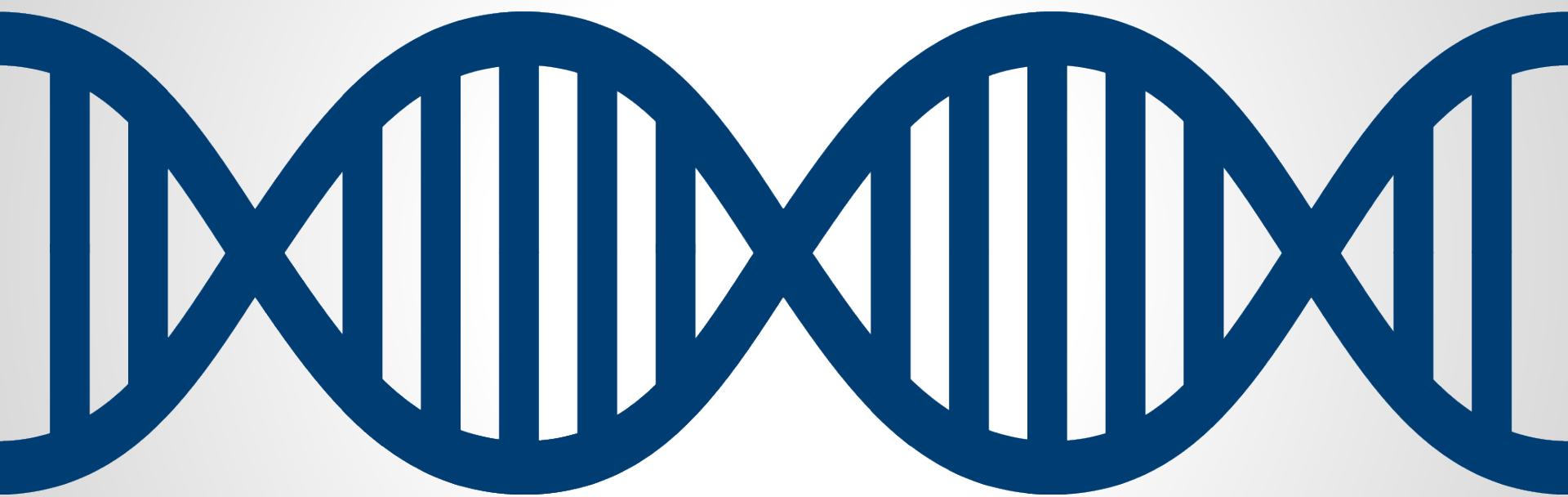


H E L I X



Project door Pieter Coeck, Anthony Hermans,
Stijn Janssens en Evert Heylen

Synopsis

- Concept
- Biologische achtergrond
- Maatschappelijke relevantie
- Functionaliteit
- Links met de cursus
- Performantie
 - Algemeenheid vs Performance (een vb)
- Praktisch
- Eindproduct



Concept

- Bio-Informatica Project
- Patroonherkenning in DNA
 - eerst conventioneel
 - daarna uitbreiding

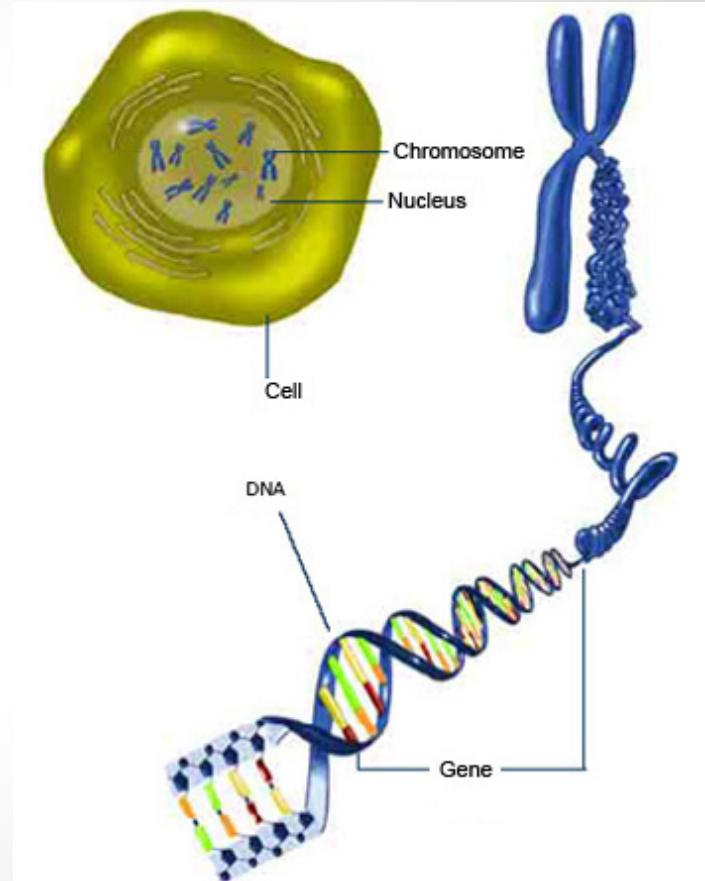


H E L I X

A stylized DNA double helix is shown at the bottom right. Above the helix, the letters H, E, L, I, and X are arranged vertically, corresponding to the segments of the DNA molecule.

