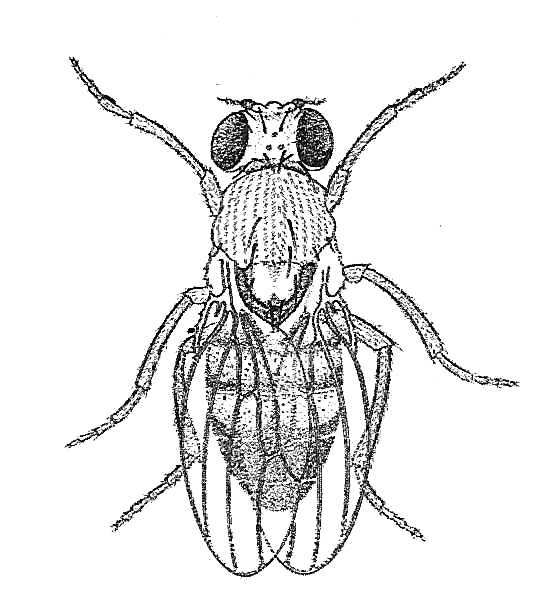
**Fruitvliegen**

Algoritmes om het ene genoom om te zetten in het andere middels inversie

Met illustratie voor Remco



Remco Blom

HvA?

Niels Pannekeet

HvA?

Renske Talsma

Vrije Universiteit

**1. Inleiding**

Fruitvliegen zijn kleine, tweevleugelige insecten die zich met name met suikers voeden. Het probleem in deze case heeft betrekking op twee soorten fruitvliegen, namelijk de Drosophila Melanogaster en Drosophila Miranda. Van deze fruitvliegen is veel over het genoom bekend. De twee soorten hebben identieke genen in het genoom, maar deze staan op een andere volgorde. Mutaties in het genoom van de vliegen vinden alleen plaats doordat series genen in hun geheel omkeren: inversie. In de onderstaande figuur is te zien hoe de genomen van respectievelijk Drosophila Melanogaster en Drosophila Miranda zijn opgebouwd. Voor deze case is het doel om uit te zoeken op welke manier de Drosophila Melanogaster in de Drosophila Miranda kan zijn veranderd middels mutaties, en wel in zo min mogelijk stappen en/of zo min mogelijk verplaatste genen totaal.



Figuur 1: De genomen van Drosophila Melanogaster en Drosophila Miranda

Doordat het genoom van de fruitvliegen 25 genen bevat, is het aantal mogelijke manieren waarop deze gesorteerd kunnen staan 25! ofwel 1.551121×1025. Het maximale aantal inversies dat nodig is om een elke mogelijke reeks van 25 genen goed te zetten is 24. Bij het omzetten van het genoom moet echter wel telkens rekening gehouden worden met het feit dat er alleen middels inversie veranderingen plaatsvinden, waardoor de minimale mutatiegrootte 2 genen bedraagt, omdat er bij een inversie altijd minstens 2 genen betrokken zijn.

**Wat is de specifieke toestandsruimte voor de vlieg? Is 24 daar ook het maximum? Wat kunnen we zeggen over de ondergrens? Wat kunnen we zeggen over het aantal inversies dat mogelijk is? Volgens mij had Niels daar iets over gezegd mbt beamsearch maar dat weet ik zelf niet precies meer. :P**

**2. Methodes**

**Elk algoritme beschrijven zodat het repliceerbaar is? Uiteindelijk ook tabel met stats toevoegen denk ik!!**

Er zijn voor dit onderzoek in totaal 8 verschillende algoritmes gebruikt. Omdat pas bij de laatste algoritmes echt sprake was van de gewenste resultaten worden de eerste versies slechts kort behandeld. De gebruikte algoritmes zijn Find and Swap, Find and Swap Reverse, Find and Swap Iterative, ChunkSwap, B-Star, BeamSearch en ElementShrink, Daarnaast is er telkens middels statistisch onderzoek geprobeerd om te achterhalen welke typen genomen zich makkelijk of juist moeilijk op volgorde laten zetten. Met deze kennis zijn de algoritmes vervolgens telkens verbeterd.

**2.1 Find and Swap en variaties**

Beschrijving….

**2.2 ChunkSwap**

Met ChunkSwap werd voor het eerst echte vooruitgang geboekt, omdat dit algoritme in tegenstelling tot zijn voorgangers tracht “chunks” van aaneensluitende getallen intact te houden, waardoor er uiteindelijk minder mutaties nodig waren om zowel het genoom van de vlieg om te zetten (klopt dit weet het niet zeker meer). Voor het totaal aantal verplaatste genen en de gemiddelde grootte van swaps bleek dit algoritme echter niet significant beter te scoren dan zijn voorgangers. Bij het 2000 maal draaien van het algoritme op random gegenereerde genomen bleek het over de gehele lijn ook niet beter te scoren dan de eerdere Find and Swap versies. Toch leek kijken naar het behouden van getallen “chunks” een goed pad om in te slaan, omdat dit bij de vlieg wel tot een verlaging van het aantal mutaties had geleid.

**2.3 B-Star en B-Star Elementshrink**

Leuk

**2.4 BeamSearch**

Ook leuk

**3. Resultaten**

**4. Conclusie**