# OOSD Content

# Week 1 – Intro, Tools, Data, Variables, print and input

## Introduction

### Software development

#### What is software?

The word software denotes instructions for a machine, organised into helping humans use the machine. The physical machine (the hardware) can be anything, but to be controlled by software it must include a microprocessor.

#### Why not just have the machine and use it? Why software?

Every machine's functionality is embodied in its hardware, but some machines are more limited than others. An old-fashioned hairdryer blows air at two or three heat settings and that is all. Controls can be implemented using fairly simple electrical circuits.   
  
A programmable washing machine is also limited in what it can do but there is scope for variation. To allow users to configure a washing machine, an electronic circuit (pure hardware) could be used, but it this would have to be more complicated than the simple circuits of the hairdryer. The alternative is to have a processor and software. The processor consists of circuits that perform a limited number of primitive operations and software is a recipe for how these primitive operations should be combined to achieve something. For all but very simple functionality, this is **cheaper to produce than a specialised electronic circuit**.   
  
However, the true value of this setup (processor + software) lies in another property it has. While a washing machine can only wash clothes (admittedly in several different ways), less specialised hardware and a more complex user interface can provide user-defined functionality through the **theoretically infinite combinatorial power of sofware as a recipe**. This is what computers and smartphones do. The versatility of their user interfaces allows for ever new functionality to be invented and their internal programmabiliy supports its implementation.

#### What is software physically?

A unit of software usable for some purpose is called a software program, application, app or product. One of these is like the contents of a book, e.g. Oliver Twist by Charles Dickens. When we call out that name we generally do not mean any particular paper or electronic copy of the story, we mean the words that Dickens wrote and that we read. It is like that with software, which, although contained in a file on disk or in the memory of a device, cannot be equated with where it is stored but consists of abstract instructional content, which can be copied and stored in multiple places at once. **It is not physical.**

#### How does software development fit into all this?

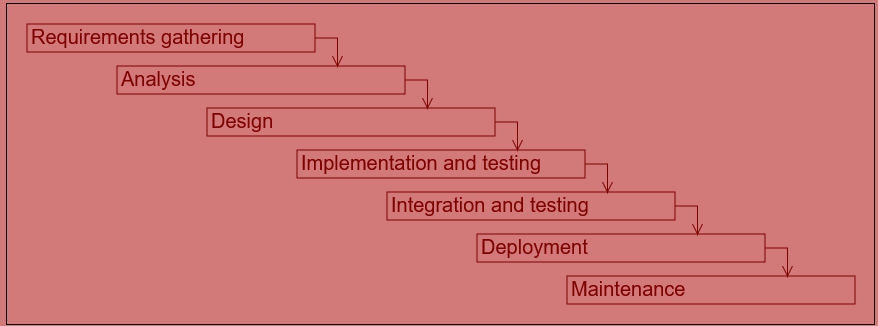
Putting together a program does not require listing of primitive instructions understood by a processor. These are packed into higher level constructs and ready-made software chunks of a programming language, which is less granular and more user-friendly for humans. Different programming languages present the underlying processor capabilities in different ways, depending on their purpose but also simply on the preferences of their designers.

What is the difference between software development, programming and coding?

Different people and even different dictionaries will give you different definitions, with the distinction between the terms varying from huge to none. However, when a distinction is made, software development is thought of as broader than programming, which in turn is treated as broader than coding. So, coding is simply writing the instructions (or code) in the language; programming involves a bit more e.g. design and testing but is still performed by one person; software development is the entire process that starts with an idea and ends in a product and could involve many people and organisations. Software development can also be used as a term for the entire industry involved in producing software.

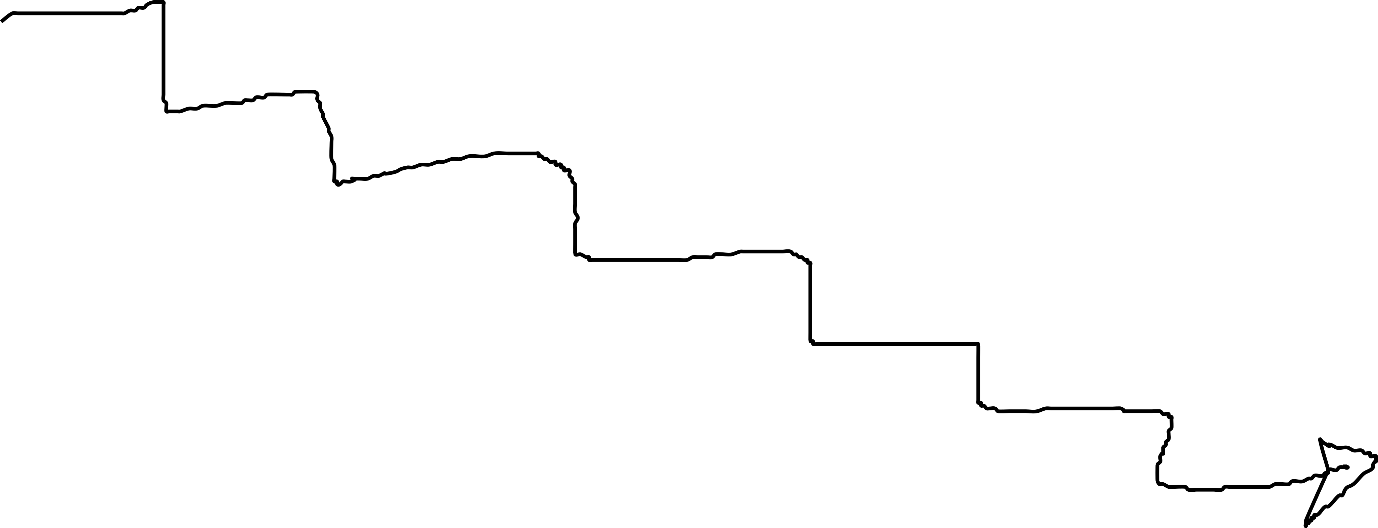
### Methodologies

Developing software is a complex undertaking that includes several different types of activity within its lifecycle (the software development lifecycle - SDLC):



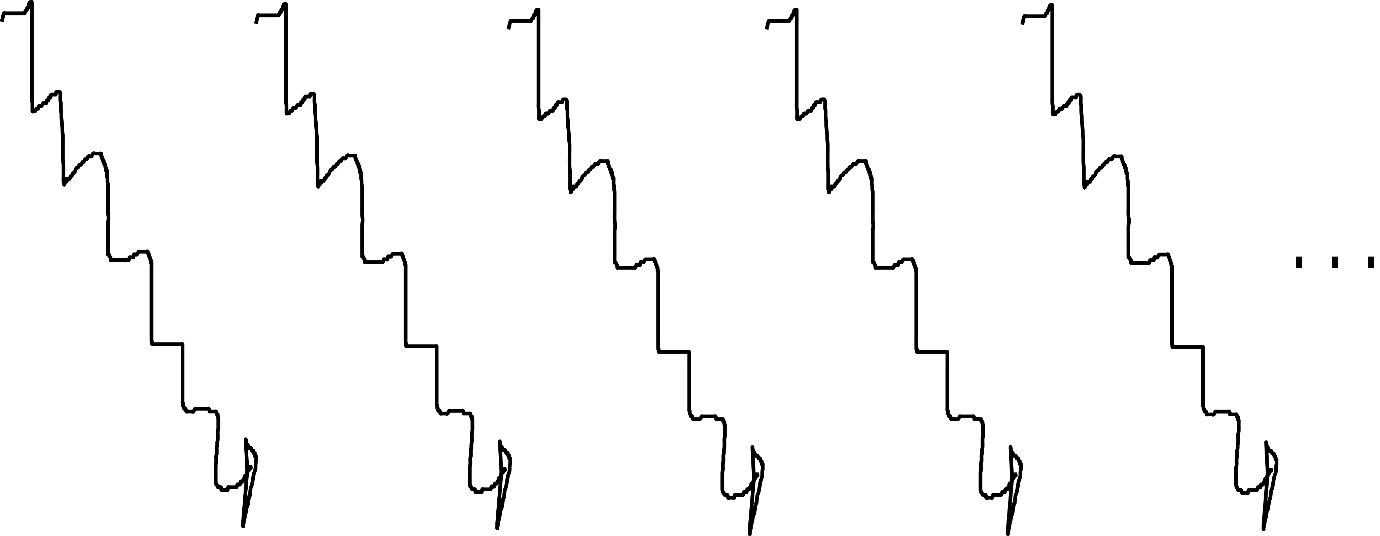
Depending on the context, the way these activities are interconnected and carried out as well as the degree of formality employed will differ greatly. However a **methodology** will generally be adhered to in one form or another.

The two main methodology styles are:



#### waterfall

* + projects are large
  + each activity is a single phase in a large project
  + phases last weeks or months
  + a project generally corresponds to a single release



#### agile (e.g. Scrum, XP)

* + projects consist of many small iterations (1-2 weeks)
  + constant updating of requirements based on feedback
  + constant communication between teams and stakeholders (internal and external)
  + continuous delivery and even deployment
  + goes hand-in-hand with DevOps (also promotes connectedness of activities and roles but is centered around *delivery* as opposed to *development* in agile, see [this](https://www.guru99.com/agile-vs-devops.html))

But there are many more that are either variations on these two themes or sit somewhere between the two, for example:

#### Rapid Application Development (RAD)

* + less focus on design and documentation
  + prototyping
  + multiple iterations

#### 'Sashimi' model - with overlapping phases

### Methods and tools

There are methods and tools for every activity in the software development cycle.

#### requirements gathering and analysis,

* using tools such as Eclipse-based Rational tools or VisualParadigm
  + use cases e.g. as prescribed by Rational Unified Process (RUP) framework
  + user stories in agile
  + source-stimulus-environment-artefact-response-measure for non-functional requirements
  + analysis-level class diagrams

#### design, using tools such as Eclipse-based Rational tools or VisualParadigm

* + component diagrams (describing high-level structure)
  + design-level class diagrams (describing structure)
  + sequence diagrams (describing behaviour)

#### implementation

* + based around programming in languages such as Python, Java, C++
  + supported by many different types of tools:
    - syntax-aware text editors e.g. Visual Studio Code, Notepad++
    - interpreters and compilers for
    - debuggers
    - program profilers and validators e.g. Valgrind
    - version control systems e.g. git implemented by GitHub, Bitbucket etc.
    - integrated development environments (IDEs) e.g. Eclipse, NetBeans, PyCharm; providing integrated access to all of the above listed tools, typically
      * text editor
      * compiler/interpreter
      * debugger
      * profiler
      * version control

#### integration, deployment, maintenance

* + tools for the automation of configuration and deployment such as Puppet, Chef, Ansible

#### planning and management (orthogonal to all of the above) also require tools for

* + defect tracking
  + project management (traditional)
  + agile tools, which include both of the above

### The substance of software and programming languages

Software has two aspects, both equally important:

* **information**, represented in a program by **data structures**
* **behaviour**, represented in a program by **algorithms**

A **programming language** consists of vocabulary, syntax and rules for representing *information* and *behaviour*.

Here are some broad classifications of programming languages:

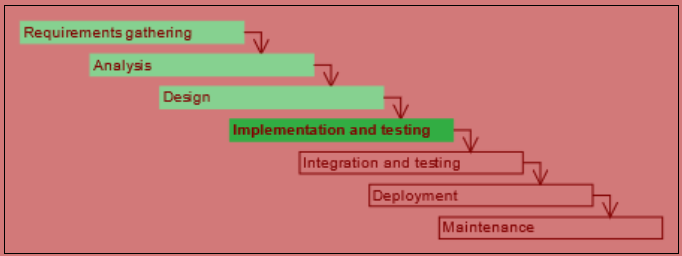
* imperative (Java) vs. declarative (SQL), by focus, on controlling behaviour vs. information
* object-oriented (C++) vs. procedural (C), by how behaviour and information are organised
* high-level (e.g. C++) vs. low-level (e.g. assembly), by closeness to the processor
* compiled (e.g. C++) vs. interpreted (e.g. Python), by mode of translation, in advance as a whole vs. during execution
* special-purpose (e.g. R for data analysis) vs. general-purpose (e.g. Python)

### What does the Object Oriented Software Development module cover?

* This module focuses primarily on
  + **implementation and testing**

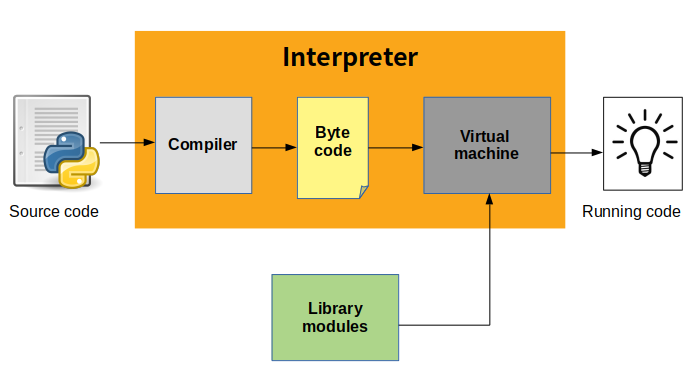
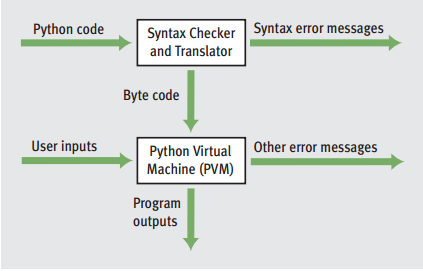
but also provides an introduction to

* + requirements gathering
  + analysis
  + design



* In the module we employ **problem-based learning (PBL)**, which engages the student in learning through solving open-ended problems. This is particularly suited to software development as it mirrors the real-life (workplace and advanced study) circumstances of software development application.
* The module is taught through **Python**, a high-level interpreted general-purpose programming language.

### Python and tools

* The Python programming language can be used
  + interactively on the command line for simple work
  + in one or more interacting scripts that can be kicked off to execute as a unit
* Python is an interpreted language, requiring no explicit compilation step on the part of the developer (traditionally this type of language would have been used for scripting). Python interpreter  
  (source: indianpythonista.wordpress.com) 
* A program in Python (or any other language) is useful only if it interacts with its environment by consuming and providing information Python program inputs and outputs  
  (source: python-nitol.blogspot.com) 

## To Begin

### Functions

Functions in Python are **chunks of functionality** that can be called to do something. The list of Python functions available with the interpreter can be found [**here**](https://docs.python.org/3/library/functions.html). We will be looking at some of these later.

#### Functions can be

* part of the standard library that is installed with the Python interpreter
* part of **modules** or **packages** that can be added to the installation
* defined and used in a newly developed Python program

For example, to exit the Python interactive interpreter, function **exit()** or **quit()** can be used.

##### Example [CS-1] - Functions for exiting the interactive interpreter (only available in interactive mode)

exit()

e# or

quit()

#### More about functions:

* Function calls are denoted with **parentheses following the function name**. The parentheses may be empty or contain arguments
* Functions **can take one or more arguments**, i.e. pieces of information that they use when carrying out whatever they are designed to do. The call to the function print() in code snippet [[CS-2]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_1_begin.html#py-func-print) is made with argument 3. The function print() prints the argument passed to it to the console.

##### Example [CS-2] - A function with an argument (run interactively)

print(3)

* Functions **can return values** and this is often their purpose (to create a value by calculating, converting or performing some other action that results in some data).

##### Example [CS-3] - A function that returns a value (run interactively)

round(4.3)

round(4.123456454323455675544, 4)

Functions will be introduced and used throughout the module. The syllabus, however, does not cover the definition of own functions.

### The Help Function

While Python documentation is available on the [**Python documents page**](https://www.python.org/doc/), it is also possible to get information on individual items programmatically, using the function help().

##### Example [CS-4] - Using the help() function (run interactively)

help(str)

help(True)

help(help)

### Objects and methods

**Roughly anything of substance,** i.e. **anything that has a value and a type**, is an **object** in Python. Thus a number or a piece of text is an object, as are more complicated structures with many pieces of individual information.

A function can be associated with a particular type of object and in that case it is called a **method**. Methods are called with the syntax

**<object name or value>.<method\_name>()**

The example in code snippet [[CS-5]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_1_begin.html#py-method-example) is that of the method upper called on an object containing the text "hello". Try it out.

##### Example [CS-5] - Example of a method (run interactively)

**"hello".upper()**

### Comments

# This line will be ignored by the interpeter

exit() # exit() runs but his part of the line is ignored

### Imports

The core functionality of Python is available in the interpreter by virtue of Python having been installed.

Some functionality will have been installed but needs to be 'brought in' with a module import.

##### Example [CS-7] - Module import (run interactively)

**# generate random number between 1 and 10**

**random.randint(1, 10) # error!**

# the random module must be imported

import random

random.randint(1, 10)

Some more specialised packages of functionality will not be installed with Python (this varies between operating systems) and in that case packages will need to be installed before contained modules can be successfully imported.

## Data Types and Operators

### Data Types and Operators

Python defines many data types. In this module we will be using only some of them, namely, boolean, integer, float, string, list and range.

The function [**type()**](https://docs.python.org/3/library/functions.html#type) can be used to check the type of a piece of data

#### Boolean datatype

* A piece of data of type *boolean* carries a *truth value*
* Possible values:

True

False

Note that True and False are written with capitalised first letter and represent *keywords* in Python.

#### Logical Operators

* Operate over *boolean* values
* List (exhaustive):

**and**

binary operation, both operands must be true for the result to be true

**or**

binary operation, at least one of the two operands must be true for the result to be true

**not**

unary operation, result is the opposite of the operand

##### Example [CS-1] - Logical operators (run interactively)

*type(True)*

*type(False)*

*True and True*

*True and False*

*False or False*

*False or True*

*not True*

*not False*

#### Integers

* A piece of data of type *integer* represents a **whole number**
* Any whole number (e.g. 0, 22, -12345) can be represented as an integer in Python

Python 2 had limits as to the integer values that it could represent. Python 3 theoretically can represent any whole number, with limits imposed only by the memory of the computer it runs on.

##### Example [CS-2] - Very large integers (run interactively)

***type(123)***

***10\*\*100 # on my computer this instantly returns the number (called a googol)***

***10\*\*1000000 # on my computer this returns the number but after about 10 seconds***

#### Floats

* A piece of data of type *float* represents **real number**
* Because of the way that floats are represented in most hardware (using double-precision binary fractions)
  + up to *15 significant digits* can be represented

##### Example [CS-3] - Float rounding (run interactively)

**type(4)**

**type(4.0)**

**# the value is implicitly rounded before the expression is written out**

**0.12345678901234567890**

**# rounding to 2 digits**

**round(0.12345678901234567890, 2)**

**# rounding to 10 digits**

**round(0.12345678901234567890, 10)**

**# rounding to 18 digits - some digits are not representable, implicit rounding**

**round(0.12345678901234567890, 18)**

In the code snippet above we have used a new function, [**round()**](https://docs.python.org/3/library/functions.html#round).

* + comparisons of floats may not yield the correct answer (see examples in comparison operator section below)

#### Arithmetic operators

* Operate over *integer* and *float* values
* List (non-exhaustive, visit [**the Python Library Reference**](https://docs.python.org/3/library/stdtypes.html#numeric-types-int-float-complex) for a full list):

+

addition

-

subtraction

\*

multiplication

/

division (yields a float result)

//

integer division (yields an integer)

%

modulo (yields the remainder)

\*\*

power

##### Example [CS-4] - Arithmetic operators (run interactively)

3 + 4

7 / 2

7 // 2

2 \*\* 3

#### Comparison operators

* Comparison operators include equality and relational operators.
* Operate over almost any types where an interpretation is possible.
* List (non-exhaustive, visit [**the Python Library Reference**](https://docs.python.org/3/library/stdtypes.html#comparisons) for a full list):

<

less than

>

greater than

<=

less than or equal to

>=

greater than or equal to

==

equal to

!=

not equal to

With the number types, these operators generally work as expected.

Comparison operators (run interactively) [CS-5]

3 < 4

3 == 3

4.0 != 5.5

However, there are some effects stemming from the way floats are internally represented (as base-2 fractions, with a limited number of bits) that we need to watch out for. For most real number values, their internal representation will be slightly 'off'. This means that when adding two real numbers there is no guarantee that the sum of representations will be equal to the representation of the sum. This effect will only appear outside of the range of 'guaranteed' consistent significant digits.

Operations with floats to watch out for (run interactively) [CS-6]

# the following gives an unexpected result

0.2 + 0.1 == 0.3

# because the value is not what we expected either

0.2 + 0.1

# always round before using the equal operator on floats

round(0.2 + 0.1, 15) == round(0.3, 15)

#### Strings

* In Python, a string is an ordered sequence of characters encoded using UTF-8 and representing some **textual data**.
* The string type is called str.
* str is a **sequence** type, along with list, tuple and range.
* A character on its own is not a 'thing' in Python. A character is represented as a string of length 1.
* A string can be of any length, including 0. Getting the length of a string (run interactively) [CS-7]
* len("Bob")

len("construction")

* String values are denoted using quotes (single or double), with the characters of the string placed between them.
* Characters such as new lines and tabs can be included in a string. These must represented with the appropriate **escape sequence**: '\n' for a new line and '\t' for a tab. Other escape sequences that we will use a lot are those for representing quotes within quotes. For reference, you might like to have a look at a [**complete list of escape sequences**](https://docs.python.org/3/reference/lexical_analysis.html#string-and-bytes-literals) (you need to scroll down a bit) Using escape characters (run interactively) [CS-8]
* print("He looked around and said \"Interesting\"")
* print("one line\nanother line\thop\thop again")

Notice that the escape sequences for new line and tab remain intact when the string is displayed. A string needs to be written to a file or screen for these characters to be rendered. We will see them rendered in section 'Printing'.

#### String concatenation

There are many useful string operations in Python, but for the moment we will look at concatenation using the +. We can concatenate two or more different strings or 'multiply' the same string.

String concatenation with + and \* (run interactively) [CS-9]

"hello" + " " + "there"

"bye bye" + ' ' + 'here'

"blah " \* 3

### Casting between types

Python provides functions for converting pieces of data from one type to another.

int()

converts any type to an integer (if possible)

float()

converts any type to a float (if possible)

str()

converts any type to a string

bool()

converts any type to a bool

Here is some general information on what to expect when casting, but when in doubt check the documentation or try it!

Almost any piece of data can be converted to a boolean:

for numeric data types, a value of zero is converted to False and everything else to True

for strings, an empty string is converted to False and any other string to True

With the flexibility of text, strings are also an easy type to convert to from anything else.

Booleans are easily converted to numeric types, but strings can be converted only if they express a number through text.

Casting (run interactively) [CS-10]

bool(34)

bool(0)

bool("")

bool("abc")

int(4.3)

int("4.3")

int("4.3abc")

float("3")

float(True)

str(2345)

str(True)

### The value and type of expressions

**An expression** in Python is an executable construct that when executed results in **a value with a type** i.e. can be evaluated. Expressions can be used as pieces of data, in all the ways that literal values, such as 5 or "Some text", can be used.

**Example 1**

The interactive interpreter console evaluates any expression passed to it and displays the value obtained (in batch mode, this automatic displaying of values does not occur and they must be explicitly shown if needed, e.g. with print()).

Expression value and type (run interactively) [CS-11]

# an expression's value will be displayed

3 + 4

# display the expression's type

type(3 + 4)

**Example 2**

A function call is also an expression and it evaluates to the returned value. The casting functions are a clear example of this. For example, int("34") returns the integer 34 and thus evaluates to 34. If we make this call inside the interactive interpreter, 34 is printed as this is the value of the expression.

