One Data Science Program Week 1

Introduction to programming, languages, and tools

Author: Katarzyna Dziopa

Teaching someone how to make a sandwich

Teaching someone how to make a sandwich

- Take two slices of bread
- Place a slice of ham over one slice of bread
- Place a slice of cheese over the ham
- Place a second slice of bread on top of the cheese

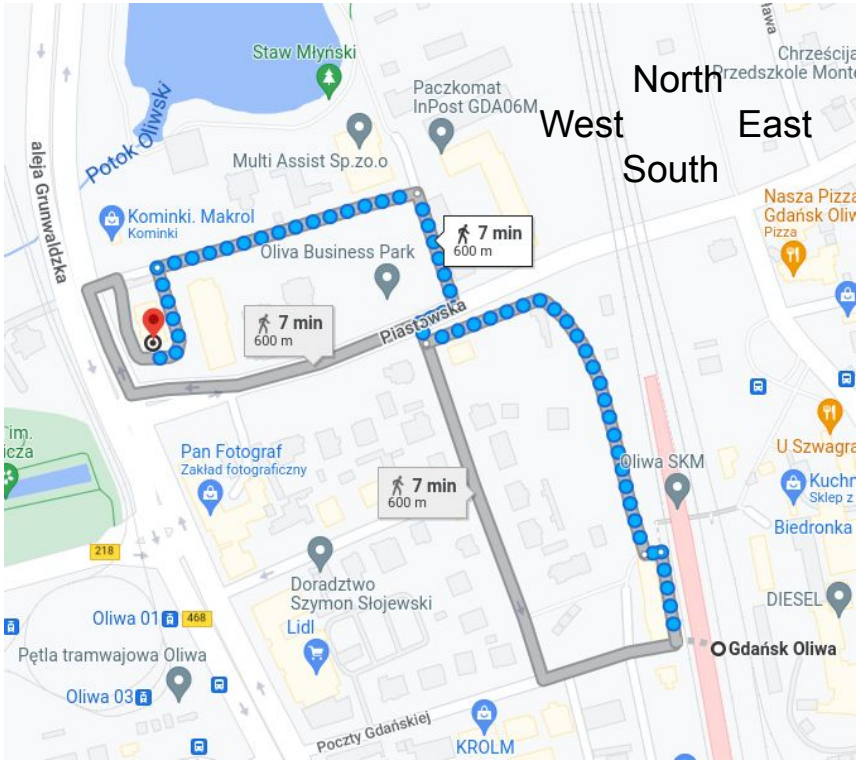
Changing a light bulb

Changing a light bulb

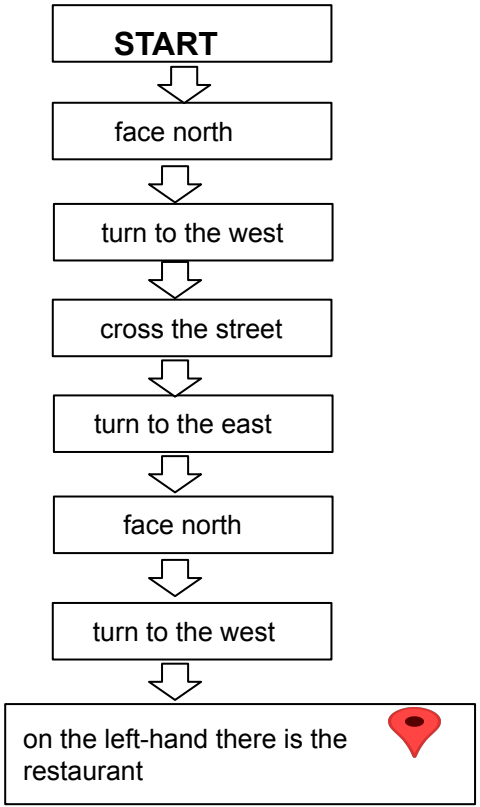
- If you do not have a new lightbulb, buy a new one
- Turn off the power
- Repeat: Twisting the lightbulb counter-clockwise; until you remove it
- Throw out the burnt lightbulb
- Place the new lightbulb into the socket
- Repeat: Turning the lightbulb clockwise; until it is tight
- Turn the power on

