



# CS 225

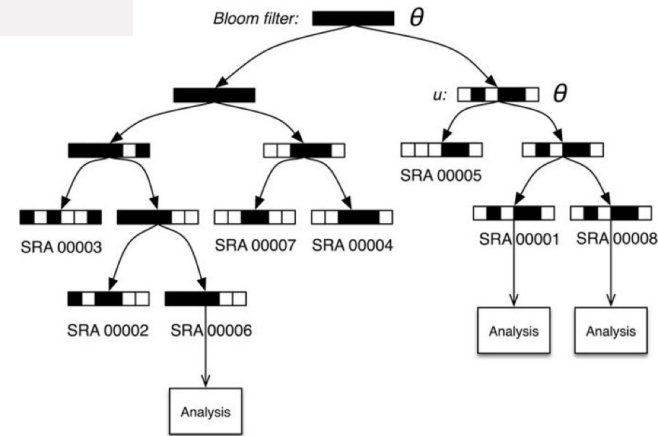
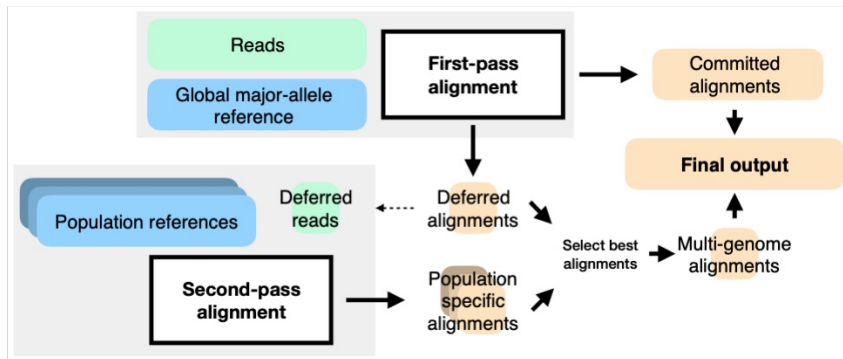
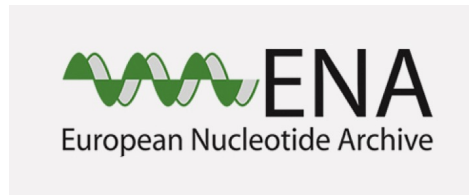
## Data Structures

*January 18 – Introduction*

*G Carl Evans & Brad Solomon*



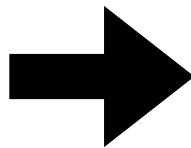
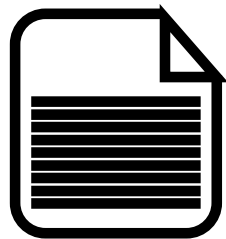
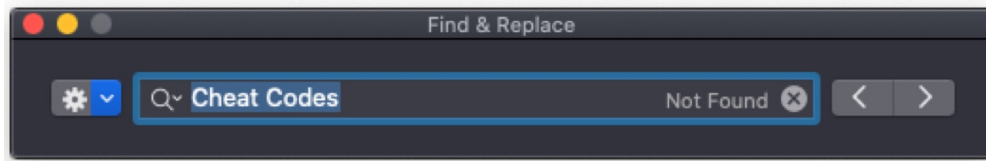
# Brad Solomon



Fast search of thousands of short read sequencing experiments. Brad Solomon and Carl Kingsford. *Nature Biotech* 2016.

Reducing reference bias using multiple population reference genomes. Chen et al. *Genome Biology* 2021

