degrés ≥ 3, mais excluant celui-ci
- Une branche est courte si elle contient 1 seul composant non orienté ; longue sinon



Les algorithmes

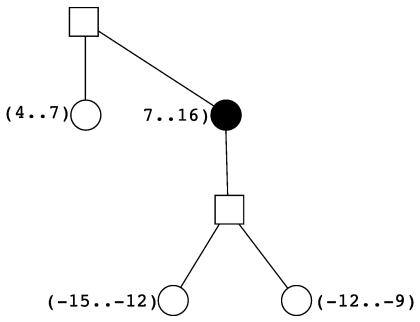
Une formule simple pour calculer t

1. Si T a $2k$ feuilles, alors $t = 2k$.
2. Si T a $2k + 1$ feuilles, et au moins une d'entre elles est sur une branche courte, alors $t = 2k + 1$.
3. Si T a $2k + 1$ feuilles, et aucune d'entre elles est sur une branche courte, alors $t = 2k + 2$.

Les algorithmes

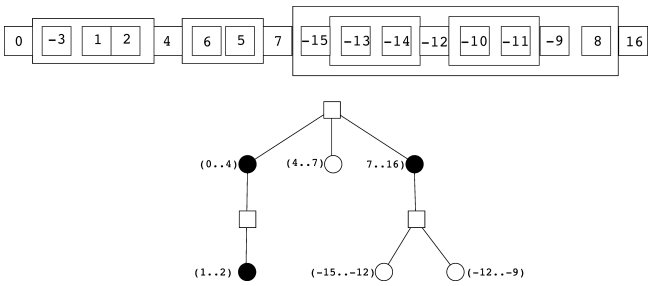
Une formule simple pour calculer t

- 3 feuilles, chacune sur des branches courtes $\Rightarrow t = 3$



Les algorithmes

Calcul de la distance d'inversion



- Pour cette permutation π : $n = 16$, $c(\pi) = 6$, $t = 3$, donc $d(\pi) = 13$

Plan

Bibliographie

Références

1. Anne Bergeron, Julia Mixtacki et Jens Stoye. *The inversion distance problem*. Chapitre 10 du livre Mathematics of Evolution and Phylogeny, 2005.
2. Michal Ozery-Flato et Ron Shamir. *Two notes on genome rearrangement*. Journal of Bioinformatics and Computational Biology, 2003