

## Introduction, administrivia

Data Structures and Algorithms for Computational Linguistics III  
ISCL-BA-07

Çağrı Çöltekin  
/ʧaˈɾɪ tʃɔlˈtɛkɪn/  
ccoltekin@fsfs.uni-tuebingen.de

University of Tübingen  
Seminar für Sprachwissenschaft

Winter Semester 2021-2022

version: 202109-2022-03-10

## What is this course about?

- An intermediate-level course on programming
- Algorithms: (good) solutions to programming problems
- Data structures: (efficient) ways to organize/store information

Prerequisites:

- Data Structures and Algorithms for CL I
- Data Structures and Algorithms for CL II

Module: ISCL-BA-07, Advanced Programming

C. Çöltekin, SS | University of Tübingen

Winter Semester 2021-2022 1/19

## What is in this course?

A bird's eye view

Introductory lectures on

- Some fundamental data structures: arrays, queues, stacks, trees, ...
- Some fundamental algorithms: searching, sorting, pattern matching, graph algorithms
- Analysis of algorithms
- Finite state automata
- Parsing

C. Çöltekin, SS | University of Tübingen

Winter Semester 2021-2022 2/19

## Why study algorithms?

- It is one of the fundamental topics in computer science: an algorithm is the way you instruct a computer to do things
- Knowing a clever, efficient solution to one problem helps designing good solutions for other, related problems
- Learning basic algorithmic techniques makes you a better programmer
- Designing good algorithms is an intellectual challenge
- The most popular interview questions for programming jobs are about algorithms

C. Çöltekin, SS | University of Tübingen

Winter Semester 2021-2022 3/19

## Course overview

- Lectures (all online):
  - Monday 14:15-15:45 (lecture)
  - Wednesday 12:15-14:45 (lecture)
  - Friday 14:15-17:45 (lab)
- Tutors:
  - Siena Biales
  - Klara Lennermann
- Public course website: <https://dsac13-2021.github.io/>
- Moodle:  
<https://moodle.zdv.uni-tuebingen.de/course/view.php?id=2057>
- GitHub: <https://github.com/dsac13-2021/dsac13>

C. Çöltekin, SS | University of Tübingen

Winter Semester 2021-2022 4/19

## Literature

- *Data Structures and Algorithms in Python*. Goodrich, Tamassia, and Goldwasser (2013)
  - Available through university library (online version):  
<https://bookcentral.proquest.com/lib/unitueb/detail.action?docID=4946360>
  - Website of the book contains source code, hints, examples:  
<http://bca.wiley.com/bw-bca/Books?action=index&csId=8029&itemId=1118290275>
- *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Jurafsky and Martin (2009)
  - Draft chapters of 3rd edition is available at  
<https://web.stanford.edu/~jurafsky/slp3/>
- Course notes will be provided for some topics

C. Çöltekin, SS | University of Tübingen

Winter Semester 2021-2022 5/19

## Coursework and evaluation

- Reading material for most lectures
- 6 programming assignments (approximately every two weeks)
  - The best 5 assignments count (as 60% of your total grade)
  - NEW: Each assignment is paired with a short quiz
- Final (written) exam (40%)
- Attendance is not required, but you are unlikely to pass without regular attendance

C. Çöltekin, SS | University of Tübingen

Winter Semester 2021-2022 6/19

## Assignments

- Assignments in Python
- Only online submissions through GitHub
- Up to one week late = half the points, more than one week late = 0 points
- Solutions will be discussed after the late-assignment deadline
- The assignments can be done in pairs (strongly recommended – knowing your classmates, and learning from them, is an important part of the university experience/education)
- This means **working together on all exercises**, not sharing and parts of an assignment and working on them independently
- We will have a match-making mechanism
- See course page for more information

C. Çöltekin, SS | University of Tübingen

Winter Semester 2021-2022 7/19

## Topics at a glance

- A recap of what you should already know: arrays, lists, maps, queues, stacks, iteration, recursion, binary search, ...
- Algorithmic analysis
- Common algorithmic patterns: brute force, greedy, divide and conquer, dynamic programming, ...
- Sorting
- Trees
- Priority queues, heaps
- Hashing
- Graphs, graph algorithms
- Pattern matching
- Tries
- Regular expressions and finite state automata
- Finite state transducers
- Parsing

C. Çöltekin, SS | University of Tübingen

Winter Semester 2021-2022 8/19

## Final remarks

- Please do not be shy, ask your questions during the lectures
- Please take the assignments seriously, learning programming requires practice
- Next:
  - a recap of basic data structures and algorithms
  - a Python tutorial
- Time for your questions.

C. Çöltekin, SS | University of Tübingen

Winter Semester 2021-2022 9/19

## Acknowledgments, credits, references

- Some of the slides are based on the previous year's course by Corina Dima.

Goodrich, Michael T., Roberto Tamassia, and Michael H. Goldwasser (2013). *Data Structures and Algorithms in Python*. John Wiley & Sons, Incorporated. [sasc: 9781118476734](https://doi.org/10.1002/9781118476734).

Jurafsky, Daniel and James H. Martin (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, second edition. Pearson Prentice Hall. [sasc: 978-0-13-504196-3](https://doi.org/10.1002/9780135041963).

C. Çöltekin, SS | University of Tübingen

Winter Semester 2021-2022 A.1