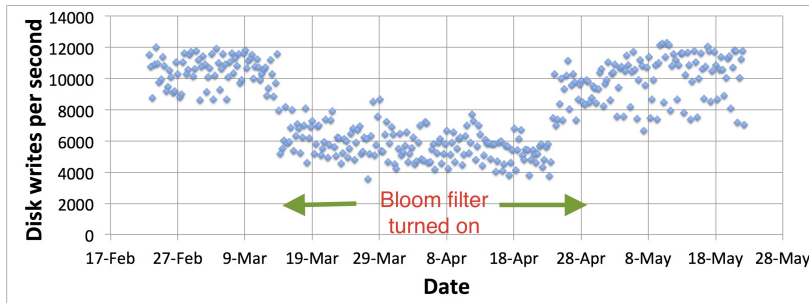
# CS 199-225: String Algorithms and Data Structures



```
GATCACAGGTCTATCACCTATTAACCACTCACGGGAGCTCTCCATGCATTTGGTATTTT
CGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCATGTC
GCAGTATCTGTCTTTGATTCCCTGCCTCATCTATTATTTATCGCACCTACGTTCAATATT
ACAGGCGAACATACTTACTAAAGTGTGTTAATTAATTAATGCTTGTAGGACATAATAATA
ACAATTGAATGCTGCACAGCCACTTTCCACACAGACATCATAACAAAAAATTTCCACCA
AACCCCCCTCCCCGCTTCTGGCCACAGCACTTAAACACATCTCTGCCAAACCCAAAA
ACAAAGAACCCCTAACACCAGCCTAACAGATTTCAAATTTTATCTTTTGGCGGTATGCAC
TTTTAACAGTCACCCCCCACTAACACATTAATTTTCCCCTCCCCTCCACTCCATACTACTAAT
CTCATCAATAACAACCCCGCCATCCTACCCAGCACACACACCCGCTGCTAACCCATA
CCCCGAACCAACCAACCCCAAGACACCCCCACAGTTTATGTAGCTTACCTCCTCAAA
GCAATACACTGACCCGCTCAAACCTCCTGGATTTTGGATCCACCCAGCGCCTTGGCCTAAA
CTAGCCTTTCTATTAGCTCTTAGTAAGATTACACATGCAAGCATCCCCGTTCCAGTGAGT
TCACCTCTAAATCACCACGATCAAAGGAACAAGCATCAAGCAGCAATGCAGCTC
AAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGAATACCTTTAGCAATAA
ACGAAAGTTTAACTAAGCTATACTAACCCCAAGGTTGGTCAATTTCAGCCAGCCACCGC
GGTCACACGATTAACCCAAGTCAATAGAAGTCGGCGTAAAGAGTGTAGATCACCCCC
TCCCAATAAAGCTAAAACCTACCTGAGTTTAAAAAACTCCAGTTGACAAAATAGAC
TACGAAAGTGGCTTTAACATATCTGAACACCAATAGCTAAGACCCAACTGGGATTAGA
TACCCCACTATGCTTAGCCCTAAACCTCAAAGTTAAATCAACAAAACTGCTCGCCAGAA
CACTACGAGCCACAGCTTAAAACCTCAAAGGACTGGCGGTGCTTCATATCCCTCTAGAGG
AGCCTGTTCTGTAATCGATAAACCCCGATCAACTCACCACCTTACAGCCTATATA
CCGCCATCTTCAGCAAACCTGATGAAGGTACAAGTAAAGGCAACACACGTAAG
ACGTTAGGTCAAGGTGTAGCCCATGAGGTGGCAAGAAATGGGCTACATTAACCCAG
AAAACCTACGATAGCCCTTATGAACTTAAGGGTGAAGGTGGATTAGCACTAAG
AGTAGAGTGCTTAGTTGAACAGGGCCCTGAAGCGGTACACACCCGCCGTCACCTC
AAGTACTTCAAAGGACATTTAACTAAAACCCCTACGCATTTATATAGAGGAGT
CGTAACCTCAAACCTCCTGCCTTGGTGATCCACCCGCTTGGCTACCTGCATAAAG
AAGCACCAACTTACACTTAGGAGATTTCAACTTAACTTGACCGCTCTGAGCTAAACCTA
GCCCAAAACCACTCCACCTTACTACCAGACAACCTTAGCCAAACCATTTACCAATAA
```