Biologische achtergrond

- DNA ligt aan de basis van hoe cellen werken
- de programmeertaal van het leven op aarde

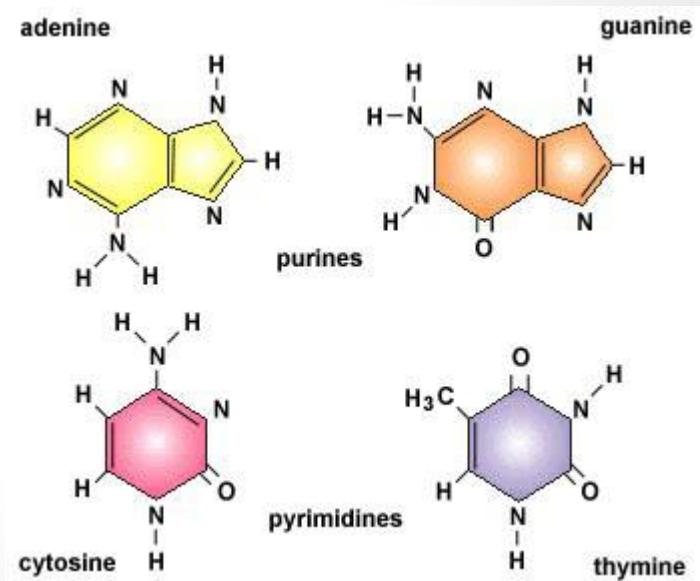


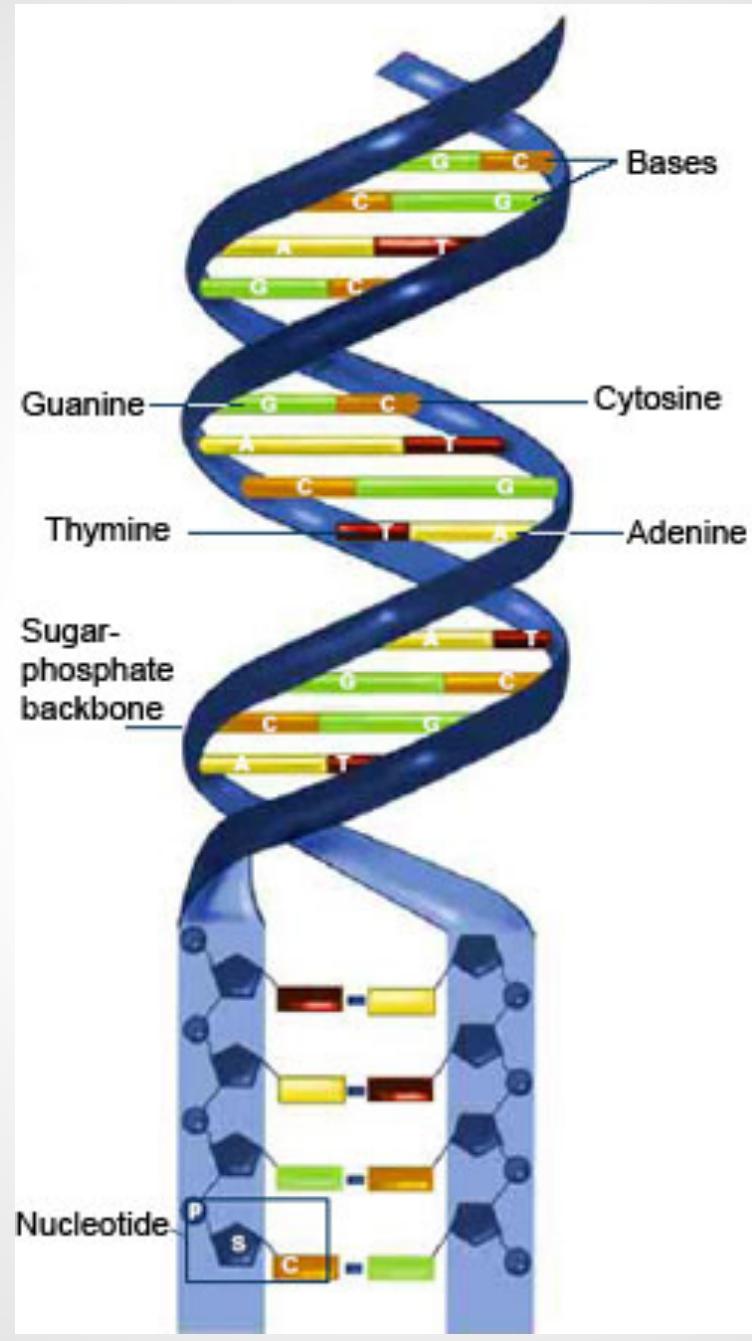
H E L I X

A stylized DNA double helix is shown, with the word "HELIX" integrated into its structure. The letters H, E, L, I, and X are placed along the helix, corresponding to the segments of the DNA molecule.