##### Example [CS-12] - Checking the types of casting function calls (run interactively)

type(bool("abc"))

type(int(4.3))

type(float("3"))

type(str(True))

**Example 3**

The function call print("a") is not really an expression, as it does not return a value and thus does not evaluate to anything. However, if we try to treat it as an expression, checking its value and type, we do not get an error but the value None and type NoneType are returned. Functions that do not return a value could thus be thought of as quasi-expressions. This differentiates them from statements, the other type of executable language construct, which represent *actions* rather than *things* and cannot be evaluated. Statements will be treated later in this module.

##### Example [CS-13] - Value and type of print()

# the function call

print("a")

# force display of function call expression value

print(print("a"))

# display the function call expression type

type(print("a"))

### Printing

The print() function writes to the console. We have encountered it already but here we have a closer look at it through examples.

* print() can take any number of arguments of different types, separated by commas The print() function - basic use [CS-14]
* print(345)
* print("Hello there")
* print(True)
* print("Hello,", "Bob")

print('I am', 100, 'years old')

In the output of multi-argument calls, a whitespace character separates the printed arguments.

* When printing, escape characters for new lines and tabs ('\n' and '\t') are rendered ([**list of escape sequences**](https://docs.python.org/3/reference/lexical_analysis.html#string-and-bytes-literals), you need to scroll down to see the table). Using escape characters (run interactively) [CS-15]

print("one line\nanother line\thop\thop again")

* By default, print() ends the output with a new line. This can be changed with the use of a **named argument**, end (see section below, 'Function arguments revisited"). Any string can be specified as the terminating string of the print output. Using the end argument (run interactively) [CS-16]
* print("This is followed by a new line")
* print("This is followed by a full stop", end=".")

print("This is followed by nonsense", end="\tas;dlkfj;alkjksjdflkdjflkj")

## Variables

So far we have learnt about data types and how pieces of data can be manipulated, individually or in groups. However, the functionality that we have used for this has all been Python built-in functionality and the instances of data manipulation carried out have been tied to particular fixed data values or **literals**. We have more or less used Python as a sophisticated kind of calculator.

To create some reusable functionality, in most cases we need to be able to manipulate **placeholders**, rather than only known data values.

##### In class exercise!! >>> [CE-A]

Can you think of two cases where a useful program can be defined without placeholders i.e. work without inputs from the outside world?

There are broadly two types of placeholders in programs:

* **function parameters** (the pieces of data that a function says it accepts), which are used as placeholders for defining what the function does and are replaced by arguments (actual pieces of data) when the function is called
* **variables**, which are, within a context, standalone placeholders for pieces of data

Writing programs using only functions, with function parameters the only kind of placeholder, is called *functional programming* and can be pursued in Python (although it is not optimised for it), but this is not what we are going to do (especially since funcitions are not even on the syllabus!). So let's look at variables.

Python has a very simple approach to variables. They are **named** and **given a value** in one expression, which can be placed anywhere in the code.

**Variable naming and value assignment**

* a variable name
  + can include only alphanumeric characters and underscores ('\_')
  + must not start with a number
  + must not be a [**Python keyword**](https://docs.python.org/3/reference/lexical_analysis.html#keywords)
  + using any built-in function names (e.g. str or len) as variable names would make the functions unavailable and needs to be avoided; you can check if something is a built in function by typing in the name into the Python command line Checking if names are used already [CS-1]
  + len
  + # expected output:
  + <built-in function len>
  + str
  + # expected output:
  + <class 'str'>
  + xyz # a string that is not already used
  + # expected output:
  + Traceback (most recent call last):
  + File "<stdin>", line 1, in
  + NameError: name 'xyz' is not defined

# we can proceed to use xyz as a variable name

* the assignment is made with the **assignment operator =** (this is different from the equality operator, ==, which tests values for equality)
* the assigned value can be of any type
* if a new value is assigned to the same name further in the program flow, the variable takes on the new value and its type
* the value assigned to a variable can be that of a *literal*, of *another variable* or of a *more complex expression*

Naming a variable and giving it a value (run interactively) [CS-2]

# setting variables with literal values

number = 3

a\_decimal\_number = 3.3

myname = "Jane"

isItTrue = True

r2d2Status = 'active'

# setting variables from other variables

ano0therNum = number

C3POStatus = r2d2Status

# some invalid names:

True = "the value of true"

2bOrNot2b = 22

# setting a new value to an existing variable

a\_decimal\_number = 5.5

**Variable use**

* a variable must be named and set (given a value) before it is used in the program flow
* a variable is used in the same way that a literal value would be used, for example to provide the value of another variable

Using variables (run interactively, after [[CS-2]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_3_variables.html#py-var-create)) [CS-3]

# display values in the Python console

a\_decimal\_number

C3POStatus

# add two numbers

number + a\_decimal\_number

# create a new variable and print it

status = "Like r2d2, " + myname + " is " + r2d2Status

print(status)

**Assignment operators**

Apart from the standard assignment operator, =, there are the **augmented assignment** operators, which combine another operation with assignment: +=, -=, \*=, /= etc.

For example, x += 3 means the same as x = x + 3, but is generally more efficiently implemented under the hood.

Augmented assignment operators (run interactively) [CS-4]

x = 3

x += 4

print(x)

name = "John"

name += " Doe"

print(name)

### Input

One way of setting variables that we will use a lot in this module is input from the console. At this beginner level of programming, reading data from the console, in conjunction with printing, is an easy way to achieve two-way communication with the outside world and make a program useful.

When the function input() is called, it waits for a line to be typed in and the Enter key pressed on the console. Then it returns the text contained in the submitted line, as a string value.

Input (run as script) [CS-5]

print("Enter some text:")

input()

# enter text on console, press Enter

print("Enter your name:")

name = input()

# enter name on console, press Enter

print("Your name is", name)

The code using input() in [[CS-5]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_3_variables.html#py-input) can be shortened by using an argument to input(). This argument is printed to prompt the user, rather than relying on a separate call to print().

###### Example [CS-6] - Input with argument (run as script)

**input("Enter some text: ")**

**# enter text on console, press Enter**

**name = input("Enter your name: ")**

**# enter name on console, press Enter**

**print("Your name is", name)**

It is important to remember that the value returned by input() is always a string. If a different type is required, a conversion needs to take place.

###### Example [CS-7] - Input with type conversion (run as script)

**x = input("Enter a whole number: ")**

**y = input("Enter another whole number: ")**

**# x and y are both of type string**

**print("The sum of your numbers without conversion is", x + y)**

**# int(x) and int(y) will be integers and so will their sum**

**print("The sum of your numbers with conversion is", int(x) + int(y))**

# Week 2 – String and Lists, Design and Principles

## Strings

### Strings as objects

The string type, str, is a **sequence** type in Python (along with *list*, *tuple* and *range*).

As with all other types in Python, instances of str are *objects*. Objects and methods were [briefly introduced](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_1_begin.html#ooswd-l1-03) at the beginning.

#### Some string methods

Here we will look at some methods that can be called on a string object. The full list can be found [**here**](https://docs.python.org/3/library/stdtypes.html#text-sequence-type-str).

* A very useful way of constructing strings is to use the [**format()**](https://docs.python.org/3/tutorial/inputoutput.html#the-string-format-method) method (methods are like functions, but are associated with a piece of data and called 'on it'). This particular method allows inserting values of any type into other strings.

##### Example [CS-1[ - String format (run interactively)

**"Hello {0}, it must be {1} years since we last met!".format("Bob", 7)**

* There are various useful text modification string methods.

##### Example [CS-2] - Various text modification methods (run interactively)

**# call on literal**

**"hello".upper()**

**# call on variable**

**name = "tyrannosaurus rex"**

**name.upper()**

**name.capitalize()**

**name.isnumeric()**

**"234".isnumeric()**

#### Operators

We have already seen the concatenation operators + and \*, but there are others that can be used with the string type. We will look at some of them but a full list is available [**here**](https://docs.python.org/3/library/stdtypes.html#common-sequence-operations).

Behind the scene, operators are implemented as methods. If you run help(str) in the interactive interpreter to get information on the string type, you will notice a number of methods listed at the beginning that start and end with double underscores (\_\_). These methods are not intended to be called directly but are implementations of operators. For example, the method \_\_contains\_\_() will be called when the operator in is used. We can use both operator and method to the same effect, **but it is the operator that should be used**.

##### Operator *in*

The return value of this operator is of type **boolean**.

###### Example [CS-3] - Checking for a substring with the in operator (run interactively)

**dinosaur = "tyrannosaurus rex"**

**"ran" in dinosaur**

**"swam" in dinosaur**

**# method works but should not be used!**

**dinosaur.\_\_contains\_\_("ran")**

**dinosaur.\_\_contains\_\_("swam")**

##### Comparison

The return value of this operator is of type **boolean**.

###### Example [CS-4] Comparing strings (run interactively)

**name1 = "Jane"**

**name2 = "Jack"**

**name1 == name2**

**name1 < name2**

**"." < ":"**

In code snippet [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_4_strings.html#py-str-comparison) the result of the last expression may be surprising, but will be explained in the next section, 'UTF-8'.

##### Operator []

Square brackets can be used to read an individual character or a range of characters from a string **(slicing).**

The return value of this operator is a **string**.

some\_string[i]

get a single character at index i

some\_string[i:j]

get a substring starting at index i and ending before index j

some\_string[:j]

get a substring starting at the beginning of the original string and ending before index j

some\_string[i:]

get a substring starting with the character at index i and containing the remainder of the original string

###### Example [ CS-5] - Accessing a substring using square brackets (run interactively)

**name = "Charlie"**

**name[0]**

**name[3]**

**name[2:4]**

**name[4:]**

**name[:3]**

**name[:]**

**name[:3] + name[5]**

Negative indices can be used to access characters and slices in a string. The last character is in negative position -1, the second last in position -2 and so on.

###### Example [ CS-6] - Accessing a substrings with negative indices (run interactively)

**name = "Charlie"**

**# these are equivalent to the first 5 expressions in** [**[CS-5]**](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_4_strings.html#py-str-sq-brack)

**name[-7]**

**name[-4]**

**name[-5:-3]**

**name[-3:]**

**name[:-4]**

#### Immutability

Strings are **immutable** objects, which means that a string value cannot be changed 'in place'. A string can be read and its value or parts thereof used as a basis for the creation of new strings, but these will be stored in a different memory location from the original string, the original string remaining intact.

###### Example [ CS-7] - Replacing substrings to demonstrate immutability (run interactively)

**name = "Johnathan"**

**shorter = name.replace("natha", "")**

**name**

**shorter**

In code snippet [[CS-7]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_4_strings.html#py-str-immut) a call is made on the variable name for the string "natha" to be replaced with an empty string "". The function does not modify the string stored in variable name (this would not be allowed because strings are immutable) but creates a new string with the new value and returns that new value. We store it in variable shorter.

### UTF-8 Encoding

In Python 3 strings are encoded using the UTF-8 character encoding system. An encoding system maps the visual appearance of characters to numeric values that can be stored and manipulated in a computer. A full list of characters mapped by UTF-8 can be found [**here**](https://mcdlr.com/utf-8).

* When strings are compared, what is really compared is the UTF-8 numbers that underlie them.
* The functions ord() and chr() convert between the character and an integer representing its UTF-8 code.

##### Example [ CS-8] - Converting between characters and their UTF-8 codes (run interactively)

**# ord() must be called on a string of length 1**

**ord("a")**

**ord("abc") # cases an error**

**chr(45)**

**chr(ord("a"))**

**ord('ć')**

* The first 128 characters in the UTF-8 mapping are exactly the same as the ASCII encoding map. ASCII (from American Standard Code for Information Interchange) was designed in the 1960s and widely used in computing ever since. Backward compatibility is thus important and UTF-8 provides it.

In code snippet [[CS-8]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_4_strings.html#py-str-utf8) the function call ord('a') yields 97, which is in the ASCII range. The call ord('ć') yields 263, which is outside of the ASCII range.

#### In-class exercise[CE-A]

Write a short program that takes two inputs: a character and an integer (call it an encryption key). The program should then encrypt the character into another character and print this new character out. The encryption algorithm is as follows: add the encryption key to the character's UTF-8 code to get the UTF-8 code of the character's encrypted value.

##### [EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/strings/str_encoding.py)

String comparison cannot be used directly for **alphabetical ordering**. Consider the example in code snippet [[CS-9]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_4_strings.html#py-str-alpha-ord). Lower case 'a' is considered to be greater than upper case 'a', because the UTF-8 code for the upper case letter is lower, but for alphabetical ordering we would like all letters 'a' to have the same value, regarless of case. This can be solved by converting all the text to either upper case or lower case before comparison.

##### Example [ CS-9] - Alphabetical ordering of strings (run interactively)

**'a' > 'A'**

**# the above is true because**

**ord('a') > ord('A')**

**name = 'Jack'**

**name\_again = 'jack'**

**name == name\_again**

**# the above yields False**

**# alphabetical ordering**

**name.lower() == name\_again.lower()**

## Lists

### What is a list in Python?

* A list is an ordered collection of objects of any type. A list is itself an object.
* The list type in Python is called list.
* list is a sequence type, along with **str**, **tuple** and **range.**
* A list literal is denoted using square brackets, with the elements of the list specified inside the brackets and separated by commas. The elements can be literals or variables.
* A list can be empty.
* A list can contain another list.

Unlike strings**, lists are mutable objects in Python**, which means that their value can be modified 'in place' i.e. at the memory location where they are stored.

##### Example [CS-1] - Defining lists (run interactively)

**a\_string = "hohoho"**

**an\_int = 12345**

**a\_list = ["hello", 4, a\_string, 5.6, "a", an\_int, True]**

**type(a\_list)**

**empty\_list = []**

**listWithListElement = ["hi", ["Jane", "Doe", 21], True, 55.3]**

### List operations and methods

As lists are mutable, they support both [**operations that do not change the sequence object**](https://docs.python.org/3/library/stdtypes.html#common-sequence-operations) and [**those that do**](https://docs.python.org/3/library/stdtypes.html#mutable-sequence-types).

#### Operations that do not change the list object

These operations are the ones also supported by the string sequence type.

##### Concatenation (using operators + and \*)

Concatenation does not modify the lists that are being concatenated but returns a new list.

###### Example [CS-2] - List concatenation (run interactively)

**["a", "b"] + ['c', 'd']**

**['1', '2'] \* 3**

**a\_string = "hohoho"**

**an\_int = 12345**

**a\_list = ["hello", 4, a\_string, 5.6, "a", an\_int, True]**

**a\_list + ['additional string']**

**# check if a\_list has changed**

**a\_list**

##### Membership checking (using operator in)

The operator in allows us to check if an object is an element of a list.

###### Example [CS-3] - Checking for list membership (run interactively)

**theList = ['abc', 3, True, '']**

**'' in theList**

**3 in theList**

**4 in theList**

##### Accessing parts of a list by index (using operator [])

As with strings, square brackets can be used to access a single element of a list or a group of contiguous elements. The way indices are used with the square brackets is exactly the same as for accessing substrings. Positive and negative index values can both be used.

###### Example [CS-4] - Accessing elements of a list (run interactively)

**laugh = "hohoho"**

**letters = ["a", "b", 'c']**

**a\_list = ["hello", 4, letters, 5.6, "a", laugh, True]**

**a\_list**

**a\_list[0]**

**a\_list[2:5]**

**a\_list[5:]**

**a\_list[:-5]**

**# accessing element in sublist**

**a\_list[2][1]**

**# accessing substring in string element**

**a\_list[5][:2]**

##### Comparison

###### Equality operators

These can be used between any two lists. For two lists to be equal, they need to be of the same length and their elements in the same position must be of the same type and have the same value.

Example [CS-5] - Equality of lists (run interactively)

**['three'] == [3]**

**['3', 4] == ['3', 4]**

**[2] != [3]**

**['one', 'two', 'three'] == ['one', 'two']**

###### Relational operators

Two lists can be compared with relational operators if their elements in the same position are individually comparable. Comparisons are carried out left to right i.e. the relation between an element pair carries more weight if the elements are at a lower index.

Example [CS-6] - Relational operations on lists (run interactively)

**['3', 4] < ['3', 4] # False**

**['2', 4] < ['3', 3] # True**

**['2', 4] < [3, 3] # Error! ('<' not supported between instances of 'str' and 'int')**

#### Operations that change the list object

With square brackets (**the [] operator**) we can:

###### Example [CS-7] - assign a new object to a list element Setting individual element values of a list

**myList = ["hello", "hi", 5, "how are you"]**

**myList**

**myList[2] = "bonjour"**

**myList**

###### Example [CS-8] - replace any contiguous portion of a list with another list Replacing a sublist

**myList = ["hello", "hi", 5, "how are you"]**

**myList**

**myList[2:3] = ["bonjour", "hola"]**

**myList**

**myList[-1:] = ["good day"]**

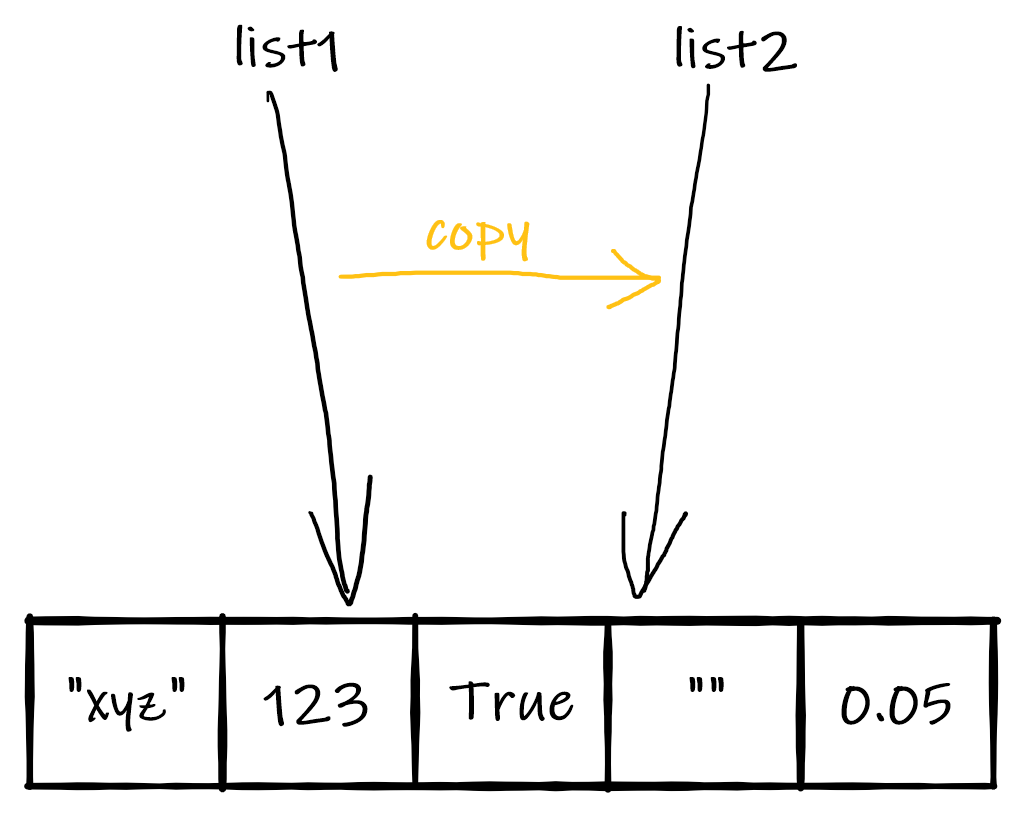
**myList**

#### Copying lists

There are several levels of copying that can be carried out with lists.

##### Creating a new name for a list vs. copying

###### New name

Giving an existing list a new name (list2 = list1) 

Assigning a list object to a new name results in the creation of a new name for the list. The new name points at the exact same list in memory that the assigned object is at. What is copied is only the connection (pointer) from name to object. It is a **copy of a reference**.

Example [CS-9] - Create a new name for a list (run interactively)

**list1 = ['abc', 123, 1.23, True]**

**list2 = list1**

**list1**

**list2**

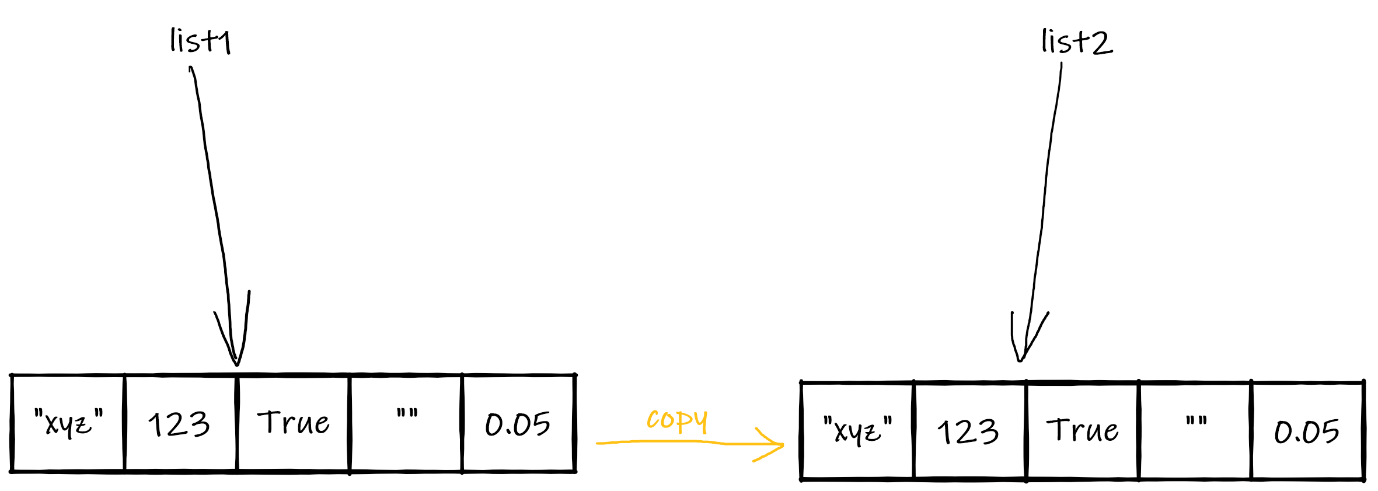
**# check it is the same list**

**list1[1] = 456**

**list1**

**list2**

###### **Copy**

**Copy (list2 = list1.copy() or list2 = list1[:])** 

A copy of a list results in the creation of a new name for the list and a copy of the content in a new memory location. The new name points to the copy of the content. It is a **copy of a value**.