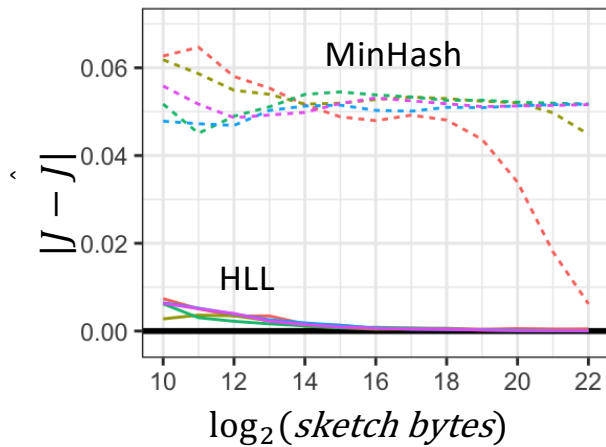


# CS 199-225: String Algorithms and Data Structures

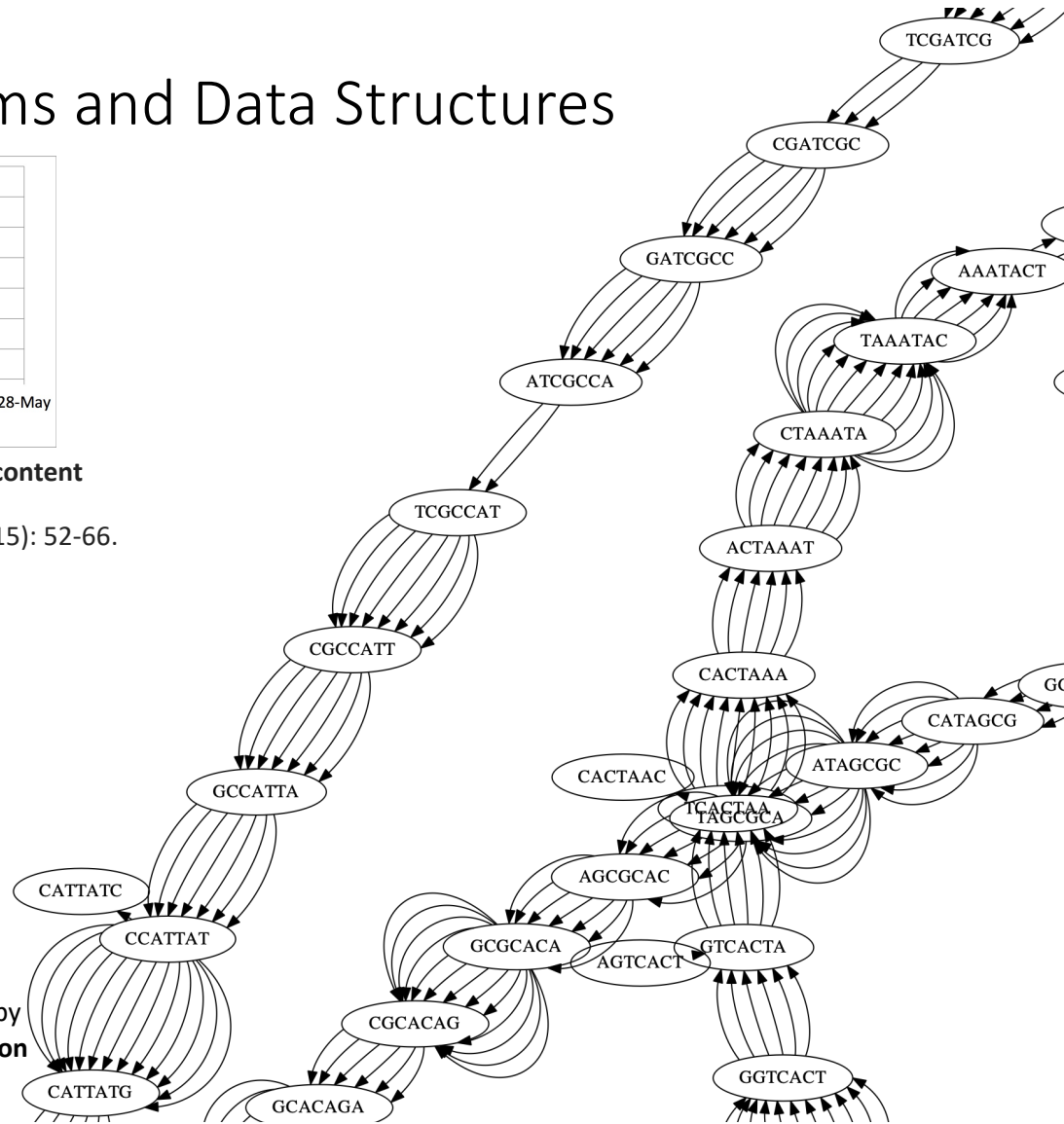


Maggs, Bruce M., and Ramesh K. Sitaraman. **Algorithmic nuggets in content**

**delivery.** *ACM SIGCOMM Computer Communication Review* 45.3 (2015): 52-66.



Ondov B, Starrett G, Sappington A, Kostic A, Koren S, Buck CB, Phillippy AM. **Mash Screen: high-throughput sequence containment estimation for genome discovery.** *Genome Biol* 20, 232 (2019)