Biologische achtergrond

- Desoxyribonucleic acid
- Bouwstenen (basen):
 - Adenine
 - Thymine
 - Guanine
 - Cytosine





H E L I X

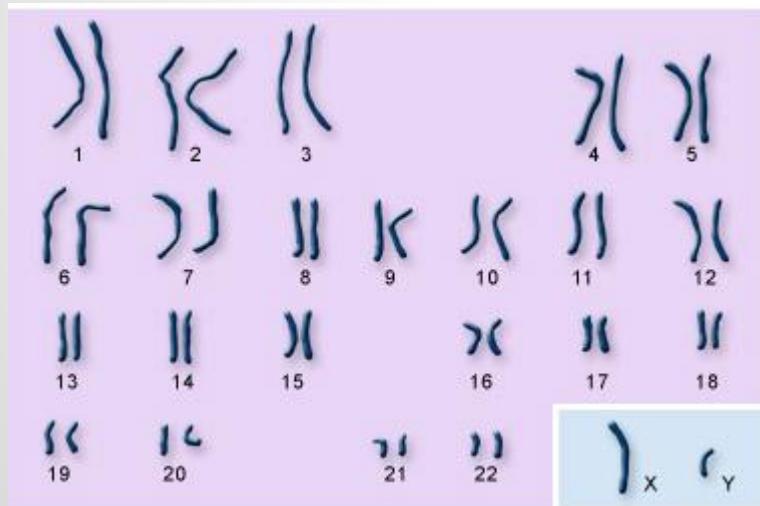
Maatschappelijke relevantie

- Gebruik in biologische sector
 - medicijnen
 - ziektes opsporen
- DNA-matching
 - ouderschap bewijzen



Functionaliteit

- Matchen DNA-patronen
 - DNA-strings == MB's data
- efficiëntie!



chromosoom 9

== 147 MB tekstfile

== 2306580 lijnen



Uitbreidbare functionaliteit

- Meer gradatie in matchen: #errors
- Complementaire string creëren of vergelijken



Links met de cursus

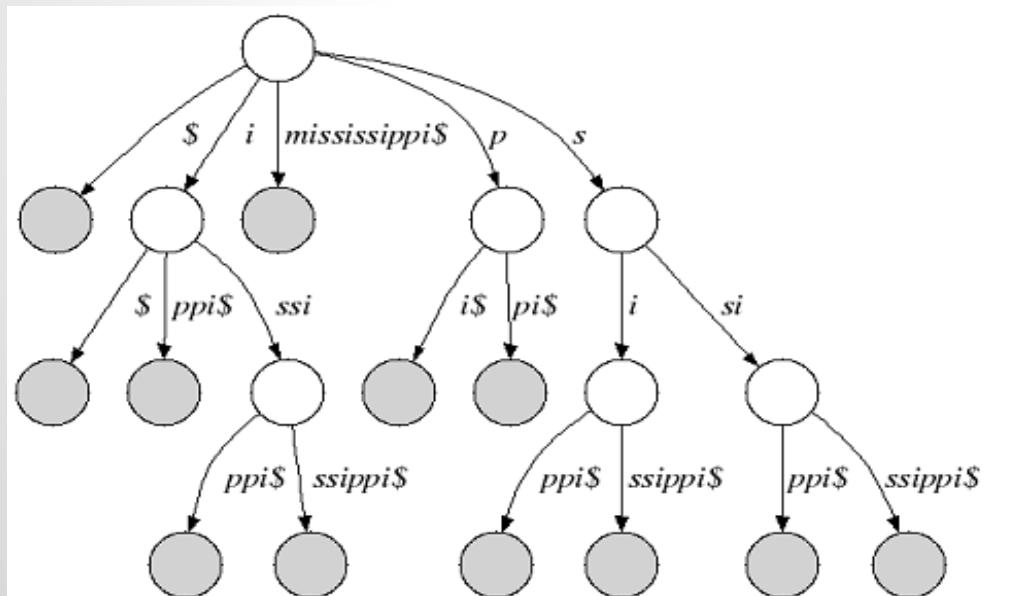
- Performantie → TFA
- Gebruik van NFA's, regex...
- *Theoretische* performantie algoritmen (Big-O)