Example [CS-10] - List copy (run interactively)

**list1 = ['abc', 123, 1.23, True]**

**list2 = list1[:]**

**list3 = list1.copy()**

**# check copies are different lists**

**list2.append('newEl')**

**list3.remove('abc')**

**list1**

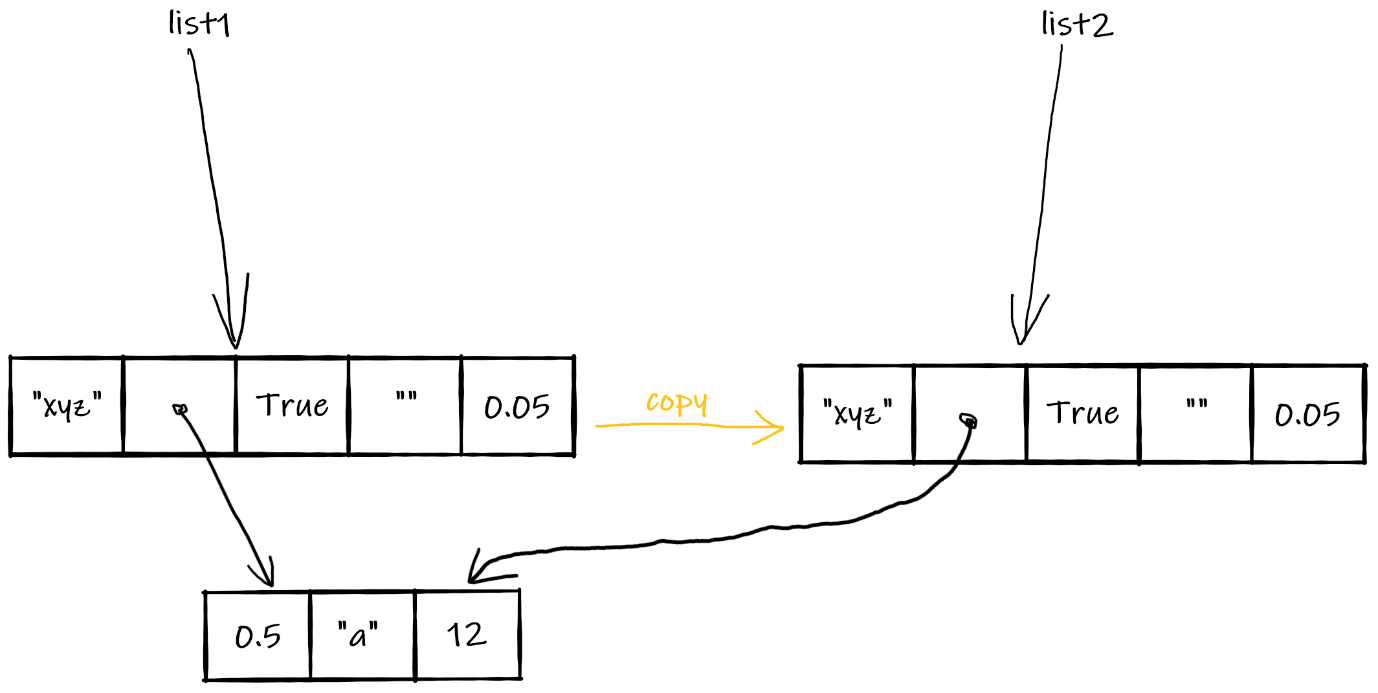
**list2**

**list3**

###### Shallow copy vs. deep copy

Copying a list with nested lists requires a special operation i.e. a **deep copy**. The copy discussed above is known as a **shallow copy**.

Shallow copy

Copy of list containing lists (list2 = list1.copy() or list2 = list1[:]) 

The figure *Shallow copy* shows what happens when the function copy() is called on (or equivalent operator [:] is applied to) a list with nested lists.

The only object copied is the top-level list. For nested lists, only the reference (pointer) is copied.

*Example [CS-11] - List copy (run interactively)*

**list1 = ['abc', [1, 2, 3], 1.23, True]**

**list2 = list1[:]**

**# see that the lists are the same**

**list1**

**list2**

**# change original list top level**

**list1[2] = 4.567**

**# change original list nested level**

**list1[1][1] = 222**

**# see list1**

**list1**

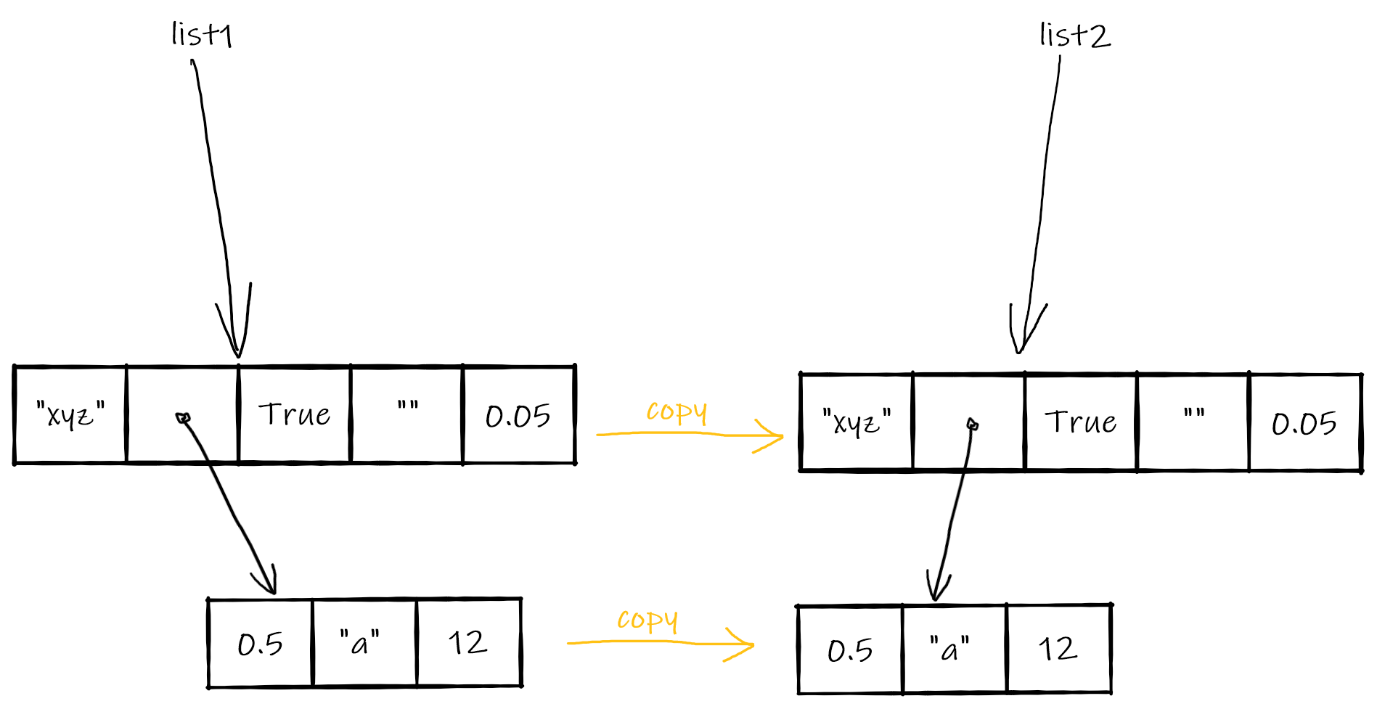
**# see list 2**

**# top level is not modified**

**# nested is modified (because shared)**

**list2**

Deep copy

Deep copy of list containing lists (list2 = copy.deepcopy(list1)) 

When a deep copy is made, the whole hierarchy of nested lists is copied, to any level of depth. For this, the function copy.deepcopy().

*Example [CS-12]--- List copy (run interactively)*

**# import the copy module**

**import copy**

**# define list1 and create a deep copy**

**list1 = ['abc', [1, 2, 3], 1.23, True]**

**list2 = list1.deepcopy()**

**# see that the lists are the same**

**list1**

**list2**

**# change original list top level**

**list1[2] = 4.567**

**# change original list nested level**

**list1[1][1] = 222**

**# see list1**

**list1**

**# see list 2 (no modifications)**

**list2**

#### Creating Lists from strings

##### Casting a string to list

Casting a string to a list returns a list with one string element for each character in the original string.

Example [CS-13]--- Casting strings to lists (run interactively

**myName = "Jane Brown"**

**list(myName)**

##### Splitting a string

With the split() method we have more control over how the string is split into list elements, in that a separator can be specified. When the separator is not specified, the split creates substrings of non-whitespace characters that were separated by whitespace characters in the original string.

Example [CS-14]--- The split method (run interactively)

**myName = "Jane Brown"**

**myName.split()**

**wordsStr = "last, first, penultimate, second, unspecified"**

**wordsList = wordsStr.split()**

**wordsList**

**type(wordsList)**

**# use a separator**

**wordsStr.split(sep=", ")**

#### Creating strings from lists

##### Casting

Casting from list to string will not produce the result we generally want. The returned string is just the list, printed out.

###### *Example [CS-15]---* Casting list to string (run interactively)

**str(["a", "b", "c"])**

##### Joining elements into a string

The string type provides a method [**join()**](https://docs.python.org/3/library/stdtypes.html#str.join) for joining elements of an iterable object (which a list is) into a string.

The syntax for calling the method is **<separator string object>.join(<list object>),** i.e. the method is called on a string object that should act as separator in the resulting string.

###### *Example [CS-16]---* Joining list elements into a string (run interactively)

**' '.join(["a", "b", "c"])**

**''.join(["A", "n", "n"])**

**separator = ", "**

**list = ["python", "javascript", "java"]**

**separator.join(list)**

### Transforming lists

#### Lambda Expressions

##### A very short introduction to lambda expressions

A lambda expression

* is an expression that defines a function i.e. a reusable piece of functionality
* returns (evaluates to) an **anonymous** function i.e. a function without a name
* can be called as a function
* can be assigned to a variable and thus get a name by which it can be called
* can be passed in as an argument to another function and called by the parameter name
* takes the form

**lambda <parameter list> : <return value created from parameter>**

* takes arguments defined by the parameter list between lambda and the colon
* returns the value to which the expression right of the colon evaluates
* can use variables defined outside of it

##### Examples [CS-17] - Examples of lambda expressions and their use (run interactively)

**# define a lambda expression**

**# that increases a number by 3**

**# and assign it to f (name it f)**

**f = lambda x : x + 3**

**# run it**

**f(4)**

**f(2.3)**

**# we don't have to name a lambda**

**# expression but can call it**

**# directly (it must be placed**

**# in parentheses)**

**(lambda x : x + 3)(4)**

**(lambda x : x + 3)(2)**

**# define and use a lambda**

**# expression with two parameters**

**(lambda x, y : x \* y)(3, 4)**

**(lambda x, y : x \* y)(23, 2)**

**# define a lambda expression**

**# that checks if a number is**

**# greater than 3 (the return**

**# type is boolean)**

**g = lambda x : x > 3**

**# run it**

**f(3)**

**f(3.1)**

##### Examples [CS-18] - Examples of lambda expressions that use variables external to them

**# some names**

**names = ["ben", "con", "ann", "dan"]**

**# define a lambda expression that returns the**

**# index in the list 'names' for the given string value**

**get\_index\_for\_name = lambda nm : names.index(nm)**

**# use it**

**get\_index\_for\_name('ann')**

**get\_index\_for\_name('ben')**

**# some scores**

**scores = [23, 67, 34, 12]**

**# define a lambda expression that returns the**

**# value in the list 'scores' corresponding by position**

**# to the given value in the list 'names'**

**get\_score\_for\_person = lambda nm : scores[names.index(nm)]**

**get\_score\_for\_person("ben")**

**get\_score\_for\_person("dan")**

**# sort the name list by scores, using**

**# sorted(), which does not modify the list;**

**# we need the original list intact as**

**# it is being used in the lambda expression**

**sorted(names, key=get\_score\_for\_person)**

**# or**

**sorted(names, key=lambda n : scores[names.index(n)]**

List index() Method Syntax

**list\_name.index(element, start, end)**

#### Sorting Lists

Lists can be sorted using **the** [**sort()**](https://docs.python.org/3/library/stdtypes.html#list.sort) **method**, which modifies the list **in place**. The **function** [**sorted()**](https://docs.python.org/3/library/functions.html#sorted) **has the same functionality but does not modify the list, creating a copy of it to sort and return.**

For a list to be sortable, either

1. the values of all its elements need to be comparable

##### Example [CS-19]---Sorting a list without a key (run interactively)

**names = ['ben', 'con', 'don', 'ann']**

**names**

**names.sort()**

**# names has changed in place**

**names**

**# now try using sorted() to preserve the original list**

**games = ['scrabble', 'ludo', 'monopoly', 'carcasonne']**

**games**

**sorted(games)**

**# check games again (should not have changed)**

**games**

**# try to sort a list with types that cannot be compared**

**mixed\_list = ['abcd', 34, '', 'aaa', 0, 4, 0.0]**

**mixed\_list.sort() # error!**

**# however if all types can be compared:**

**another\_mixed\_list = [2, True, 0.2]**

**another\_mixed\_list.sort**

**# True compared as 1 and the sort works  
another\_mixed\_list**

b) the key argument is a function (either named or [lambda expression](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_6_lambda.html)) that converts all the elements to comparable values

##### Example [CS-20] - ---Sorting a list with a key (run interactively)

**# use the bool() casting function as key**

**#-------------------**

**mixed\_list = ['abcd', 34, '', 'aaa', 0, 4, 0.0]**

**mixed\_list.sort() # error**

**# pass key to convert to boolean**

**mixed\_list.sort(key=bool)**

**# elements that equate to False are listed first**

**mixed\_list**

**# use the len() function as key**

**#-------------------**

**word\_list = ['hello', 'hi', 'how are you']**

**word\_list**

**word\_list.sort(key=len)**

**# words are sorted by length**

**word\_list**

**# use a lambda expression as key**

**#-------------------**

**# some names**

**names = ["ben", "Con", "ann", "Dan"]**

**# try to sort them**

**names.sort()**

**# order is not alphabetical because**

**# because some names are capitalised;**

**# we use a lambda expression to**

**# convert to all-lower before sorting**

**names.sort(key=lambda n : n.lower())**

(lambda (some\_variable): what ever you need to do, (mapped to whatever list you want to apply itto)

**Lambda is an expression that generates a new function to be called later – as it is an expression it can be used in places where def cannot i.e. lists and dictionary litterals… (p505)**

#### Map and filter

While sorting results in a list with the same elements but possibly in a different order, the [**map()**](https://docs.python.org/3/library/functions.html#map) function maps each element of the list to create a list of new elements, while the [**filter()**](https://docs.python.org/3/library/functions.html#filter) function applies a filter to each element of the list to create a possibly shorter list.

* The syntax of the map function is **map(<mapping function>, <list to be mapped>).** The function returns an iterable object, which means that we need to cast it to a list to get a list with the elements resulting from the mapping. Thus the full expression for mapping a list to a list is so **list(map(<mapping function>, <list to be mapped>)).**

##### Example [CS-21] - Mapping a list to a new list

**names = ['jim', 'john', 'jane', 'jen']**

**# map to name lengths**

**list(map(len, names))**

**# map to capitalised names**

**list(map(lambda s : s.upper(), names))**

* The map function can work on several input lists in parallel. They all have to be the same length.

##### Example [CS-22] - Mapping two lists to a new list

**names = ['jim', 'john', 'jane', 'jen']**

**surnames = ['lock', 'smith', 'stone', 'mason']**

**# join name and surname, capitalising both**

**list(map(lambda n, s : n.capitalize() + " " + s.capitalize(), names, surnames))**

* The syntax of the filter function is **map(<filter function>, <list to be filtered>).** The filter function must return a boolean value. Similarly to the map function, the result of the filter function must be explicitly transformed to a list.

##### Example [CS-23] - Filtering a list

**names = ['jim', 'john', 'jane', 'jen']**

**# find the names that contain an 'n'**

**list(filter(lambda n : n.find('n') >= 0, names))**

#### List comprehensions

List comprehensions are a special construct that combines mapping and filtering of lists. In most cases, it is easier to read than the map/filter combination. The syntax of a list comprehension is

**[** expression1 **if** condition1 **else** expression2 **for** list\_element **in** list **if** condition2 **]**

* The black text in the syntax specification constitutes the **fixed parts** of the comprehension, including keywords and brackets.
* The statement for list\_element in list specifies the list with expression list and a variable to store the current element (list\_element) during iteration over the list , much like in a usual for-in loop.
* expression1 is the mapping expression, which transforms each list element to an element of the new list (it corresponds to the function or lambda expression passed to the map function).
* The part with the blue background is the filter, with condition2 as the filtering condition, which must return a boolean value. This part is optional.
* The part with the green background is a construct that allows conditional mapping: expression1 is used if condition1 is true, while expression2 is used if not. This part is also optional.

##### Example [CS-24] - Examples of list comprehensions

**names = ['jim', 'jill', 'john', 'jane', 'jen', 'jack']**

**# examples with no filter or conditional mapping**

**# -----------------------------------------------**

**# map to string length**

**[ len(n) for n in names ]**

**# map to names in all-capitals**

**[ n.upper() for n in names ]**

**# examples with filter only**

**# --------------------------**

**# map to string length, including only names containing an 'e'**

**[ len(n) for n in names if 'e' in n ]**

**# map to names in all-capitals, including only names longer than 3**

**[ n.upper() for n in names if len(n) > 3 ]**

**# examples with conditional mapping only**

**# ---------------------------------------**

**# map to all-capitals if name length is less than 4, otherwise don't change**

**[ n.upper() if len(n) < 4 else n for n in names ]**

**# map to length+1 if name contains an 'e', otherwise map to length-1**

**[ len(n) + 1 if 'e' in n else len(n) - 1 for n in names ]**

**# examples including all parts of the construct**

**# ----------------------------------------------**

**# map to all-capitals if name length is less than 4, otherwise don't change,**

**# including only names with an 'e' (i.e. only if it has an e in the name)**

**[ n.upper() if len(n) < 4 else n for n in names if 'e' in n ]**

**# map to length+1 if name contains an 'e', otherwise map to length-1,**

**# including only names longer than 3 (only if**

**[ len(n) + 1 if 'e' in n else len(n) - 1 for n in names if len(n) > 3]**

Comprehensions are easier to write and to the trained eye are much easier to read. There are some scenarios where map/filter is still more concise: when the mapping function is built in (e.g. len or int) and when multiple lists need to be transformed.

In the list comprehension in the example below, the **zip** function needs to be **called to combine the names and surnames into tuples (n, s),** which are then transformed. This is a sort of a hack, since it creates an intermediate data structure that is needed only for processing.

**Maps vs comprehension(neither is correct/incorrect, whatever you are used to )**

* **Comprehension can only work on one list, so need to use zip function**
* Map>> use as Jelena used to them
* Filter >>
* Map+ filter is mapping function is already in existence
  + i.e. len function a spre-defined
* mapswill work with more than one list…

##### Example [CS-25] - Example of list comprehension using two input lists

**names = ['jim', 'jill', 'john', 'jane', 'jen', 'jack']**

**surnames = ['lock', 'smith', 'stone', 'mason', 'bell', 'ringer' ]**

**# join name and surname, capitalising both, using map**

**#**

**list(map(lambda n, s : n.capitalize() + " " + s.capitalize(), names, surnames))**

**# the same transformation using a list comprehension**

**[ n.capitalize() + " " + s.capitalize() for (n, s) in zip(names, surnames) ]**

### Tuples and enumerate

In the section on list comprehensions, we have seen a structure called the **tuple**. [Tuples](https://docs.python.org/3/tutorial/datastructures.html#tuples-and-sequences) are sequences, like lists, but are immutable. They are used in contexts where multiple values are needed instead of one value. Tuple literals can be defined either as a comma-separated list or a comma separated list in parentheses.

Tuples Is a list of variables – for cases where we need more than one variable, can contain any number of elements, really very much like lists but **not modifiable – are immutable like strings**

#### Example [CS-26] - Defining tuples

**t1 = 1, 2, 3**

**t2 = (1, 2, 3)**

**t3 = 3, 4, 5**

**t1 == t2 # expected output: true**

**t1 == t2 # expected output: false**

Python has a built-in function called **enumerate,** which returns index-value tuples of a sequence. This is very useful where the **indices of list elements are needed for processin**g. Like map and filter, this function's output is not a list but can be converted to one.

#### Example [CS-27] - Enumerate function

**names = ['jim', 'john', 'jane', 'jen']**

**# create an enumerated list**

**list(enumerate(names))**

**# get a list of strings preceded by 1-based indices**

**[ str(x[0] + 1) + " " + x[1] for x in enumerate(names) ]**

**# or**

**list(str(x[0] + 1) + " " + x[1] for x in enumerate(names))**

### Picking a random element from a list

Many programs will require the generation of some sort of random information. Usually, a random number is generated in a set range.

Python has a function that allows picking a random item from a list directly, **random.choice().**

#### Example [CS-28]--- Picking a random element (run interactively)

**aList = ['hello', 4, ['a', 'b', 'c'], 5.6, 'a', 'hohoho', True]**

**aList**

**import random**

**random.choice(aList)**

**random.choice(aList)**

**random.choice(aList)**

**random.choice(aList)**

**random.choice(aList)**

## Simple program design steps (roughly!)

1. Identify inputs and outputs
2. Identify the order of inputs and outputs
3. Identify which inputs affect which outputs
4. Design the flow of control
5. Identify repetitions and variables
6. If necessary, go back to step 4
7. Fill the programmatic details
8. After implementation and testing, fix any imperfections

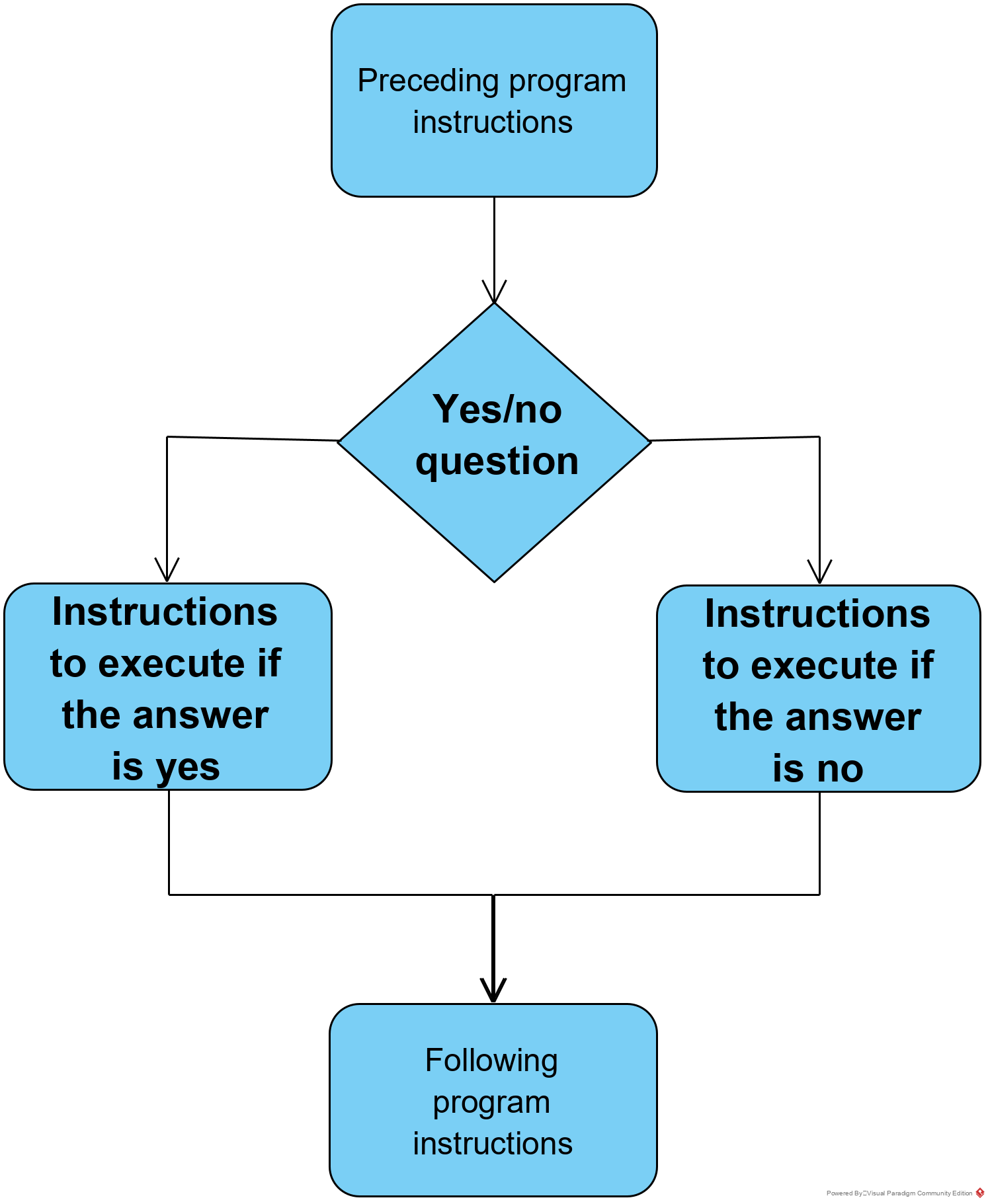
# Week 3 – Conditionals, while loops

## Conditionals

To conditionally execute code need a conditional statement… condition the flow…

### What is a conditional statement

Exists in all imperative programming languages.. and functional progamming langs.. can’t program without it