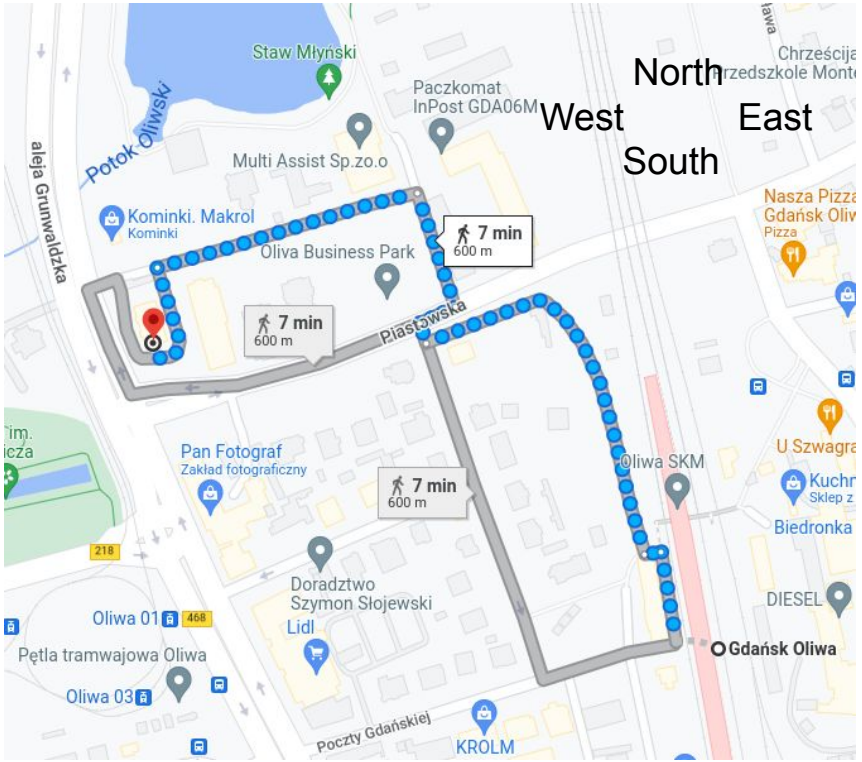
Representation of algorithms

Try to give directions to a person who is currently at the Gdansk Oliwa train station and wants to get to the MacDonald's restaurant (red pin).



Representations of algorithms - a flowchart





What is programming?

Programming is a way to "instruct the computer to perform various tasks".

1. "Instruct the computer": providing the computer a set of instructions that are written in a language that the computer can understand. Examples: adding 2 numbers
2. "Perform various tasks": for example adding two number or rounding off a number (simple task) or more complex tasks such as finding a shortest path between two cities.

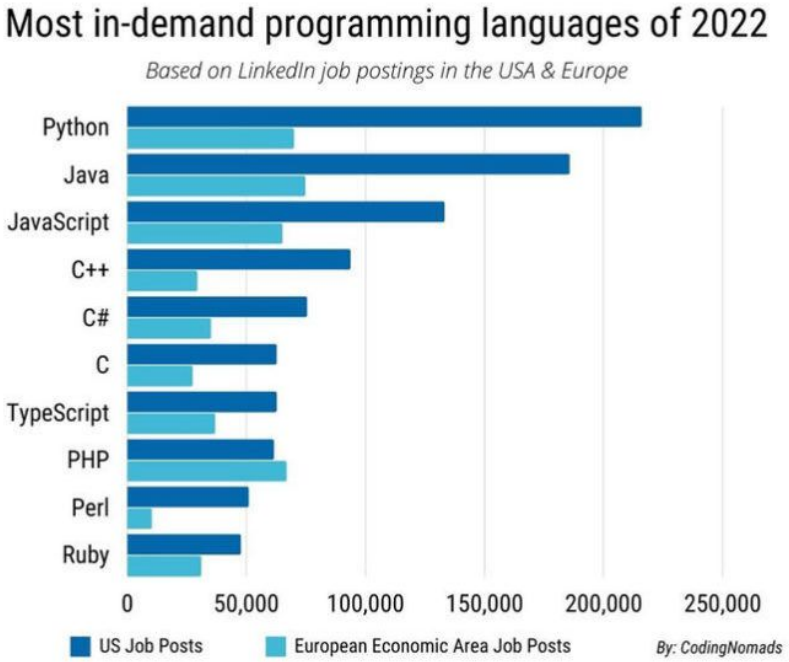
What is a program?

- A program is a set of instructions that a computer uses to perform a specific function.
- Contains a list of ingredients and a list of directions that tell the computer how to execute a specific task.

```
> myString <- "Hello, World!"
> print ( myString)
[1] "Hello, World!"
```

Types of programming languages and their popularity

- Python - tasks range from advanced web application and artificial intelligence to simple scripting.
- Java - general-purpose language, it is used to develop desktop, mobile apps, big data processing, and so on.



R programming language

Used in:

1. Statistical computing
2. Data analysis
3. Scientific Research

Popular language used by statisticians, data analysts, researchers and marketers to retrieve, clean, analyze, visualize and present data.

Major strengths: comprehensive set of high-quality, specialized open-source packages for solving/calculating complex problems in various domains e.g. bioinformatics, healtcare.