# CS 199-225: String Algorithms and Data Structures

Course Schedule: Monday 5:00 — 5:50 PM

1214 Siebel Center for Computer Science

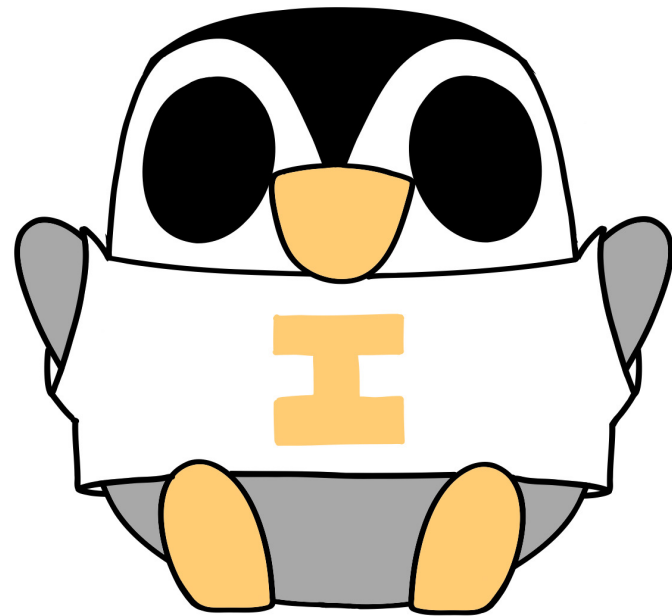
First class: January 30th

<https://courses.engr.illinois.edu/cs225/sp2023/pages/honors.html>

Thierry  
Ramais

---





**CS225 STAFF**



# How to contact us?

- Admin Email

[cs225admin@lists.cs.illinois.edu](mailto:cs225admin@lists.cs.illinois.edu)

- Discord

<https://discord.gg/YuEwhnR>

Do not DM Staff on Discord!



ECE 220

CS 128

CS 225

CS 173

MATH 213





# Everything about CS 225

<https://courses.engr.illinois.edu/cs225/>

## **Information on:**

Staff

Communications

Lab Sections

MPs

Exams

Grading

Academic Integrity



# Grading Notes

Points	Grade	Points	Grade	Points	Grade
[930, $\infty$ )*	A+	[930, $\infty$ )*	A	[900, 930)	A-
[870, 900)	B+	[830, 870)	B	[800, 830)	B-
[770, 800)	C+	[730, 770)	C	[700, 730)	C-
[670, 700)	D+	[630, 670)	D	[600, 630)	D-
		(600, 0]	F		

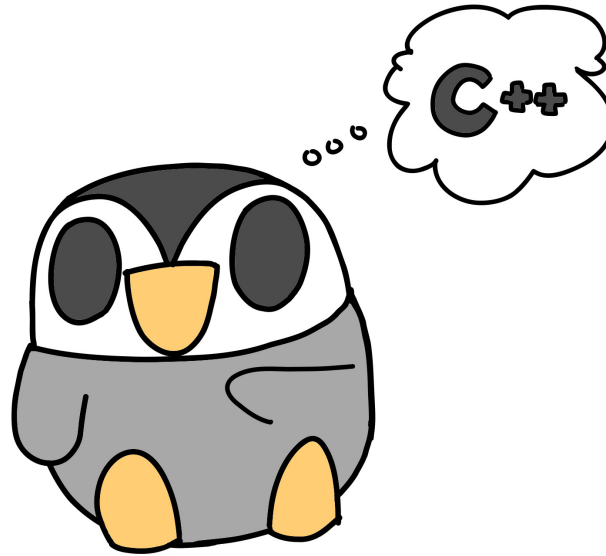
\* To get an A+ you need the score of 930 but you also need to have your final project or for some other work in this course to be seen as exceptional by course staff.



What is this course about?



# What about C++



Material from CS 128

<https://learncpp.online/lessons>



# Problems of the Day (PotDs)



# Open Lab This Week

This week lab are open office hours to help you get your machine setup and can be reached at this link.