The flowchart of basic conditional statement in a computer program – basic chart that applies to all langs not just python

The conditional statement is one of the three essential ingredients needed for expressing algorithms:

* Assignments (assign to variables/lists etc)
* **conditional statements**
* loops

### The basic conditional statement in Python

##### Example [CS-1] - Example of a basic conditional statement in Python

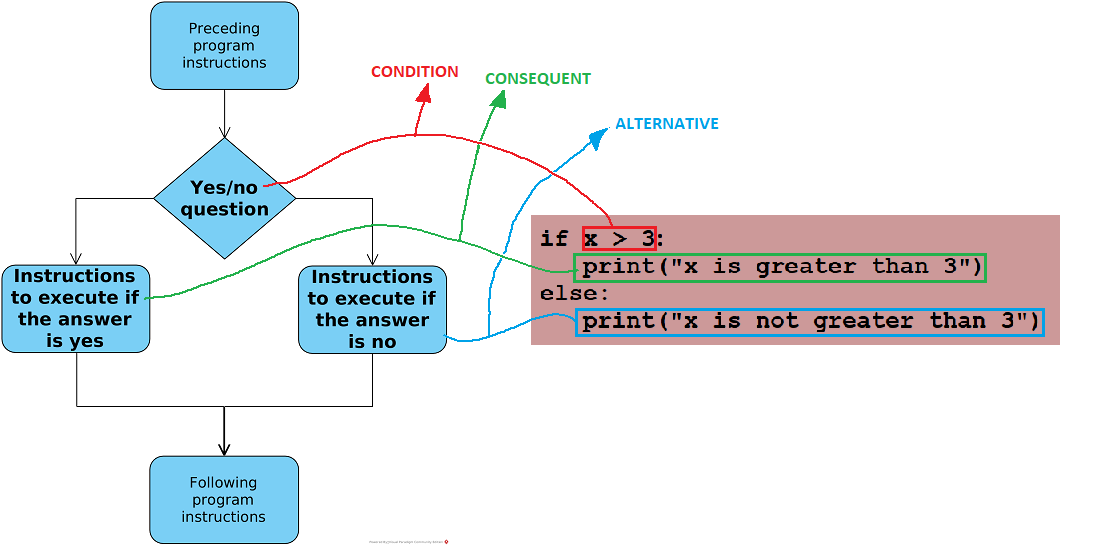
Indentation ism meaningful – must be indented by same amount in python. Is binary (doesn’t matter how much you indent -indentation is meaningful in python) is Binary.. is either met or it isn’t…)

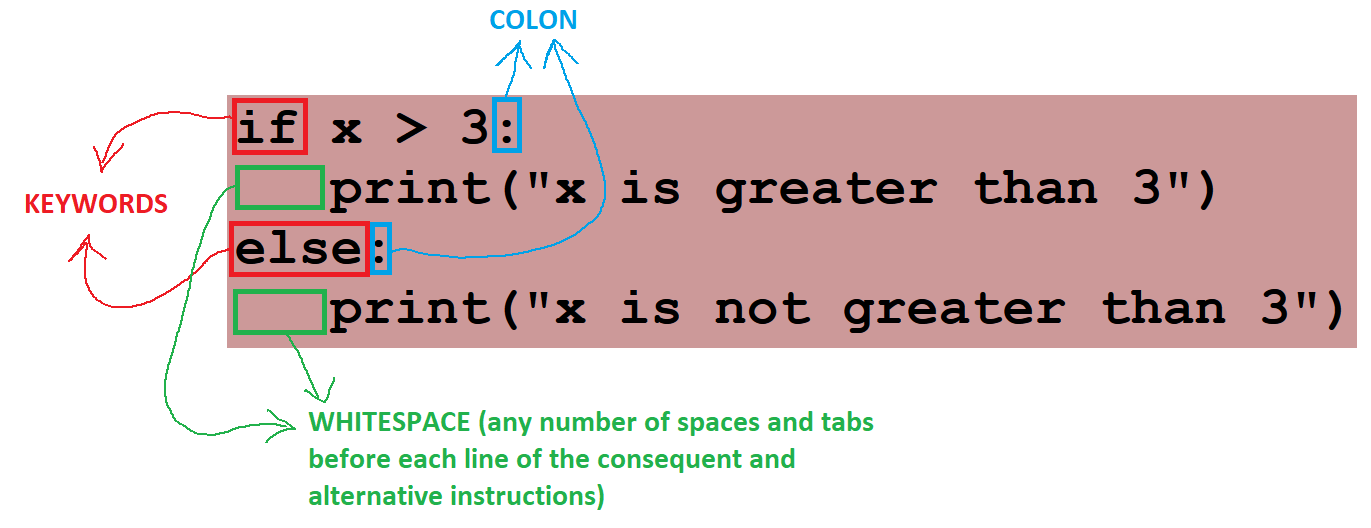
**if x > 3:**

**print("x is greater than 3")**

**else:**

**print("x is not greater than 3")**

Relating a Python conditional statement to the generic flowchartThe syntax of a Python conditional statement

Answer **must translate as a Boolean** i,e. true or false (condition must be yes or no)

* + The **condition** must be a yes/no question, which in programming language terms means that it must **evaluate to a boolean**.
  + Both the **consequent** and the **alternative** can be any sequence of programming instructions.

#### In-class exercise [CE-A]

Use the basic conditional statement example above in a runnable script that checks a number input by the user and tells the user if the number is greater than 3.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/conditionals/cond_basic.py)

#### In-class exercise [CE-B]

What will the following program print?

**if True:**

**print("It's true")**

**else:**

**print("Not a chance")**

### Conditional statement variations

#### Without an alternative ('if' statement)

1. The else keyword and the alternative instructions listed below it can be left out entirely

###### Example [CS-2] - Example of an 'if' statement in Python (run interactively)

**if x > 3:**

**print("x is greater than 3")**

* In this case the alternative consists of no instructions i.e. **nothing is done if the condition is not fulfilled.**
* By writing this kind of conditional statement, we are implicitly stating the alternative and should always keep that in mind. Not considering it can lead to errors.

##### Example [CS-3] - A script with a problem caused by an omitted branch

**x = 4**

**if x == 3:**

**y = x \* 2;**

**print(y)**

#### In-class exercise [CE-C]

Write a script to try out the code above. What happens when you run it?

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/conditionals/cond_no_alt.py)

#### With more than two branches

* + Conditions evaluate to booleans and can have only two outcomes, hence the only way to split the flow into more than two branches is by using **nested conditional statements** that split the already existing branches.

##### Example [CS-4] - Example of creating multiple branches of execution

**if x > 3:**

**if x > 10:**

**print("x is greater than 10")**

**else:**

**print("x is greater than 3 but not greater than 10")**

**else:**

**if x > 0:**

**print("x is 1, 2 or 3")**

**else:**

**print("x is not a positive number")**

#### The elif keyword

* When a nested conditional statement is inside the *alternative* part of the conditional statement i.e. in the block that follows the else keyword, the else and if keywords can be collapsed into the keyword **elif**. With the use of elif the code snippet above turns into the following.

###### Example [CS-5] - Example of using the elif keyword

**if x > 3:**

**if x > 10:**

**print("x is greater than 10")**

**else:**

**print("x is greater than 3 but not greater than 10")**

**elif x > 0:**

**print("x is 1, 2 or 3")**

**else:**

**print("x is not a positive number")**

The logic of code snippets [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_7_conditionals.html#py-multi-branch) and [[CS-5]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_7_conditionals.html#py-elif-intro) is identical.

* + In order to 'get rid of' the first nested conditional statement in code snippet [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_7_conditionals.html#py-multi-branch), we would have to **change the order** in which the conditions are checked.

By reordering the checking of the conditions and using the elif keyword, write a code snippet logically equivalent to [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_7_conditionals.html#py-multi-branch) that does not have any nested conditional statements.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/conditionals/cond_elif.py)

### Covering all bases

When programming conditional statements it is important to **stay in control** by accounting for both branches of every condition, implicit or explicit.

#### Implicit alternatives

* + An alternative branch may be implicit with good reason, but deciding this must be part of the programming process.
  + For example, I may want to issue a warning inside a stock-keeping application if the stock of a product is very low. If the stock is not low I *do not want the program to do anything*. This is a valid candidate for a conditional statement without an alternative branch. I am sure that nothing needs to be done if the condition is not met i.e. if there is enough stock.
  + Cases where it makes sense to not have the “else” as don’t’ want it to do anything

##### Example [CS-6 ] Example of a conditional statement validly without an alternative

**stock = 0 /\* calculated in some way \*/** *(this is java format for comments)*

**if stock < 2:**

**print("WARNING: Stock is very low. Put in an order with supplier!")**

**/\* no else needed \*/**

* + Nested conditional statements can make it harder to see 'un-covered' branches of execution. Consider the following example.

##### Example [CS-7 ] - Example of a missing alternative

**age = int(input("Please enter your age:"))**

**hasProvisional = input("Do you have a provisional licence (y/n):")**

**if age >= 17:**

**if hasProvisional == "y":**

**print("You are eligible for a driving licence.")**

**else:**

**print("You are not eligible for a driving licence.")**

Spot the problem in code snippet [[CS-7]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_7_conditionals.html#py-if-no-alt-incorr) and rewrite it so that it works correctly. Start by using a flowchart to identify the required algorithm.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/conditionals/cond_licence_missing_alt.py)

#### Hidden alternatives (don’t use if if if if… )

* + For all conditions, the condition and it's negative can be programmed into separate conditional statements. While this approach can be logically correct, it is
    - inefficient (uses more computing power)
    - harder to write
    - harder to understand and maintain
  + Look at the following example where we want the program to do one thing when x is equal to 3 and another when it is not equal to 3. The two conditional statements 'cover all bases' but inefficiently and 'messily'.

###### Example [CS-8] - Example of hidden alternatives

**print("Guess my lucky number (it's an integer): ", end="")**

**x = int(input())**

**if x == 3:**

**print("You have guessed my lucky number!")**

**if x <> 3:**

**print("Sorry, that's not my lucky number :-(")**

##### In-class exercise [CE-F]

Rewrite code snippet [[CS-8]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_7_conditionals.html#py-if-no-alt-bad) so that it takes advantage of the Python conditional statement's built in alternative handling.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/conditionals/cond_luck_num_hidden_alts.py)

### Complex conditions

1. We use the logical operators **AND**, **OR** and **NOT** (keywords **and**, **or** and**not**) to combine conditions into complex conditions.
2. Depending on the logic, complex conditions can be used to 'flatten', partially or fully, a nested conditional statement structure.

#### In-class exercise [CE-G]

Rewrite the driving licence example to use a complex condition instead of a nested conditional statement.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/conditionals/cond_licence_complex_cond.py)

### Exercises

#### In-class exercise [CE-H]

Write a Python script for Icy Icecream Parlour's automated cash register. It should:

* + Communicate appropriately with the customer, printing prompts when needed
  + Read in the number of icecream scoops that the customer wants to buy
  + Calculate the amount due, based on a scoop price of €1.20
  + Ask the customer if they are over 65 and apply a discount of 20% to the price if they are
  + Print the amount due

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/conditionals/cond_icecream.py)

#### In-class exercise [CE-H]

Write a Python script for Cringy Cinema film listing machine. The script should not contain any nested conditional statements. It should:

* + Have three films stored as string variables: a G-rated, a 15-rated and an 18-rated film
  + Communicate appropriately with the customer, printing prompts when needed
  + Read in the customer's age
  + Print a list of suitable films

[EXAMPLE SOLUTION 1](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/conditionals/cond_cinema.py)

[EXAMPLE SOLUTION 2](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/conditionals/cond_cinema2.py)

**Note on the example solutions**

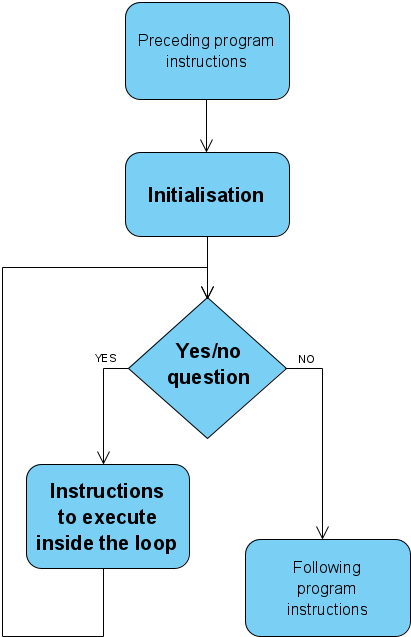
Both solutions are logically correct. The first one uses one conditional decision tree to decide which list of films to print. The second one uses two conditional statement trees, one for the over-15 and another for the over-18 film.

The first example is a larger program i.e. would take up more space in memory. The other example (owing to the two decision trees) is slower. While the difference in size (three function calls) and speed (ratio 5:6 for conditional checks) are negligible in these small examples, in real-life projects design decisions of this kind are made as part of code optimisation.

*In class the main goal is to design****correct****solutions. If a particular form (e.g. an elif 'flat' decision tree) is required, this will be stated in the assignment specification but otherwise any correct solution is satisfactory.*

## While loops

### What is a loop in programming?

The flowchart of a generic loop in a computer program

The loop is one of the three essential ingredients needed for expressing algorithms:

* + assignments
  + conditional statements
  + loops

A loop is **used when the same instructions need to be iterated (repeated several times**):

* + to improve the development process (speed, readability and maintainability)
  + to optimise the program (smaller code size)
  + to make use of the processing power of a computer (i.e. get things done on a scale that humans cannot achieve)
  + to allow the program to be written at all:
    - the number of iterations may not be known in advance if it depends on environment communication or is random
    - if the program represents a 'server' (runs infinitely)

### The while loop in Python

The main function of a loop is to **repeat some group of instructions over and over again**. The loop in code snippet [[CS-1]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-inf) does exactly that: it prints out the number 1 over and over again. The condition is number == 1 and in this context always true. The call to time.sleep() has been added to space the iterations by 1 second, to avoid looping at a speed that hogs the CPU and freezes the computer running it. Once running, this little program will have to be stopped externally (**Ctrl-C** on the command line or stop button in PyCharm).

##### Example [CS-1] - Example of an infinite while loop in Python

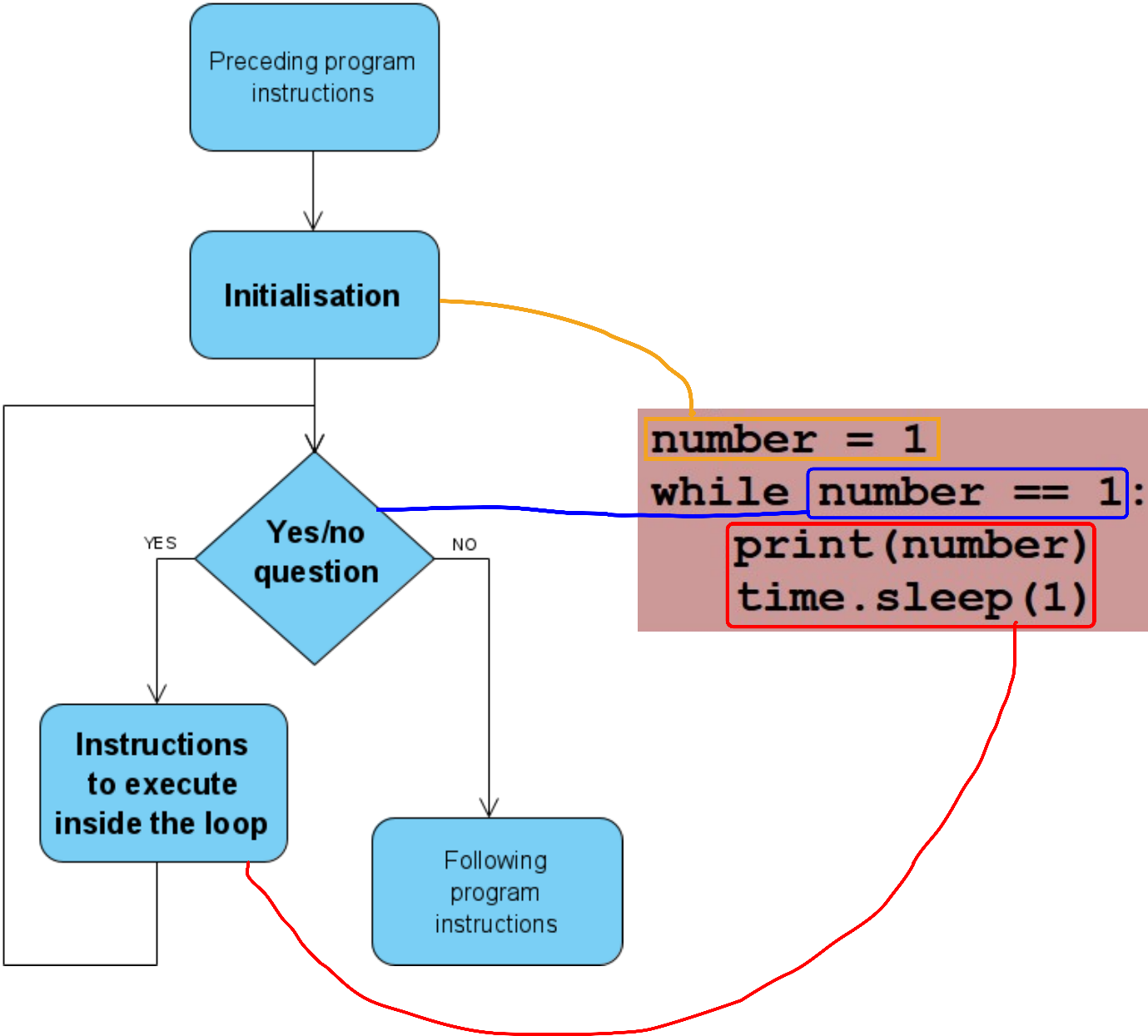
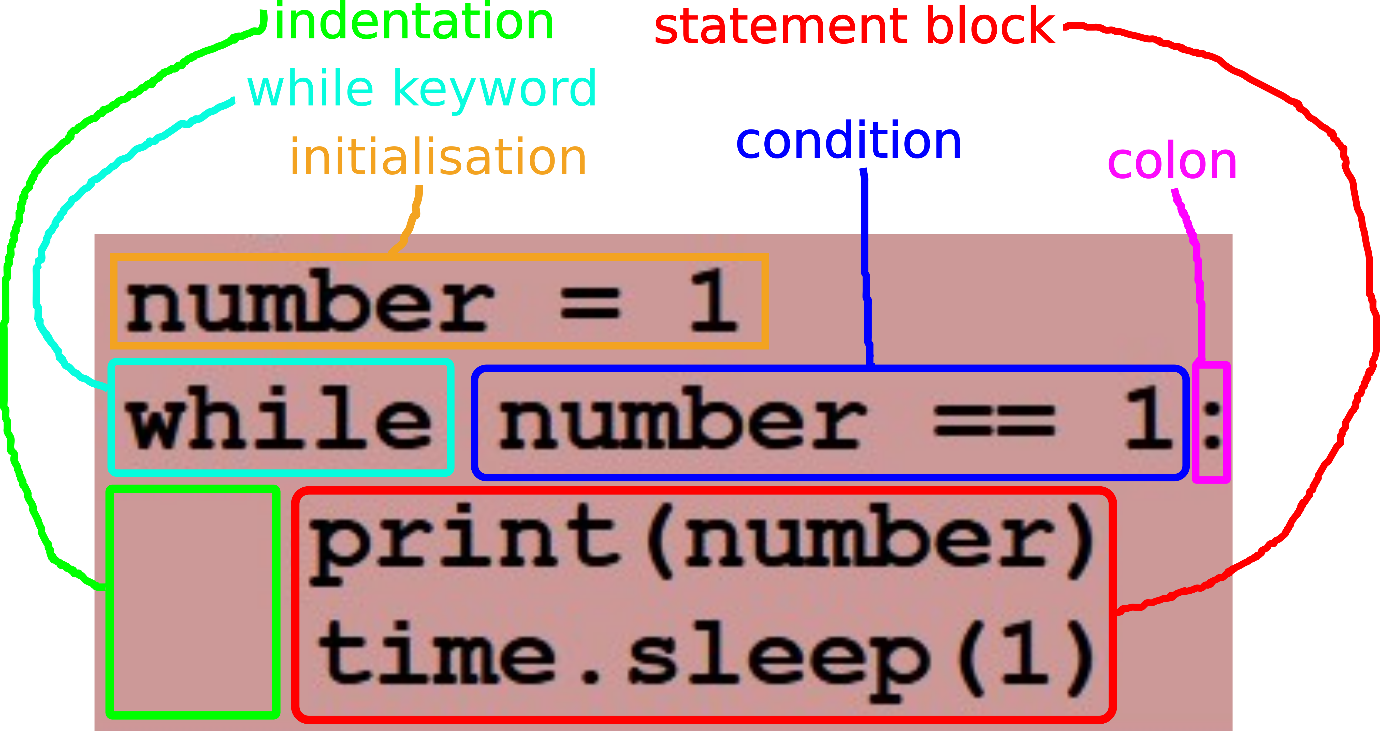
**import time**

**number = 1**

**while number == 1:**

**print(number)**

**time.sleep(1)**

Relating a Python while loop to the generic flowchartThe syntax of a Python while loop

### Exiting the loop

Most of the time, loops do not run infinitely and are written so that the conditional expression evaluates to False after a finite number of iterations. This can happen in different ways.

* + The variable checked in the conditional expression may be one that assumes values from a **fully determined sequence**, eventually reaching a value that makes the expression False. In this case the number of iterations is known in advance.

##### Example [CS-2] - While loop with the number of iterations fully determined at the time of coding

**number = 1**

**while number < 11:**

**print(1)**

**number = number + 1**

#### In-class exercise[CE-A]

Write a Python script that produces the output below. Use a while loop to print the lines between the first and the last one.

**Wherever repetition occurs that’s a candidate for a loop**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**\* \***

**\* \***

**\* \***

**\* \***

**\*\*\*\*\*\*\*\*\*\*\*\*\*\***

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/x4_print_rect.py)

#### In-class exercise [CE-B]

Write a loop that prints 10 random numbers between 1 and 11 (inclusive).

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/x10_print_rand.py)

Often, the number of iterations cannot be determined at the time of coding but **at run-time is known before the loop is executed**. This is the case in code snippet [[CS-3]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-det-rt). Notice that if the user enters 1, the statement block of the loop will never be executed (0 iterations).

##### Example [CS-3] - A while loop with iterations fully determined at run-time

**iterCount = input("Please enter a positive whole number: ")**

**number = 1**

**while number <= iterCount:**

**print(1)**

**number = number + 1**

#### In-class exercise [CE-C]

Write a Python script that:

* generates a random number between 1 and 10 (inclusive) and stores it in a variable called iterationCount
* prints out the number 1 iterationCount times using a while loop

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/xrand_print_1.py)

The variables checked in the conditional expression may, alternatively, depend on some **factor external to the program** and subsequently not be known to it before the loop runs. In this case the number of iterations cannot be known in advance. For example, a random number generator or user input, if used as the variable checked in the condition, would result in this scenario. The value obtained externally that halts the execution of the loop is often called the **sentinel** value.