Extra links met cursus

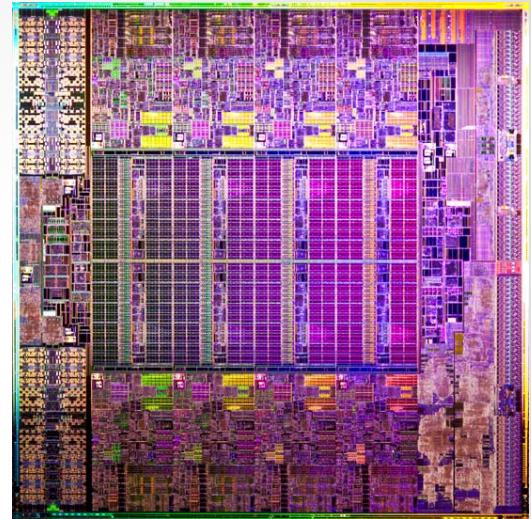
- prefix tree
- suffix tree

→ Geïnspireerd door automaten



Performantie?

- snelheid en geheugenverbruik
- multithreading
 - meerdere cores gebruiken van de processor

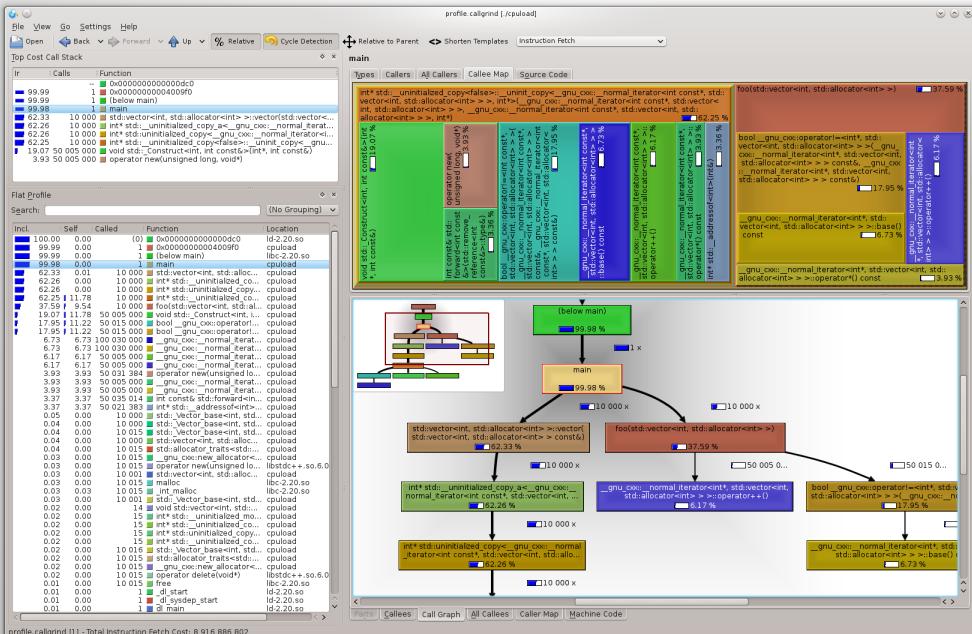


H E L I X

A logo consisting of a stylized DNA double helix. The letters 'H', 'E', 'L', 'I', and 'X' are positioned above the corresponding segments of the helix, spelling out the word 'HELIX'. The helix is rendered in a dark blue color against a white background.

Performantie

- (memory) optimization
→ profiler

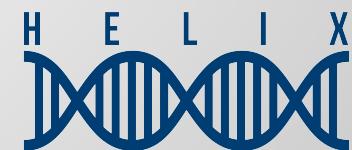


(valgrind, kcachegrind)



Algemeenheid vs performantie (een voorbeeld)

- Algemeen → std::string als type symbool
- Iets specifieker → char
 - DNA-string met x basen → x bytes



Algemeenheid vs performantie (een voorbeeld)

- Algemeen → std::string als type symbool
- Iets specifieker → char
 - DNA-string met x basen → x bytes
- DNA → A, C, T, G
 - 4 mogelijkheden → 2 bits
 - DNA-string met x basen → $x/4$ bytes



Hoe gaan we te werk?

- Make it work
- Make it beautiful
- Make it fast



Praktisch

- C++
- (focus op) Linux
- Extra tools
 - git (en github)
 - GNU Make



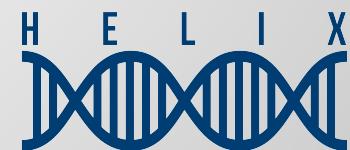
Eindresultaat

- GUI met verschillende opties
 - complement
 - matchen regex
 - makkelijk bekijken resultaten



Conclusie

- Nuttig project
- Gebruikt theorie
- Uitbreidbare functionaliteit
→ meer werkuren indien nodig



Einde

Brent Nederhand™