##### Example [CS-4] - While loop with an unknown number of iterations owing to a random process

**import random**

**number = 1**

**while number != 11:**

**print(1)**

**number = random.randint(1, 11)**

#### In-class exercise [CE-D]

Rewrite code snippet [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-indet-rand) so that the loop statement block prints out the value of number rather than 1. The loop in [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-indet-rand) always prints out 1 at least once, because the random number is not generated or checked until the first iteration. Modify the script to generate a random number before entering the loop, resulting in no numbers printed if the first number 'drawn' is 11.

Test your script. It will take a few runs, but you will eventually get one where no numbers are printed out.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/sent_rand_from_start.py)

##### Example [CS-5] - While loop with exit conditioned on user input

**number = int(input("Please enter a number between 1 and 10 or 11 to stop: "))**

**while number != 11:**

**print(1)**

**number = int(input("Please enter a number between 1 and 10 or 11 to stop: "))**

##### Example [CS 6] - While loop with exit conditioned on data from a file (requires a text file called data.txt in the directory where it is running [[sample contents]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/data.txt))

**numList = []**

**f = open("data.txt", 'r') # opens file data.txt for reading**

**l = f.readline() # reads a line from the file**

**isNum = ''.join(l.strip().split('.', 1)).isnumeric()**

**while isNum:**

**numList += [ float(l) ]**

**l = f.readline()**

**isNum = ''.join(l.strip().split('.', 1)).isnumeric()**

**f.close() # closes the file**

**print(numList)**

### Iteration initialisation and choosing conditional variables

#### Desiging initial conditions for all iterations (not only the first one)

While initialisation is generally explicitly named as part of the standard while loop structure, variables (both conditional and other) must be set up in each iteration in preparation for the next one. The initial conditions of the loop are available to the first iteration as it is executed. The conditions created in the first iteration are available to the second iteration and so on.

For example, in code snippet [[CS-2]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-det-ct), the variable number must be 'set up' for each iteration:

* + for the first iteration, it is set to 1 (before the loop starts)
  + for subsequent iterations, the preparation consists of number being incremented

In this example number represents the iteration number and is not used in any way except as a conditional variable.

##### In-class exercise[CE-E]

Examine all the code snippets and in-class exercise solutions and identify all the set up that is performed for the loop iterations.

#### Choosing conditional variables

It is often a good idea to design the statement block (the body) of the while loop first, then identify conditional variables. Most of the time the conditional variables can be chosen from variables already used in the statement block. We will look at this process through some examples.

##### Example 1 - processing a predetermined sequence

Let's say we want to print out all the numbers from 1 to 10. We want the first iteration to print 1, the second to print 2 etc. (10 iterations in all).

* + First, let's **design the statement block**:
    - Because the number printed is changing from iteration to iteration, it needs to be a variable, e.g. n
    - Then we can easily write the main part of the statement block as print(n)
    - The last thing that is missing in the statement block is some code to set up n for the next iteration: n = n + 1
    - Also, n needs to be set up for the first iteration. The statement n = 1 placed before the loop achieves this.
  + Now we need to **design the while loop condition**. This takes a small bit of analysis. The variable n is moving upwards by one in each iteration and we want the looping to stop when n goes beyond 10. So n naturally lends itself as the basis for the condition, n < 11.

We can now write the short program. It is provided as code snippet [[CS-7]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-des-det-ct).

###### Example [CS 7] - A loop that prints out the numbers from 1 to 10

**n = 1**

**while n < 11:**

**print(n)**

**n = n + 1**

##### Example 2 - processing a sequence defined at run-time

Let's say we want to print out all the characters in a word input by the user, each character on a separate line followed by its UTF code in square brackets.

* + Again, let's design the statement block first:
    - The character changes between iterations so must be a variable, e.g. c
    - This makes it easy to write the printing code: **print("{0}[{1}]".format(c, ord(c)))**
    - Between iterations, we need the script to move from character to character in a word entered by the user. Let's call the word variable theWord and we'll worry about obtaining it later.
    - Now is there a quantity related to the characters in the word that can be updated consistently (with the same operation) for all iterations? The index of the characted in the word fits the bill and the update operation is index = index + 1. Thus we have to add a variable representing the index and use that to extract the characters from the word with c = theWord[index].
    - We also need to set index to 0 for the first iteration.

Our design so far looks like this:

**index = 0**

**while <condition>:**

**c = theWord[index]**

**print("{0}[{1}]".format(c, ord(c)))**

**index += 1**

* + To design the condition, we must analyse the expected sequence of indices. It starts at 0 and ends at len(theWord) - 1. We want the loop to stop after that and can achieve it with the condition index < len(theWord).

We can now write the short program. It also needs to include the reading of the word as user input. It is provided as code snippet [[CS-8]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-des-det-rt). The length of the word is stored in a variable for optimisation (i.e. so that the length is not calculated in every iteration).

###### Example [CS 8] - A loop that prints out the characters of a word

**score**

##### In-class exercise [CE-F]

Write a Python script with a while loop that reads a list of words input by the user and prints them one per line, each word preceded by its ordinal number (1, 2, 3 etc.).

**To DO** > Do this above

[Example solution](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/list_print.py)

##### Example 3 - processing a sequence defined by a sentinel

Let's say we want to collect some sentences from the user and add them to a list, which is then printed at the end.

* + Statement block design:
    - The number of sentences is not predefined so we need to introduce a sentinel (special sentence) that signals the end of the entry process. We could use an empty string but that would be prone to mistakes. Instead let the sentinel be a full stop (.).
    - The action we need to perform on the entered sentences is to add them to a list: sentenceList += [ currSentenceStr ]. So we need a list variable and we need a string to hold the latest sentence input by the user.
    - To create conditions for the next iteration, we need to read in the next sentence: currSentenceStr = input("Enter your next sentence or full stop (.) if finished: ")
    - To create the initial conditions we need to initialise all variables:
      * for the sentence, use the same code as above: currSentenceStr = input("Enter your next sentence or full stop (.) if finished: ")
      * for the list, start with an empty one: sentenceList = []
  + Condition design: this can be a simple check for the sentinel value: currSentenceStr != "."

The full program, including the printing at the end, is given as code snippet [[CS-9]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-des-sent).

###### Example [CS 9] - A loop that collects sentences until a 'stop' value is entered

**sentenceList = []**

**currSentenceStr = input("Enter your next sentence or full stop (.) if finished: ")**

**while currSentenceStr != '.':**

**sentenceList += [ currSentenceStr ]**

**currSentenceStr = input("Enter your next sentence or full stop (.) if finished: ")**

**print("Your sentences:\n------------------\n" + "\n".join(sentenceList))**

##### In-class exercise [CE-G]

Write a Python script that reads in three words separated by whitespace, entered by the user. The program should check the number of words entered and that they contain only alphabetic characters. If the input does not meet these requirements the program should ask the user to try again and repeat the process in a while loop until the input is valid.

Next, the program should print the words, each in a new line and followed by its length in square brackets.

To DO >> do this above

[**Example solution**](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/three_words.py)

#### Condition independent from iteration functionality

All the code snippets in section [*Exiting the loop*](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#ARTCLID000018) ([[CS-2]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-det-ct), [[CS-3]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-det-rt), [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-indet-rand) and [[CS-5]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-indet-input)) are examples of loops where the task performed in the while loop statement block is independent of the condition. In all those examples, the task in an individual iteration is to print the number 1.

The statement block core (print(1)) does not define any variables that can be used in the condition. Variables are defined specifically for the purpose of controlling the iteration and updated in the statement block to reflect the iteration requirements.

##### In-class exercise [CE=H]

Write a Python script that fetches a random number between 1 and 3, asks the user to guess it and prints if the guess was correct. It should use a while loop to repeat this 3 times and print, at the end, how many out of three guesses were correct.

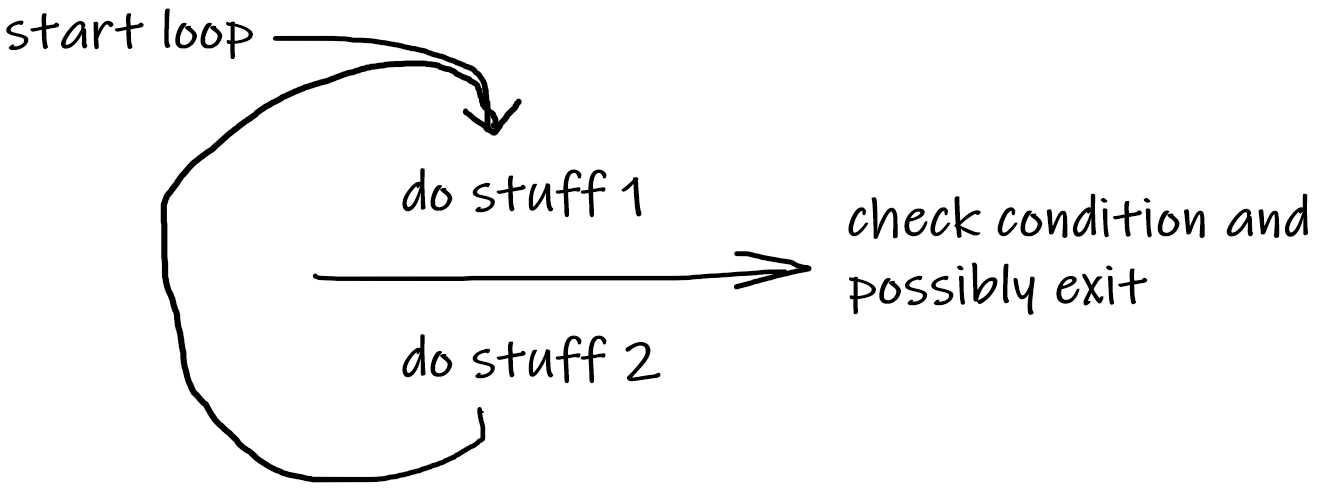
To DO >> do this above

[Example solution](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/independent_cond.py)

### While loop variations

#### Statement block variation

If we step away from the code for a minute we can have a look at iterative processes in a slightly more general form, rather than in terms of Python constructs. Consider the following model of an iterative process:

A general model of iterative processes

Also consider how the iterations of a simple loop such as that in code snippet [[CS-2]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-det-ct) might be thought of on first analysis (model and pseudocode):

**do stuff 1 set n**

**-----------------> check -----------------> check**

**do stuff 2 print number 1**

The while keyword and condition correspond to the check in the general model i.e. the while loop iterations (condition + statement block) must start where the arrow is. This means that set n for the first iteration of the general model has to happen before the while loop is entered (this is initialisation) and that set n for subsequent iterations must happen at the end of the while statement block (iteration initialisation). After mapping this model onto the Python while loop in [[CS-3]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-det-rt), we get:

**set n (iter 1) n = 0**

**-----------------> check while number < 11:**

**print number 1 print(1)**

**set n (iter 2 etc.) number = number + 1**

We can map any loop onto this model but it is particularly useful for analysing a problem and translating it into a while loop. We will work through several examples, introducing the new constructs where suitable along the way.

#### While loop without initialisation

If *stuff 1* of the model can be packed into the expression that *check* looks at, then there is no need for initialisation before the while loop. In that case the model would look more like:

**-----------------> do stuff 1 and check**

**do stuff 2**

While this can be implemented using functions (which we will be covering later in the module), otherwise self-contained conditionals do not occur often in Python because Python keeps statements and expressions separate (this is not the case in all programming languages), not allowing assignment as part of an expression, which a loop most often needs to keep track of its state using a variable. For loops handle this implicitly and are the more suitable choice of loop in a lot of cases, but let's stay with the while loop for the moment.

Examples of while loops that do not need initialisation can be found though, and the following code snippet represents one.

##### Example [CS10]- While loop with no initialisation

**import random**

**while random.randint(1, 11) != 11:**

**print(1)**

#### In-class exercise [CE-I]

Rewrite the loop in code snippet [[CS-10]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-no-init) to use a value input by the user in the condition, rather than a random value.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/while_no_init.py)

#### The front-loaded loop statement block (do-while loop)

Let's have a look at a loop that has nothing in the *stuff 2* part of the model.

A user is entering a password and they are allowed 5 mistakes before they get locked out of an account. The stored passwords are hashed, which means that the entered value must be hashed before it is compared with the stored value.

**do stuff 1 ask user to enter password**

**read input password**

**hash the password**

**set attempt count**

**-----------------> check ----------------------------> check**

**do stuff 2 [nothing]**

We know that the check corresponds to the beginning of a while loop, which means that developing the pseudocode for implementation in Python we get:

**set attempt count (1) |**

**ask user to enter password (first time) |--> initialise**

**read input password |**

**hash the password |**

**-----------------> check input vs. stored pwd**

**and that attempt < 6**

**set attempt count (add 1) |**

**ask user to enter password (again) |--> statement**

**read input password | block**

**hash the password |**

this eventually turning into something like this in Python:

##### Example [CS 11]- A while loop with identical initialisation and loop body [CS-11]

**storedHash = hash("abcde")**

**attempt = 1**

**pwd = input("Please enter your password: ")**

**hashed = hash(pwd)**

**while hashed != storedHash and attempt < 6:**

**attempt += 1**

**pwd = input("Please enter your password: ")**

**hashed = hash(pwd)**

With the similarity between the initialisation and statement block (which may contain a lot more code in a real case), it would be cleaner to have the instructions appear only once in the program. A *do-while* loop would be what we need here but Python does not have this construct. The way to emulate it is to give a dummy pass value to the conditional variables to force the first iteration of the loop to execute. After that proper checking of the conditional takes over.

##### Example [CS 12] - Using dummy conditional variable values to emulate a do-while loop

**storedHash = hash("abcde")**

**attempt = 0**

**hashed = hash("") # dummy: hash of "" (not allowed as password)**

**while hashed != storedHash and attempt < 6:**

**attempt += 1**

**pwd = input("Please enter your password: ")**

**hashed = hash(pwd)**

It is often the case that something slightly different, particularly in prompting the user, needs to be done in the first iteration. The dummy variable can be used to check if the iteration is the first one or not. Code snippet [[CS-13]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-dummy-for-diff) is the same as [[CS-12]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-dummy-cond) except that it prints a slightly different prompt the first time around.

##### Example [CS 13] - Using dummy variable value to check if in first iteration

**storedHash = hash("abcde")**

**attempt = 0**

**hashed = hash("") # dummy: hash of "" (not allowed as password)**

**while hashed != storedHash and attempt < 6:**

**attempt += 1**

**if hashed == hash(""):**

**pwd = input("Please enter your password: ")**

**else:**

**pwd = input("Password incorrect. Please try again: ")**

**hashed = hash(pwd)**

A 'do-while' pattern can be recognised from requirements. It emerges from scenarios where the sentinel value needs to be treated the same as any values before it. Password entry is such a case: both the incorrect passwords and the correct password (the sentinel) are not processed further (i.e. are treated exactly the same).

##### In-class exercise[CE-J]

Write a script including a while loop that reads in words from the user, putting them together into a sentence and printing out the sentence when finished. A word with a full stop at the end signals the end but should be included in the sentence.

Write two versions of the script: one that does not employ a dummy conditional variable and one that does.

Sample expected interaction:

--python--

Enter the next word for the sentence (with . at end if last): I  
Enter the next word for the sentence (with . at end if last): am  
Enter the next word for the sentence (with . at end if last): happy.  
Your sentence: "I am happy."

--python--

[EXAMPLE SOLUTION FOR V1](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/do_while_not.py)   [EXAMPLE SOLUTION FOR V2](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/do_while_done.py)

#### The loop statement block with checking in the middle

Let's look at an example where both the *stuff 1* and the *stuff 2* parts of the model are represented by some instructions in the loop statement block. In code snippet [[CS-14]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-mid-check-no-break) the initialisation (corresponding to *stuff 1* of the model) is repeated as the second part of the loop statement block, so that the natural loop can be fitted to a while loop.

##### Example [CS 14] - A while loop implementing a loop with condition naturally occurring mid-block

**numList = []**

**numStr = input("Please enter a number between 1 and 10 or 11 to stop: ") #**

**if not numStr.isnumeric(): #**

**print("The input is not a whole number. It will be ignored.") #--> stuff 1**

**number = 0 #**

**else: #**

**number = int(numStr) #**

**while number != 11: #**

**if number < 1 or number > 10: #**

**print("The number is not in range. It will be ignored.") #--> stuff 2**

**else: #**

**numList += [ number ] #**

**numStr = input("Please enter a number between 1 and 10 or 11 to stop: ") #**

**if not numStr.isnumeric(): #--> stuff 1**

**print("The input is not a whole number. It will be ignored.") #**

**number = 0 #**

**else: #**

**number = int(numStr) #**

**print(numStr)**

In the case of a loop like this, the break construct (see following section) can be used to exit an iteration in the middle, with the actual condition set to True and not really acting as a condition. The code snippet [[CS-15]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-mid-check-with-break) shows how this can be implemented in Python.

##### Example [CS-15] - Using the break construct for mid-block exit from iteration

**numList = []**

**while True:**

**numStr = input("Please enter a number between 1 and 10 or 11 to stop: ") #**

**if not numStr.isnumeric(): #--> stuff 1**

**print("The input is not a whole number. It will be ignored.") #**

**number = 0 #**

**else: #**

**number = int(numStr) #**

**if number == 11: #--> check**

**break # w/ break**

**if number < 1 or number > 10: #**

**print("The number is not in range. It will be ignored.") #--> stuff 2**

**else: #**

**numList += [ number ] #**

**print("Your number list: ", ' '.join(map(str, numList))) #**

##### In-class exercise[CE-K]

Rewrite code snippet [[CS-6]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-while-indet-file) so that obtaining the next number from the file features only once in the code (in the while loop statement block).

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/mid_check.py)

#### Fluidity of form and transformations

Now that we have seen some statement block forms, it is important to say that in almost all cases it is possible to transform between them, particularly with the use of functions to implement some parts of the functionality. The chosen form will be dictated by optimisation, readability and other considerations.

##### In-class exercise [CE-L]

Rewrite code snippet [[CS-15]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-mid-check-with-break) to subsume 'stuff 2' into the conditional structure in 'stuff 1' and introduce a conditional variable (end) that is prepared in the conditional structure, turning the loop into one with a do-while pattern.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/do_while_if_else.py)

#### Additional while loop constructs and their use

The keywords **break, continue** and **else** are used in Python for some additional constructs that can be used with the while loop.

* + break makes the flow of execution 'escape' the while loop and resume from the first instruction outside the loop
  + continue makes the flow of execution 'escape' from the current iteration and resume at condition checking for the subsequent iteration
  + else follows the loop immediately and contains a statement block to be executed after the while loop, but only if it did not come to an end due to a break instruction

##### Using break and else

If a while loop can exit both by the condition being false and by break being called, different behaviour may be required depending on how the exit happened. The else clause at the end of the Python while loop allows the inclusion of a statement block to be executed only if the while loop exits by False condition.

Code snippet [[CS-13]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-dummy-for-diff) is a good example of this. After the looping ends, it would be good to know whether it happened because the user exhausted the attempts of because they entered the correct password. In the example [[CS-16]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-break-and-else) break is used if the number of password entry attempts has been exhausted. The else clause is executed only when the password is correct and the loop exits as a consequence of that.

###### Example [CS-16] – Using break and else

Advanced so don’t worry if don’t do them… good to know

**storedHash = hash("abcde")**

**attempt = 0**

**hashed = hash("") # dummy: hash of "" (not allowed as password)**

**while hashed != storedHash:**

**attempt += 1**

**if attempt == 7:**

**print("You have run out of tries, sorry.")**

**break**

**if hashed == hash(""):**

**pwd = input("Please enter your password: ")**

**else:**

**pwd = input("Password incorrect. Please try again: ")**

**hashed = hash(pwd)**

**else:**

**print("You are now logged in.")**

rare for while loop to have else statement

###### In-class exercise[CE-M]

Write a while loop with break and else statements that reads words input by the user and prints out the list of collected words at the end. Do not add a word to the list if it contains non-alphabetic characters. Use sentinel values:

* full stop (.) to signal the end of input
* three dashes (---) for cancelling the entire process

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/break_and_else.py)

##### Using continue

Let's say we want to print out all the numbers from 1 to 100, leaving any numbers that contain zeroes. One way to escape from the iteration that we do not want to complete is by calling continue.

###### CS-17 – Example - Printing numbers with no zeroes [CS-17]

**count = 0**

**while count < 101:**

**count += 1**

**if '0' in str(count):**

**continue**

**print(count)**

*(comment - continue means hop out of this particular iteration and don’t execute anything below continue but go on to the next iteration)*

###### In-class exercise[CE-N]

Write a script in Python that uses a while loop with a break and a continue construct. It should read words input by the user, with sentinel value '.' for ending. In the case that the word is made out of alphabetic characters only, it should be converted to uppercase and its characters printed out separated by spaces, otherwise nothing else should be done in that iteration.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/break_and_continue.py)

(Comment - Don’t look at while loop variations except those extra key words unless you want more practice with them – get a feel for different way of setting up an iterating thru stuff)

## For Loops

(Comment – to work with for loops you need to know what a range is)

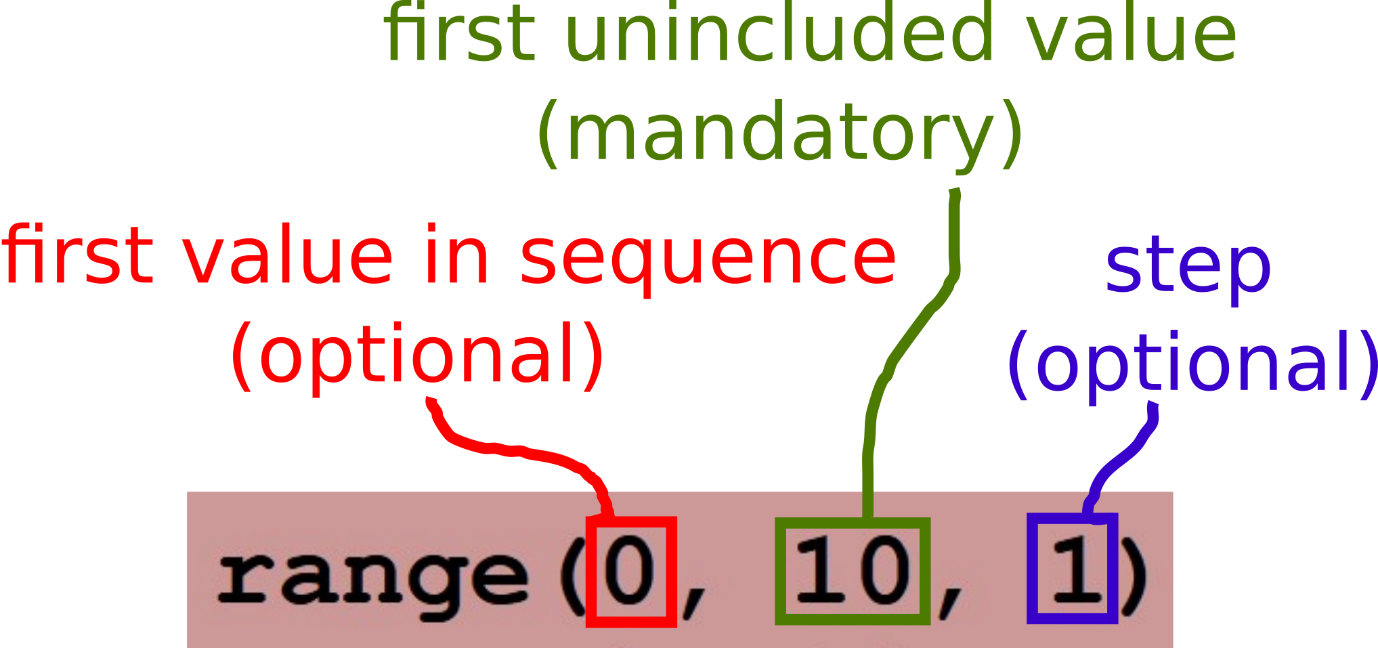
Similar to whiel only difference is we are iterating over pre-set of objects as in numbers ina range or avlues in a list)

### Ranges

The range is an **immutable** sequence type that allows easy definition of **integer sequences**. We introduce it here as for loops in Python are often defined with the use of ranges.

The syntax of a range expression (BNF)

**range(<first value, optional>, <first not-included value>, <step>)**

The syntax of a range expression (example)Examples of ranges [CS-1]

**range(0, 10, 1) # contains 0 1 2 3 4 5 6 7 8 9**

**range(0, 10) # same as above (step 1 is default)**

**range(10) # same as above (start value 0 is default)**

**range(0, 10, 3) # contains 0 3 6 9**

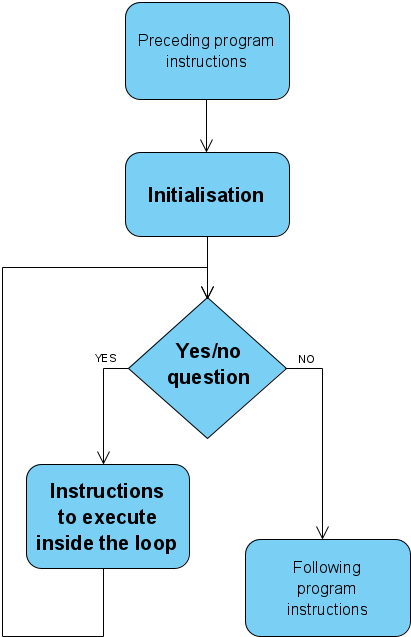
**range(10, 0, -1) # contains 10 9 8 7 6 5 4 3 2 1**

### For Loop syntax in Python

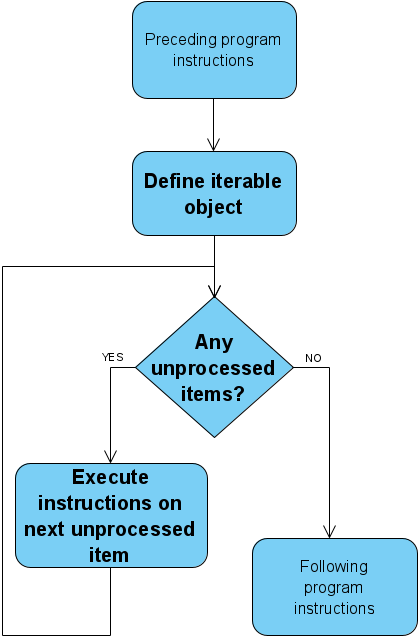
In Python, a for loop iterates over the members of any iterable object. A statement block is defined, like in the while loop, with instructions to be executed at each iteration.

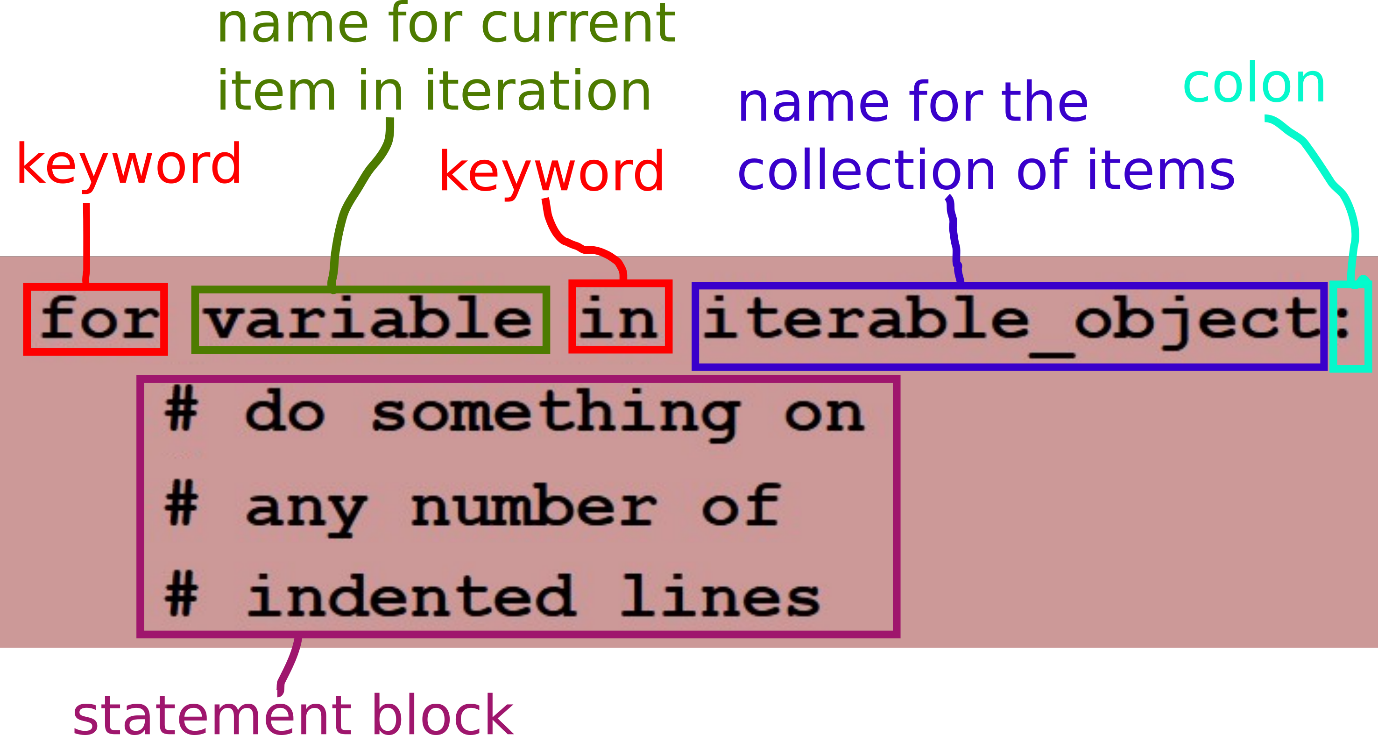
The picture shows the flowchart for a for loop, alongside the generic flow chart.

Generic loop flowchart

Generic loop flowchart

For loop flowchart

For loop flowchart

For loop syntax in Python

### Using for loops

Sequence types are iterable objects in Python. This means that we can design for loops that iterate over any of the sequence types we have covered so far.

In terms of the types mentioned in chapter *While loops*, section [*Exiting the loop*](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#ARTCLID000018), it is the iteration processes for which the iterations are determined before execution starts (either at coding time or at run time) that are suitable for implementation using for loops.

#### Iterating over ranges

###### Example [CS-2] - Iterating over ranges with for loops

**for number in range(0, 10, 1):**

**print(number)**

**for number in range(1, 11, 2):**

**print(number)**

**for number in range(11, 0, -3):**

**print(number)**

##### In-class exercise {CE-A]

Write a script that uses a for loop to print out 11 random integers between 1 and 11.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/for_range_rands.py)

##### In-class exercise {CE-B]

Write a script that uses a for loop to print out all the even integers between 1 and 100, inclusive, separated by spaces.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/for_range_ints.py)

#### Iterating over strings

###### Example [ CS-3] - Iterating over strings with for loops

**for character in "TU Dublin":**

**print(character)**

**word = "computing"**

**for character in word:**

**print("{}: {}".format(word.index(character), character))**

**# try the above with word = "Tallaght"**

**# better solution for above:**

**word = "Tallaght"**

**for index in range(0, len(word)):**

**print("{}: {}".format(index, word[index]))**

##### In-class exercise [CE-C]

Write a Python script that reads in a character string input by the user into a variable called **inputString**. Then, using a for loop, it should invert the order of the characters in the word. Finally, it should print out the string resulting from the inversion, the output of expression inputString[::-1] and the result of comparing the two.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/for_string.py)

#### Iterating over lists

###### Example [CS-4] - Iterating over lists with for loops

**nonMotorised = [ "rollerblades", "bicycle", "scooter", "sail wagon", "legs" ]**

**maxWordLen = max(map(len, nonMotorised))**

**for mode in sorted(nonMotorised):**

**print("--< " + mode.center(maxWordLen, ' ') + " >--")**

##### In-class exercise [CE-D]

Write a Python program that defines a list of items and stores it in a variable. Print out a menu, using the list items as menu items, each on a separate line and preceded by its ordinal number (1, not 0, for the first item etc.). You must use a for loop, iterating over the list rather than over list indices.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/for_list.py)

### Using break, continue and else with for loops

The break, continue and else constructs can be used with a for loop in the same way as they are used with while loops.

#### In-class exercise [ CE-E]

Rewrite code snippet [CS-16](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_8_loops_while.html#py-continue) with a for loop instead of the while loop.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops/for_continue.py)

# Week 4 – Nested Loops and Functions

* Nested loops do not involve any new language constructs.
* Both while and for loops can be nested within either while or for loops.
* Loops can be nested to any depth, with loops nested within other loops that are already nested themselves

## Nested Loops

### Nesting for loops

Nested for loops are fairly straightforward as a concept. They can be viewed as representing multi-dimensional tables, with an execution of the statement block in every cell. An example of printing out such a table is shown in code snippet [[CS-1]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_A_loops_nested.html#py-for-2dimtab).

#### Example [CS-1] - Printing a table with nested for loops

**cellWidth = 6**

**columnCount = 3**

**rowCount = 7**

**for rowNum in range(1, rowCount + 1):**

**for colNum in range(1, columnCount + 1):**

**print(("r" + str(rowNum) + "c" + str(colNum)).center(cellWidth, " "), end="")**

**print("")**

Use the nested loops in code snippet [[CS-1]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_A_loops_nested.html#py-for-2dimtab) to create a 2-dimensional list of elements that each contain the product of the row number and column number for the cell they are stored in. Print out the 2d list.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops_nested/for_nested_2dlist.py)

### Choosing the loops to use

Whether nested or not, we can be guided by the following when deciding if a loop should be 'while' or 'for':

* use a for loop when the iterations will be determined before execution of the loop
* use a while loop when the iterations depend on an external or random process

#### Example 1

For example, let's say we want our program to draw random numbers between 0 and 5, looping until a 0 is drawn. Every time a number is drawn, the sequence of numbers from 1 up to the drawn number is printed.

* The loop drawing random numbers until 0 needs to be a while loop (exit defined by random process)
* Printing out the sequence from 1 up to the drawn number needs to be done within the statement block of the while loop and requires a loop of its own, one that has it's iterations enumerated before it starts [1..drawn\_number], hence a for loop.

##### Example [CS-2] --Example of a for loop nested in a while loop

**import random**

**n = random.randint(0, 5)**

**while n != 0:**

**for i in range(1, n + 1):**

**print(str(i) + " ", end="")**

**print("")**

**n = random.randint(0, 5)**

Rewrite code snippet [[CS-2]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_A_loops_nested.html#py-for-in-while) so that, instead of printing the range from 1 to the drawn number, it prints numbers in the range [<drawn number>, <random number between 6 and 10>].

Did you have to change the types of loop used? Why?

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops_nested/while_for.py)

#### Example 2

Now let's write a script that repeats the following 10 times: draws random numbers between 0 and 5 until it gets a 5 and prints how many draws it took to get to a 5. A solution is in code snippet [[CS-3]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_A_loops_nested.html#py-while-in-for).

* At the top level we have the requirement to do something 10 times. That is a for loop, as the number of iterations is defined.
* Inside the loop there is a repetitive (drawing random numbers) with a random process determining the end. This needs to be in a while loop.
* This is also a case where we do not need to store the value of the random draw but just check it, which means we can use the call to r**andom.randint** directly in the conditional expression.
* We need to add a counter for the number of draws, which we set to 1 to account for the final draw of a 5.

##### Example [CS-3] - Random draws until 5, 10 times

**import random**

**for i in range(0, 10):**

**counter = 1**

**while random.randint(0, 5) != 5:**

**counter +=1**

**print(counter)**

Change the requirements for code snippet [[CS-3]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_A_loops_nested.html#py-while-in-for) so that a while loop is a better fit for the outer iteration. Write the corresponding code.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/loops_nested/while_while.py)

## Functions

### What is a function?

In programming, a function is a reusable piece of functionality.

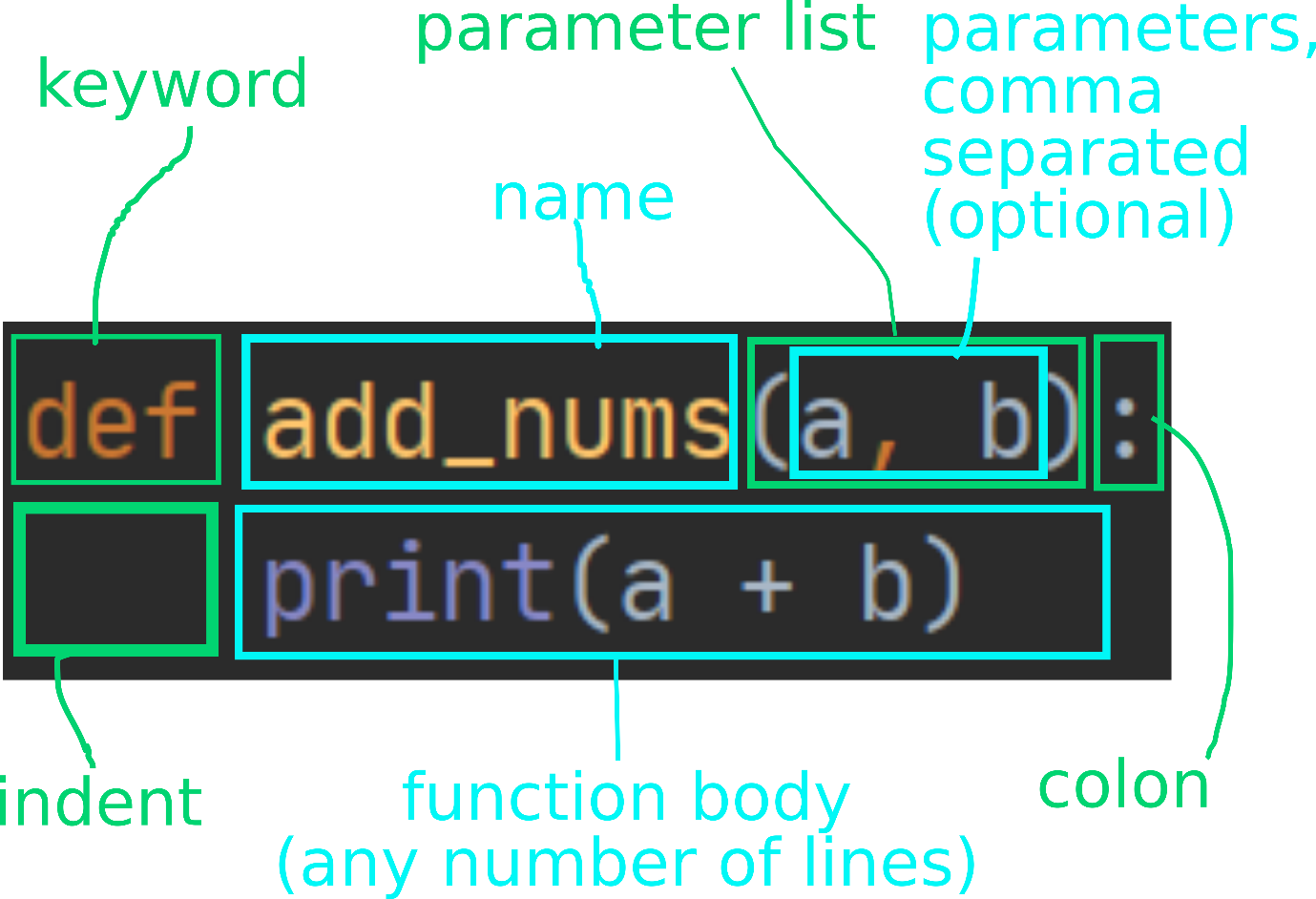
* First, the function must be **defined**. By this we mean the specification of what the function does.
* Once defined, the function can be **called**. By this we mean the invocation of the function i.e. getting the function to do whatever it does.

### Defining a function

To define a function we must provide:

* a name - must be a [valid Python identifier](https://docs.python.org/3/reference/lexical_analysis.html#identifiers)
* a parameter list - this may be empty
* a function body - this is code in one or more lines, indented in relation to the position of the function heading

The **body** of a function contains its behaviour. The **parameters** are placeholders for data that will be passed in when the function is used (called). Parameters allow the use of the function with different data, making it a very flexible code reuse mechanism.

Python function definition syntax 

### Calling a function

A function

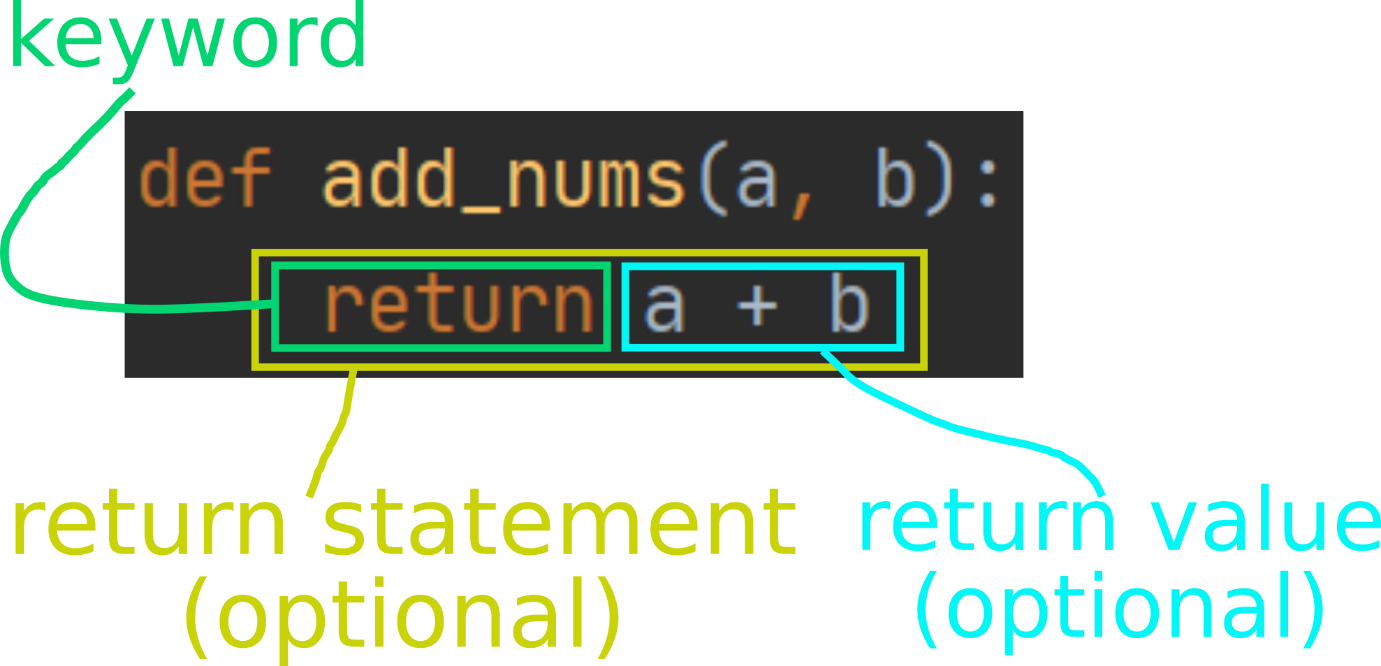
* can be called only if it has been defined
* is called using
  + the function name
  + arguments, which are values given to the function parameters for that call

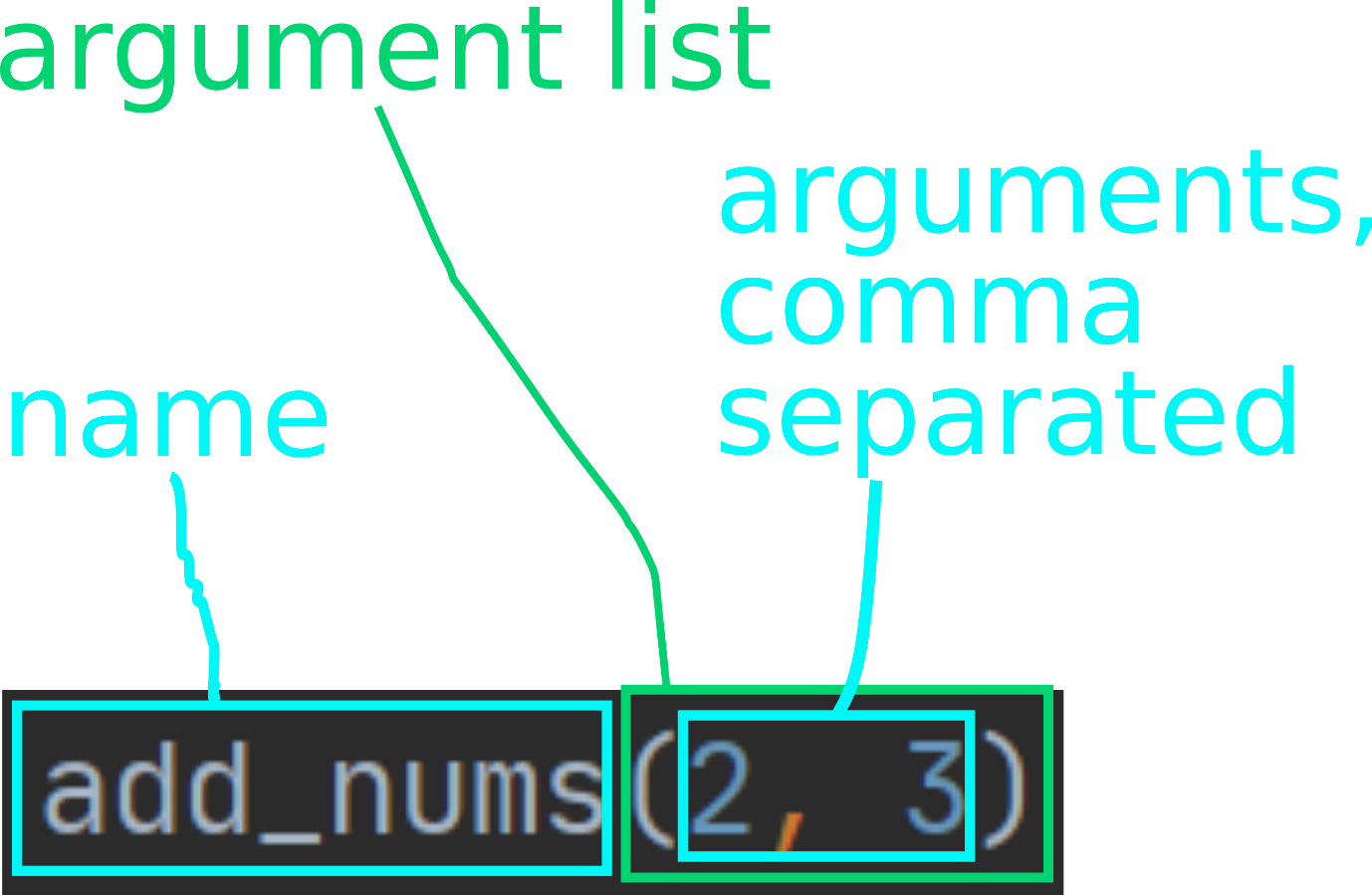
The **arguments** represent the concrete data values (in the form of any kind of valid expression) that are passed to the function to act on. In the function definition these data were represented by placeholders i.e. the parameters.

### The function return value

A function may also contain one or more **return statements**. A return statement

* ends the function execution
* can be accompanied by a **return value**, which is what the function **evaluates to** when it ends via this return statement
* a function that returns a value is like any other expression in that it can replace a value of that type in other code

Python function return statement 

Python function call syntax 

### Scope of variables and functions defined inside functions

* Any kind of code can be placed inside a function definition, including variable definitions and other function definitions.
* Being defined inside a function limits the scope of a variable or function. It is only visible within the function in which it is defined.

#### Example [CS-1] - Scope examples

**# variable definitions inside a function**

**x = 0**

**def printxy():**

**x = 3**

**print(x)**

**y = 4**

**print(y)**

**printxy() # expected output: 3 4**

**print(x) # expected output: 0**

**print(y) # error should be reported (y is not defined)**

**# nested functions**

**def print\_num(num):**

**def get\_message():**

**return "The number is: "**

**print(get\_message() + str(num))**

**print\_num(6) # expected output: The number is 6**

**print(get\_message() + 5) # error should be reported (get\_message is not defined)**

### More on function parameters and arguments

#### Parameter default values

Parameters can have default values. These are specified by using an assignment statement with the parameter in the function definition.

##### Example [CS-2] - Parameter default values

**def add\_nums(x, y=5, z=1):**

**return x + y + z**

Arguments can be left out for any number of contiguous parameters at the end of the parameter list, if those parameters have default values.

##### Example [CS-3] - Leaving out end arguments

**add\_nums(1, 2, 3) # expected output: 6**

**add\_nums(1, 2) # expected output: 4**

**add\_nums(1) # expected output: 7**

**add\_nums() # an error should be reported (x is not defined)**

#### Positional vs. named arguments

The arguments seen in function calls so far in this lesson are called **positional**. The interpreter knows how to match them up with the function parameters by position. In order to pass arguments out of order, we can use use the parameter names with the arguments, turning them into **named arguments**.

##### Example [CS-4] - Named arguments [

**add\_nums(z=3, x=1, y=2) # expected output: 6**

**add\_nums(1, z=3) # expected output: 9**

#### Allowing an arbitrary number of arguments

Python has a mechanism that allows a function to receive arguments unknown in number until runtime. The arbitrary-sized group of arguments is marked with a star (\*) character in front of the argument name. Only one such group is allowed in a function definition, otherwise the interpretation of the arguments would be ambiguous.

##### Example [CS-5] - Arbitrary number of arguments

**def concat\_strings(msg, \*list):**

**msgWithList = msg**

**for l in list:**

**msgWithList += l + ", "**

**if len(list) == 0:**

**msgWithList += "NONE :-("**

**return msgWithList.rstrip(" ,")**

**concat\_strings("My favourite letters: ", "x", "y", "z") # expected output: My favourite letters: x, y, z**

**concat\_strings("My favourite letters: ") # expected output: My favourite letters: NONE**

**names = ['jim', 'john', 'jane', 'jen']**

**concat\_strings("My favourite names: ", \*[ n.capitalize() for n in names ])**

**Note:** \* before the list in the last statement unpacks the list into individual arguments.

# WEEK 5 - Classes

## Classes

### What is a class?

In a programming language, a class:

* is a specific **blueprint** for objects
* represents the **type** of an object that is based on it as a blueprint
* bundles **data and associated behaviour** together
* is an object itself, of type type, and can be thought of as a 'meta-object' (an object that is the blueprint for other objects)

#### In Python

* Everything is an object and therefore has a type underlying it.
* What defines the type of an object is its class.
* Even classes have a class, type.

##### In-class Exercise [ CE-A]

Try the following expressions in the Python command line, or wrapped in a print() call in PyCharm.

**type(4)**

**type(int)**

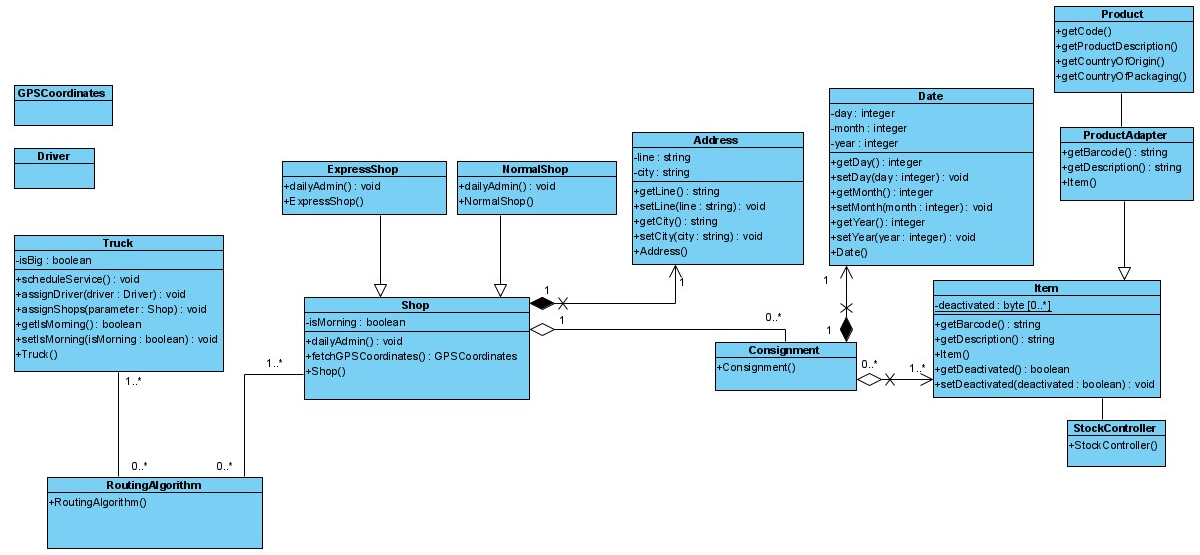
**type("abcd")**

**type(str)**

**type([2, 3, 4])**

**type(list)**

#### Why classes?

* With functions, we have seen one way of organising code to make it reusable, clearer and easier to maintain.
* Classes do the same, but they are essentially 'things' as opposed to functions, which represent 'behaviour'.
* Classes specific behaviour relevant to the thing that they are to be attached to them.
* Classes were designed to reflect reality, wherein we have things and associated behaviours. Consider the example in the picture below (it is a UML class diagram). Each box in the diagram shows a class. The top section contains the class name, the next section the class data and the bottom section the class methods (functions attached to a class). Example class diagram in UML 
* In any project but particularly in complex scenarios like the one shown in the picture *Example class diagram in UML* classes
  + break the problem down into domain concepts makes it easier to think about requirements
  + make design and implementation more intuitive
  + allow for easy division of labour in team projects
  + allow for easy reuse of concept implementations thanks to
    - the encapsulation of data
    - well defined interfaces for manipulating those data

### Defining a class in Python

We use the class construct to define a new object type.

In Python a class is defined using the following syntax:

#### Example [ CS-1] - A simple class defined in Python

class Simple:

x = "a"

y = "b"

def print(self):

print("x is: ", self.x)

print("y is: ", self.y)

|  |  |
| --- | --- |
| class | the keyword |
| **Simple** | user-defined name for the class (CamelCase by convention) |
| the colon (:) | mandatory part of the syntax |
| everything below first line | body of the class |
| x and y | class member variables |
| print() | a member function (method) |
| self | the keyword used within methods to refer to the object (not yet created at definition time) on which the method was called at runtime |

### Using a class to create objects

To create a class, we use the class name as a function (this will be explained later):

#### Example [ CS-2] - Creating an object from a class

**s = Simple()**

**s.print()**

|  |  |
| --- | --- |
| s = Simple() | creates an object of class Simple and assigns it to variable s |
| s.print() | uses dot-notation to run the method print of class Simple on object s |

### The constructor

* When an object is created, it is initialised with a call to a method called the **constructor**.
* If the constructor is not defined explicitly, the default constructor is called and this essentially does nothing.
* To add variables that will belong to the objects (instance member variables), we must define them in the constructor.

###### Example [CS-3] - The constructor method

**class Simple:**

**x = "a"**

**y = "b"**

**def \_\_init\_\_(self):**

**self.num1 = 1**

**self.num2 = 3**

**def print(self):**

**print("x is: ", self.x)**

**print("y is: ", self.y)**

**print("num1 is: ", self.num1)**

**print("num2 is: ", self.num2)**

|  |  |
| --- | --- |
| \_\_init\_\_ | the constructor method |
| num1 and num2 | instance member variables (pertaining to object) |

##### In-class Exercise [ CE-B]

Create a Python file in PyCharm called *simple.py* and copy in the code from figure [[CS-3]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_C_classes.html#py-constructor). Then add a line to create an object of type Simple and a line to call the print method. Run the code.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/simple1.py)

### Accessing and modifying member variables

Like member functions (methods), member variables are accessed using the dot notation.

###### Example [CS-4] - Reading and writing member variables

**s = Simple()**

**print("\n---- print variables individually from object ----")**

**print(s.x, s.y, s.num1, s.num2)**

**print("---- object before modification ----")**

**s.print()**

**s.x = "new a"**

**s.y = "new b"**

**s.num1 = 10**

**s.num2 = 30**

**print("---- object after modification ----")**

**s.print()**

##### In-class Exercise [CE-C]

Clear your *simple.py* file of all code bar the class definition from [[CS-3]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_C_classes.html#py-constructor). Add the statements from code snippet [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_C_classes.html#py-member-vars) and run it. Make sure you understand where the output is coming from.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/simple2.py)

#### Adding instance member variables

While this is **not recommended for the moment**, instance variables can be added to already created objects.

##### In-class Exercise [ CE-D]

Copy the following code into a Python script file. You will notice that the Simple class print method has been modified for easier addition of a printout heading and that it prints a (nonexistent) member variable z. Run the script.

**class Simple:**

**x = "a"**

**y = "b"**

**def \_\_init\_\_(self):**

**self.num1 = 1**

**self.num2 = 3**

**def print(self, heading):**

**print("----- {} -----".format(heading))**

**print("x is: ", self.x)**

**print("y is: ", self.y)**

**print("num1 is: ", self.num1)**

**print("num2 is: ", self.num2)**

**print("z is: ", self.z)**

**s1 = Simple()**

**s2 = Simple()**

**print("")**

**s1.z = "c"**

**s1.print("s1 after adding variable z to s1")**

**s2.print("s2 after adding variable z to s1")**

An error is reported. Why do you think this happened? Correct the code so that it runs without errors.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/simple3.py)

#### Class vs. instance member variables

There is a difference between the variables defined directly in the class and variables defined in the constructor:

* class variables, which are defined directly in the class, are shared by all objects
* each object has its own copy of a variable defined in the constructor

###### Example [CS-5] - Class and instance variables

**class Simple:**

**x = "a"**

**y = "b"**

**def \_\_init\_\_(self):**

**self.num1 = 1**

**self.num2 = 3**

|  |  |
| --- | --- |
| x and y | class variables (also called static variables), shared by all objects of the class, initialised at class definition time |
| num1 and num2 | instance variables, with a separate copy for each object, initialised in the constructor, at object creation time |

Let's have a look how access to these variables works.

###### Example [CS-6] - Class vs. instance variables (part 1)

**class Simplissima:**

**x = "a"**

**s1 = Simplissima()**

**s2 = Simplissima()**

**print(s1.x)**

**print(s2.x)**

What happens?

**class: x = "a"**

**object s1: NA object s2: NA**

**|| ||**

**|| ||**

**\/ \/**

**print(s1.x) print(s2.x)**

**own x not available own x not available**

**so get class x so get class x**

**output: a output: a**

Adding onto the code in [[CS-6]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_C_classes.html#py-class-vs-inst-1):

###### Example [CS-7] - Class vs. instance variables (part 2)

**s1.x = "hello"**

**s2.x = "world"**

**print(s1.x)**

**print(s2.x)**

What happens now?

**class: x = "a"**

**object s1: x = "hello" object s2: NA**

**|| ||**

**|| ||**

**\/ \/**

**print(s1.x) print(s2.x)**

**own x available own x available**

**so print that so print that**

**output: hello output: hello**

But can we still access the class (static) variable? Yes. We must scope the variable with the class name. Adding the following line onto the code in the previous two snippets will print 'a'.

###### Example [CS-8] - Class vs. instance variables (part 3)

print(Simplissima.x) # output: a

#### Class vs. instance variable summary

Class variables are useful for declaring constants that will not be modified e.g.pi=3.14.

Although sometimes they are used to set initial values for instance variables that will be modified later, this is not recommended as it obsucres the intention of the programmer. It is much better to initialise instance variables in the constructor e.g. **self.x = "hohoho".**

### Constructor parameters

Our example constructor from the previous section always initialises num1 and num2 to the same values. However, objects can be initialised with different values with the use of constructor parameters.

###### Example [CS-9] - Constructor parameters

**class Simple:**

**def \_\_init\_\_(self, n1, n2):**

**self.num1 = n1**

**self.num2 = n2**

**def print(self, heading=""):**

**print("")**

**if heading != "":**

**print("----- {} -----".format(heading))**

**print("num1 is: ", self.num1)**

**print("num2 is: ", self.num2)**

**s = Simple(3, 4)**

**s.print()**

Constructor parameters listed after the self keyword are treated like any other function parameters.

##### In-class Exercise [ CS-E]

Copy the code snippet [[CS-9]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_C_classes.html#py-constructor-params) into an empty Python script and run it to check if the outputs are as expected.

Then modify the script to include another parameter in the constructor and create an instance variable from its value. Modify the print method to include the new instance variable in the printout.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/simple4.py)

### Member naming convention

Some programming languages have constructs that allow controlling access to variables and methods in classes. In Python all variables and methods in a class are accessible (public). However, there is a convention in naming them that indicates to class users what the intention of the class developer was with respect to variable and method levels of access.

###### Example [ CS-10] - Naming convention for variables and methods

**class MixedUp:**

**def \_\_init\_\_(self)**

**self.x = "a" # public**

**self.\_y = "b" # protected**

**self.\_\_z = "c" # private**

|  |  |
| --- | --- |
| x (name with no underscores) | a public variable, intended for access by anybody |
| \_y (one leading underscore) | a protected variable, intended for access only by objects of the same class |
| \_\_z (two leading underscores) | a private variable, inteded for access only by the object itself |
| ***Note that the same rules apply to method names.*** | |

While nothing stops developers from not complying with the convention, doing this could be considered 'hacking', as it means not using someone else's code in the way it was intended.

### Getters, setters and deleters OPTIONAL

The constructs described in this section are presented for information only. The student is not expected to use them in this module.

###### Example [ CS-11] - Getter, setter and deleter example

**class WithProp:**

**def \_\_init\_\_(self):**

**print("calling constructor")**

**self.\_x = "a"**

**def getx(self):**

**print("calling getx")**

**return self.\_x**

**def setx(self, value):**

**print("calling setx")**

**self.\_x = value**

**def delx(self):**

**print("calling delx")**

**del self.\_x**

**x = property(getx, setx, delx)**

**def prhd(num):**

**print("--------- {} -----------".format(num))**

**prhd(1)**

**wp = WithProp()**

**prhd(2)**

**print(wp.x) # output: a**

**prhd(3)**

**wp.x = "b"**

**prhd(4)**

**print(wp.x) # output: b**

**prhd(5)**

**del wp.x**

**prhd(6)**

**print(wp.x) # error**

|  |  |
| --- | --- |
| x = property(getx, setx, delx) | defines x as an instance variable of class WithProp |
| getx, setx and delx | methods that wrap getting, setting and deleting, performing additional stuff (in this case printing "calling ...") |
| Notice that the getter, setter and delter must access x with a prepended underscore, which from the outside it is accessed a just x. | |

You might like to copy the code snippet [[CS-11]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_C_classes.html#py-get-set-del) into a script and experiment with it.

### Classes as types of member variables

Objects that are instances of classes defined as discussed in this lesson can be used like values of any other type. Thus they can be used as the values of object member variables.

In typed programming languages, the type of a member variable can be declared to be some class.

In Python we do not declare the type of a variable, but class methods can be written with the assumption that a particular variable will contain an object of a certain class.

###### Example [ CS-12] - A class as the type of a member variable value

**class Address:**

**def \_\_init\_\_(self, n, s, t):**

**self.number = n**

**self.street = s**

**self.town = t**

**class Person:**

**def \_\_init\_\_(self, n, s, a):**

**self.name = n**

**self.surname = s**

**self.address = a**

**def print(self):**

**print(self.name, self.surname)**

**print(self.address.number, self.address.street, self.address.town)**

In the example in code snippet [[CS-12]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_C_classes.html#py-class-as-member) the class Person includes a member variable called address. The print method assumes that address has the type Address and accesses the variables of that class.

##### In-class Exercise [ CE-F]

Copy the code in [[CS-12]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_C_classes.html#py-class-as-member) into a Python script, then edit it to create an object of class Address and an object of class Person (passing in the just created address). Print the person object.

Now create another person object but instead of an address of type Address pass in a string. What happens when you try to print the second person object? Why?

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/clsincls.py)

In the previous in-class exercise the Person class's print function accesses the Address class's mamber variables directly. This is not good practice. To get the most out of object-oriented programming a developer should strive towards encapsulation, exposing as little of a class's variables as possible. Do the next exercise to improve that code.

##### In-class Exercise [ CE-G]

Working with the same file as in the previous exercise, delete the second person and modify the two classes so that the Address class gets a print method and the Person class calls it, rather than accessing individual variables in the address object. Mark all the member variables in both classes as private.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/cicencapsulate.py)

### Importing modules

If we started defining many classes and placed them all in the same file things would start getting pretty unmanageable very quickly.

We can place our code in different files and import a file (or module) in another one that depends on its contents.

###### Example [ CS-13] - Module import

Contents of address.py:

**class Address:**

**# code defining the class**

Contents of person.py:

**from address import Address**

**class Person:**

**# code defining the class and using Address**

##### In-class Exercise [ CE-H]

Download [this file](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/import/import_set.py) onto your computer. Then reorganise it into three files called *address.py*, *person.py* and *main.py*. The first two should include a class definition only and the last on the code that creates Person objects and prints them. Remember that you must use imports.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/import.zip)

##### In-class Exercise [ CE-I]

Starting with the solution files for the previous exercise on imports, modify the person class to add a date of birth as a member variable of type datetime.

The class datetime is available in Python from module datetime (same name). Add a method to the Person class that calculates and returns the person's age. Write some code in file *main.py* to test your solution.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/age.zip)

### Collections of objects

Objects of any type can be in collections such as lists.

##### In-class Exercise [ CE-J]

Starting with the files given [here](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/team_set.zip), modify the Team class in file *team.py* with methods to add a person to the team, remove a person from the team and print a list filtered by role in team (or whole list if role is ""). You will notice that the class Person now has a member variable indicating the member's role ('lead', 'dev' or 'ops'). Write some code in file *main.py* to test all three methods.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/team.zip)

# WEEK 6 – Classes Ctd, Design Principles

## 23rd Oct 2024

## Classes Continued

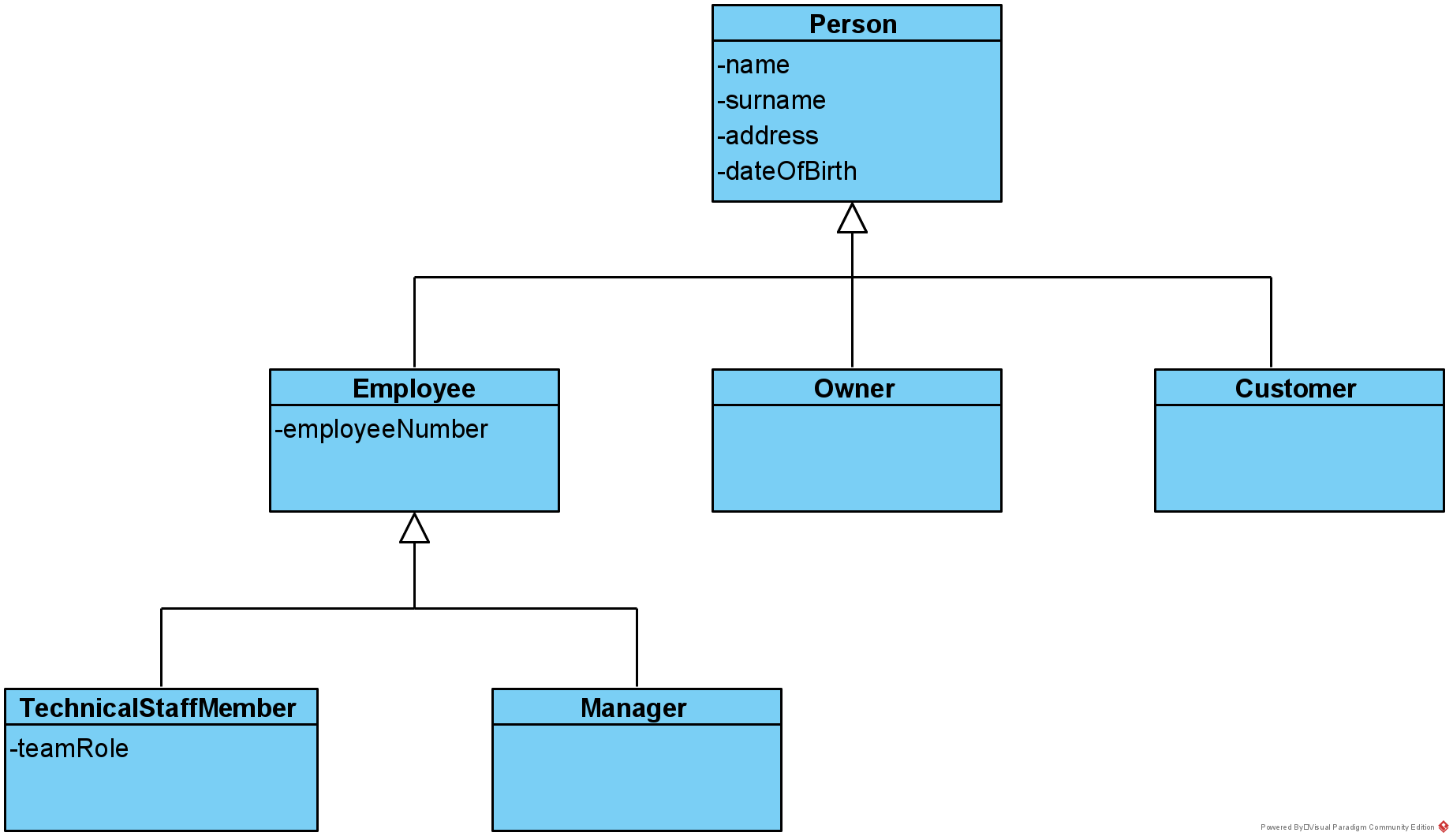
### Class Inheritance

Classes can be defined as subtypes of other classes. This concept is called **class inheritance**.

In an inheritance relationship

* one class is a **subclass** of another class, called the **superclass** class
* the superclass is always more **general** and the subclass more **specific**
* the subclass **inherits** the properties (variables and methods) of the superclass
* thus, an object of the subclass type **is** an object of the superclass type
* **in addition** to inherited properties, the subclass has properties that are specific to itself
* relationships can extend over many levels of an inheritance **hierarchy**
* synonymous terms
  + superclass, base class, parent class
  + subclass, derived class, child class

In the picture below, class Employee is a derived from class Person. It inherits properties name, surname, address and dateOfBirth. An object of type Employee has all these properties **and** its own property employeeNumber, which does not pertain to a person in general.

A class inheritance hierarchy 

### Inheritance syntax in Python

##### Example [CS-1] - Class inheritance with Python

class Person:

def \_\_init\_\_(self, n, s):

self.name = n

self.surname = s

def print(self):

print("")

print("name: ", self.name)

print("surname: ", self.surname)

def get\_fullname(self):

return self.name + " " + self.surname

class Employee(Person):

def \_\_init\_\_(self, n, s, en):

super().\_\_init\_\_(n, s)

self.employeeNumber = en

def print(self):

super().print()

print("employee number: ", self.employeeNumber)

|  |  |
| --- | --- |
| Class Employee(Person) | a new class called Employee is declared as a subclass of Person |
| super() | returns an object of the base class type |
| super().\_\_init\_\_() | base class constructor being called from the subclass to construct its inherited part |
| get\_fullname() | inherited and callable on the subclass |
| print() | on the subclass, this is called instead of the base class print method |

#### In class Exercise [ CE-A]

Copy the code from code snippet [[CS-1]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_D_classes.html#py-inheritance) into a script file. Then write some code to test the classes:

* create an object of type *Person*
* call the function *print* on it
* call the function *get\_fullname* on it
* repeat the previous steps with an object of type Employee
* think about the outputs until you are happy you understand how they were produced

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/inheritance.py)

#### In class Exercise [ CE-B]

Use the answer files for the team/person/address exercise (found [here](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/team.zip)) as the starting point for this exercise. Reorganise the code so that that it retains the functionality of the existing code, but with a new class Employee included. In the current setup, the Person class has a member variable for their team role, which really does not belong on Person, but could belong on Employee. The team should now consist of employees.

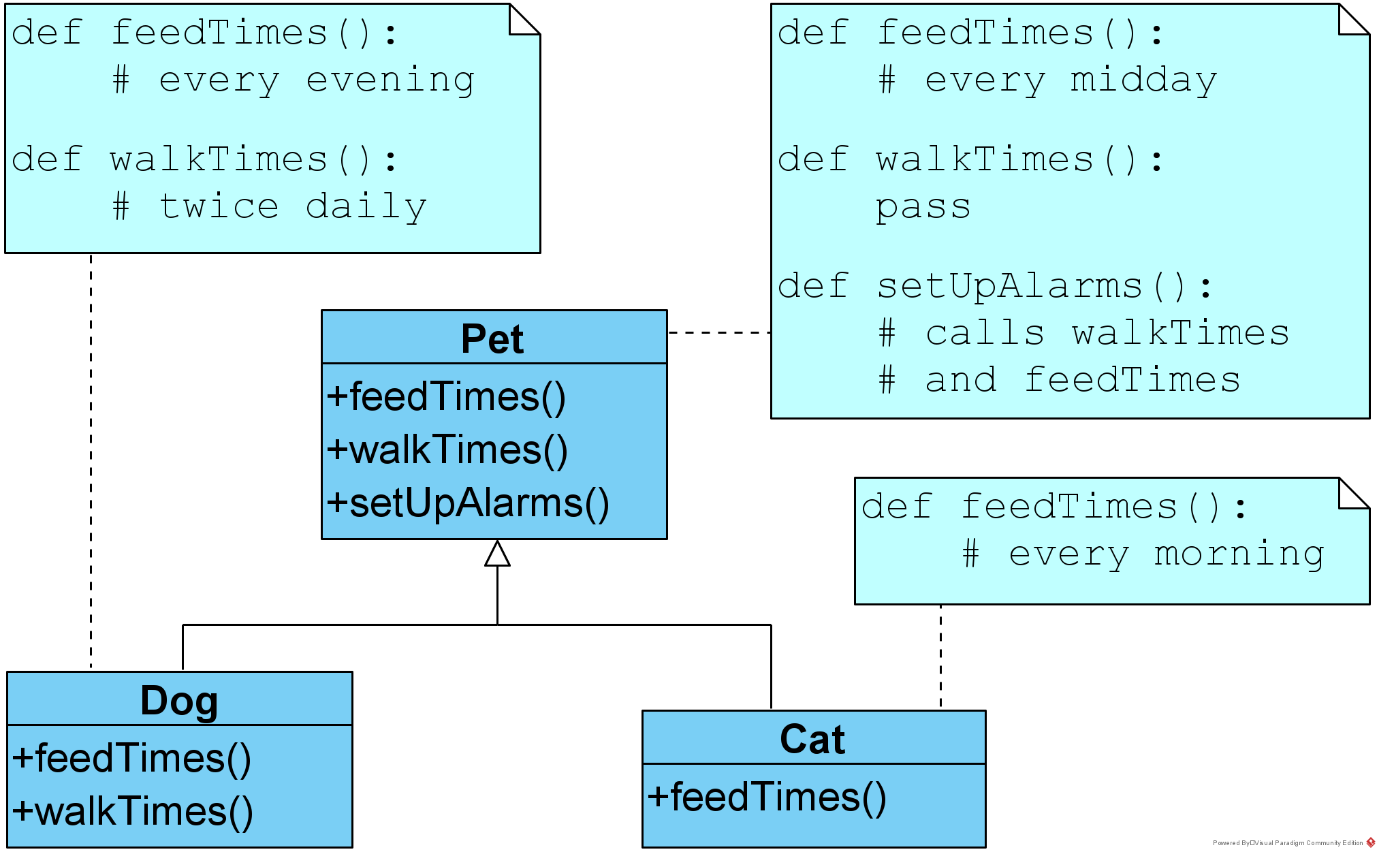
[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/team_reorg.zip)

### Method overriding

In code snippet [[CS-1]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_D_classes.html#py-inheritance) we have seen the subclass Employee define the function print, which has the same name and signature as the function print in the base class Person. This is called **method overriding** (or function overriding). As a result of overriding, the definition/execution mapping is as shown in the table.

If the base class has a function called func, then the following applies.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **object type** | **subclass has function func** | **func called on** | **executed** | **examples in *Pets class diagram*** |
| base | does not matter | base | base class func | petObj.walkTimes() petObj.feedTimes() petObj.setUpAlarms() |
| subclass | ✔ | subclass | subclass func | dogObj.feedTimes() dogObj.walkTimes() catObj.feedTimes() |
| subclass | X | subclass | base func | dogObj.setUpAlarms() catObj.walkTimes() cat\_obj.setUpAlarms() |
| subclass | ✔ | base | subclass func | feedTimes() called from dogObj.setUpAlarms() walkTimes() called from dogObj.setUpAlarms() feedTimes() called from catObj.setUpAlarms() |

Pet class diagram 

Implement the classes in *Pets class diagram* and test the four different scenarios that can happen when overriding/overridden functions are called. Place each class in a separate file and use a file called *main.py* for the test code.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/pets.zip)

**A couple of other things to note**

* Function overriding is at the heart of the **polymorphism** found in inheritance hierarchies. The word polymorphism roughly means "many shapes" and in object-oriented programming refers to how the same thing (the base class) can behave in many different ways (specific behaviour of subclasses, implemented in overriding functions).
* It is important for the overriding function to have the **same signature** (name and parameters) as the base class function. Using different overridden and overriding signatures can be made to work sometimes (or can work by chance) but this would be using polymorphism in an inintended way. Python is very permissive, which in many contexts makes it easy to work with, but this should not be abused by the programmer.

### Car example from class

[Car example from class](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/classes/car_with_owner.zip)

## Coding Practice

### Simple program design steps

#### Simple program design steps (roughly!)

1. Identify inputs and outputs
2. Identify the order of inputs and outputs
3. Identify which inputs affect which outputs
4. Design the flow of control
5. Identify repetitions and variables
6. If necessary, go back to step 4
7. Fill the programmatic details
8. After implementation and testing, fix any imperfections

## Code quality attributes

The first goal when learning to program is to get it right, in the sense of fulfilling any **functional requirements** with the code.

However, the **non-functional quality attributes** of a program are just as important. Quantitative measures of these are used to set quality attribute (non-functional) requirements for programs. Some of these attributes are:

* **usability**: the user-friendliness of the program

Does the program communicate fully with the user and in a way that is easy to understand?

* **understandability**: the clarity of the code for other coders

How easy would it be for someone else to review the code or take over its development?

* **maintainability**: the ease with which the code once in production can be modified

Can bug fixes or extensions be applied and tested easily and quickly?

* **reusability**: the possibility of reusing the code if needed

Can the code easily be utilised in a different context?

* **optimality**: as low as possible use of resources

Is the program written in a way that does not waste memory and/or CPU?)

* **environmental friendliness**: similar to above but with focus on environmental concerns
* **scalability**: the possibility of adjustment for greater processing bulk

Can this code be deployed so as to serve more users or handle a bigger database?

* **configurability**: changeability of high-level properties

Can the program easily be used in different ways or for different purposes that are all equally useful?

The list goes on, but mentioned here are only those that we will consider in our module.

## An example

Let us start with a function that prints out an instruction for volunteers giving out t-shirts at an event. The function prints out the size and colour that needs to be handed out, given the age and height of an event-goer. It is shown in code snippet [[CS-1]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_6b_design.html#py-quality1).

#### Example [CS-1] - Code quality example - version 1

**def print\_instruction(height, age):**

**if height >= 180 and age >= 18:**

**print("This person needs a dark t-shirt, size large.")**

**if height < 180 and height >= 150 and age >= 18:**

**print("This person needs a dark t-shirt, size medium.")**

**if height < 150 and age >= 18:**

**print("This person needs a dark t-shirt, size small.")**

**if height >= 180 and age < 18:**

**print("This person needs a bright t-shirt, size large.")**

**if height < 180 and height >= 150 and age < 18:**

**print("This person needs a bright t-shirt, size medium.")**

**if height < 150 and age < 18:**

**print("This person needs a bright t-shirt, size small.")**

The Python function above is **logically correct**, fulfilling its functional requirements, which are

* to choose a t-shirt size based on height (with particular height boundaries for sizes small, medium and large)
* to choose a t-shirt colour based on age (one colour for adults and another for young people)
* to print a message with the size and colour information

However, there are various problems with this function, which we will examine and correct one by one.

### Code that causes superfluous execution

This is bad practice because:

* it is a drain on CPU resources
* it confuses the reader by obscuring the code purpose

#### Instances in the example:

1. There are six *disjunctive* (non-overlapping) combinations of age and height. Each is checked for with a new **if-**statement, meaning that even if- an if-statement condition is true and its block is executed, further if-statements are also executed, unnecessarily, since they definitely cannot be true due to the disjunctiveness. An **if-elif** chain should be used here.
2. The six combinations are also *exhaustive* (are all the combinations that exist), meaning that the last **elif** condition check is superfluous. An **else** clause should be used here instead of the **elif** one.
3. Finally, the combinations are *contiguous* (share boundaries), which means that a check for a value's position with respect to a boundary need to be carried out only once (the check height **< 180** in the first **elif** clause is superfluous because the if clause has the check height **>== 180** etc.). These superfluous checks can be left out. We have two height boundaries (180 and 150) and one ege boundary (18). The code, when running, should carry out at most 3 checks.

The unnecessary if-statements cause the processing power needed for this function to double on average, affecting *optimality* and *environmental friendliness*. The other two problems also have an efficiency reduction effect, if a lesser one. All the unnecessary conditionals affect the *understandability* and *maintainability* because the functional intention is not expressed clearly by the code.

Version 2 of the function, including the three improvements discussed above, is shown in code snippet [[CS-2]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_6b_design.html#py-quality2). Notice that even though five condition checks are found in the code, at run time no more than three will be executed, which is the minimum possible for this problem.

#### Example [CS-2] - Code quality example

**def print\_instruction(height, age):**

**if age >= 18:**

**if height >= 180:**

**print("This person needs a dark t-shirt, size large.")**

**elif height >= 150:**

**print("This person needs a dark t-shirt, size medium.")**

**elif:**

**print("This person needs a dark t-shirt, size small.")**

**else:**

**if height >= 180:**

**print("This person needs a bright t-shirt, size large.")**

**elif height >= 150:**

**print("This person needs a bright t-shirt, size medium.")**

**else:**

**print("This person needs a bright t-shirt, size small.")**

### Repetition of code

This is bad practice because:

* it is a drain on storage and network resources
* it is hard to maintain, given that many-target changes are prone to error

#### Instances in the example:

1. In the example the print call is almost identical in the six conditional blocks. There are a couple of words that are different, but this can be easily parameterised.
2. The height boundary checks are written into the code twice. There will be cases where this is necessary, if there is a dependency between the two variables that are being checked (age and height). However, in our example there is no dependency (the height and age categories are always the same, regardless of the other variable's value), which means that we can categorise the age and height separately and avoid this repetition.

Repetition of code affects the code's *maintainability*, because of the multiple places in which changes would have to be made to modify it. Again, *understandability* can be affected simply because there is more to read in the listing. Repetition means more text in the program files and the unnecessary use of storage space (which in scenarios such as a mobile application can be scarce) and network resources (e.g. when client-side code of a web application is being transferred over the Internet).

Version 3 of the example function, with repetitions removed, is shown in code snippet [[CS-3]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_6b_design.html#py-quality3).

#### Example [CS-3] - Code quality example

**def write\_message\_to\_supplier(height, age):**

**if age > 18:**

**colour="dark"**

**else:**

**colour="bright"**

**if height > 180:**

**size="small"**

**elif height > 150:**

**size="medium"**

**else:**

**size="small"**

**print("This person needs a {} tee-shirt, size {}.".format(colour, size))**

### Hard-coded constants

This is bad practice because:

* it makes code harder to maintain, with literals that might need to change embeded deep in the program
* it prohibits program configuration in terms of the hard-coded constants
* literals are much harder to interpret than named constants

#### Instances in the example:

1. The age and height boundaries can be declared with the use of variables at the beginning of the program.

The final version (4) of the example function is shown in code snippet [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_6b_design.html#py-quality4).

#### Example [CS-4] - Code quality example

**LM\_SIZE\_BOUNDARY = 180**

**MS\_SIZE\_BOUNDARY = 150**

**AY\_AGE\_BOUNDARY = 18**

**def write\_message\_to\_supplier(height, age):**

**if age > LM\_SIZE\_BOUNDARY:**

**colour="dark"**

**else:**

**colour="bright"**

**if height > LM\_SIZE\_BOUNDARY:**

**size="small"**

**elif height > MS\_SIZE\_BOUNDARY:**

**size="medium"**

**else:**

**size="small"**

**print("This person needs a {} tee-shirt, size {}.".format(colour, size))**

# Week 7 – Review week

# Week 8 – Exceptions, Design and Principles

## Exceptions

### What is an exception?

* Literally, an [exception](https://docs.python.org/3/tutorial/errors.html) is a class (Exception) that represents a **runtime error**.
* Such errors typically make it impossible for the program to fulfill its intended purpose and prevent further execution.
* The exception class in Python has many [built-in subclasses](https://docs.python.org/3/library/exceptions.html) that represent different types of errors.
* Examples of such errors are
  + when the + operator is used between a string and a number (TypeError)
  + when the access is attempted of a list element at a non-existent index (IndexError)
  + when a non-existent variable is referenced (NameError)

#### In class exercise [CE-A]

For each of the three error types above, write some code that causes it.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/exceptions/examples.py)

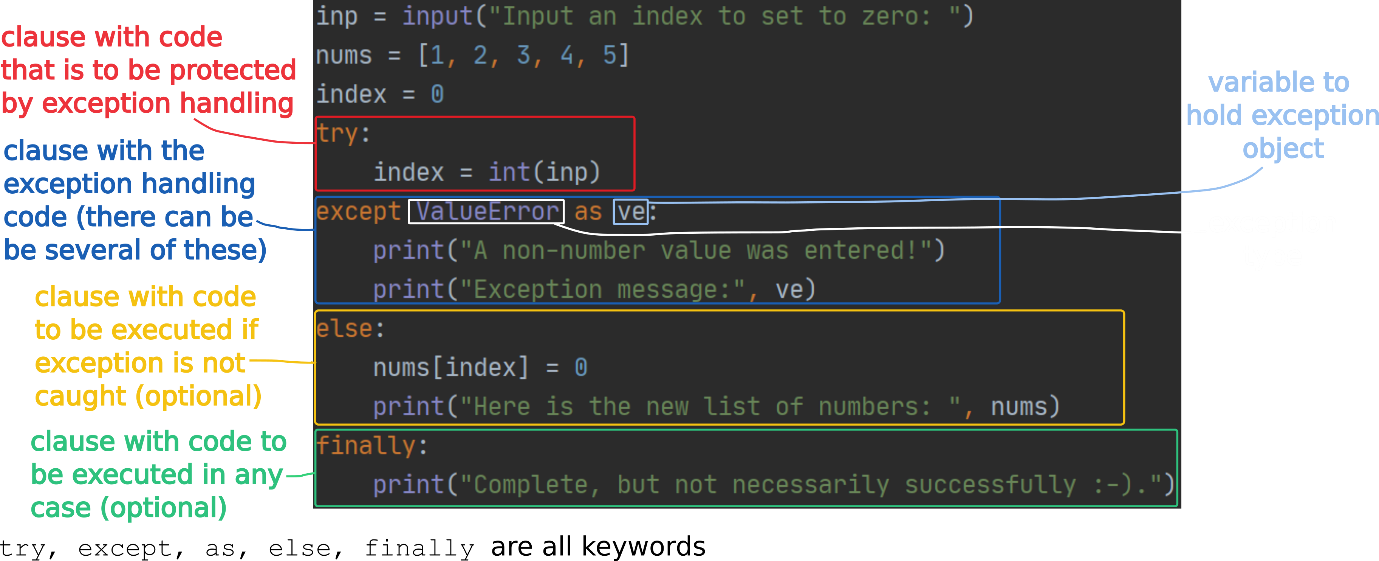
**How do exceptions work?**

* When a runtime error occurs, an **exception may be raised** (or *thrown*) using a special language construct: *raise an\_exception\_object*
* If the program **does not handle the exception**
  + the current execution frame (e.g. current function) is terminated
  + the exception **is propagated** to the next higher execution frame (e.g. another function that called the current function)
  + that next higher execution frame is terminated
  + the exception is propagated to the following higher execution frame
  + this is repeated until the exception 'reaches the surface' and is presented to the user with any information contained in the exception object when it was raised.

We saw this happening with the three examples earlier on. Those exceptions were raised by built-in Python functionality.

#### In class exercise [CE-B]

Run the Python script found [here](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/exceptions/abcd.py) (***abcd.py***) in the PyCharm debugger, with a breakpoint on the line where the exception is raised with *raise*. Look at the Frames portion of the Debugger panel inside the Debug view in PyCharm. This shows the functions currently on the call stack (in our case, main, a, b, c and d). Step over the line with the raise statement. You will see how the *print* calls at the end of each function are never made, because the exception is propagated through the frames in the call stack, terminating the execution of each function and surfacing to be reported in the console. The whole journey of the exception can be seen in the stack trace.

* However, **exceptions can be handled**. This is done with the try/except/else/finally construct: Exception handling construct 

#### In class Exercise [CE-C]

Copy the following code into a file and run it three times. Once enter a correct index, once an index that is out of range and once a non-number value. Explain what happens.

#### Example [CS-1] - try/except/else/finally

**inp = input("Input an index to set to zero: ")**

**nums = [1, 2, 3, 4, 5]**

**index = 0**

**try:**

**index = int(inp)**

**except ValueError as ve:**

**print("A non-number value was entered!")**

**print("Exception message:", ve)**

**else:**

**nums[index] = 0**

**print("Here is the new list of numbers: ", nums)**

**finally:**

**print("Complete, but not necessarily successfully :-).")**

An unhandled exception is raised when the index is out of bounds. Correct the code to handle this exception as well, by adding an *except* clause that handles exceptions of type *IndexError.* (*secondx.py*)

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/exceptions/secondex.py)

* + Note that all the clauses in the Python exception handling statement constitute the same scope. Any variables defined in clauses above will be accessible to the clauses below.
  + While the else clause is optional, it is a good idea to make use of it for readability and the case that additional exceptions are raised.
  + The finally clause is also optional, but it should be used if clean-up or other work using the same scope as the try clause needs to be completed regardless of the outcome.
* Exceptions can be explicitly raised in a program.

#### Example [CS-2] - Exception raising syntax

**raise object\_of\_some\_error\_class\_type**

|  |  |
| --- | --- |
| raise | keyword for raising exceptions |
| *object\_of\_some\_error\_class\_type* | an object that for its type has a subclass of Exception |

In the following example, an exception of type Exception is raised.

#### Example [CS-3] - Raising an exception

**raise Exception ("Something went wrong!)**

|  |
| --- |
| Exceptions take one or two arguments. The first argument is always a message. |

* New exception types can be defined

#### Example [CS-4] - Defining a new exception

**class ProblemWithAge(Exception):**

**pass**

In most cases, it is sufficient to define an error type, then use the Exception class message that is already implemented.

### Using exceptions

Exceptions as an error handling mechanism

* are, in most cases, easier to write into the code than many nested conditional statements
* make code a lot more readable
* allow for error reporting and handling to be concentrated rather than strewn around the code

#### In-class exercise [CE-D]

The code in example [[CS-5]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_E_exceptions.html#py-error-handling-with-conditionals) is relatively simple and yet suffers from lack of clarity as to what functionality it implements.

#### Example [CS-5] - Error handling without exceptions

**ageStr = input("Please input your age: ")**

**if not ageStr.isdigit():**

**print("ERROR: Age is invalid.")**

**else:**

**age = int(ageStr)**

**if age < 18:**

**print("ERROR: You are not old enough to use this website.")**

**else:**

**quantityStr = input("How many bottles of craft beer do you want to buy? ")**

**if not quantityStr.isdigit():**

**print("ERROR: Quantity is invalid.")**

**else:**

**quantity = int(quantityStr)**

**payment = input("Would you like to pay with Revolut or credit card? ")**

**if payment != 'Revolut' and payment != 'credit card':**

**print("ERROR: Invalid payment type.")**

**else:**

**print("Thank you for your order!")**

Rewrite the code above to use exceptions, defining exception classes where needed.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/exceptions/converted.py)

Exceptions **should not** be used to implement conditional behaviour that is part of the normal, error-free functioning of the program. The following example shows some logically correct code that employs exception handling to provide two different execution paths that are taken depending on a condition. This is not good practice.

#### **Example [ CS-6] - BAD PRACTICE EXAMPLE:** Exceptions in error-free functionality

**# if the entered op is 1, the list is sorted**

**# otherwise, the list items are converted to uppercase**

**l = ['apple', 'banana', 'cherry', 'orange', 'peach', 'apricot', 'raspberry']**

**op = int(input("Enter op number: "))**

**try:**

**if op == 1:**

**print(sorted(l))**

**else:**

**raise Exception()**

**except Exception:**

**print(list(map(lambda x: x.upper(), l)))**

## Design and principles

### Simple program design steps (roughly!)

1. Identify inputs and outputs
2. Identify the order of inputs and outputs
3. Identify which inputs affect which outputs
4. Design the flow of control
5. Identify repetitions and variables
6. If necessary, go back to step 4
7. Fill the programmatic details
8. After implementation and testing, fix any imperfections

### Code quality attributes

The first goal when learning to program is to get it right, in the sense of fulfilling any **functional requirements** with the code.

However, the **non-functional quality attributes** of a program are just as important. Quantitative measures of these are used to set quality attribute (non-functional) requirements for programs. Some of these attributes are:

* **usability**: the user-friendliness of the program

Does the program communicate fully with the user and in a way that is easy to understand?

* **understandability**: the clarity of the code for other coders

How easy would it be for someone else to review the code or take over its development?

* **maintainability**: the ease with which the code once in production can be modified

Can bug fixes or extensions be applied and tested easily and quickly?

* **reusability**: the possibility of reusing the code if needed

Can the code easily be utilised in a different context?

* **optimality**: as low as possible use of resources

Is the program written in a way that does not waste memory and/or CPU?)

* **environmental friendliness**: similar to above but with focus on environmental concerns
* **scalability**: the possibility of adjustment for greater processing bulk

Can this code be deployed so as to serve more users or handle a bigger database?

* **configurability**: changeability of high-level properties

Can the program easily be used in different ways or for different purposes that are all equally useful?

The list goes on, but mentioned here are only those that we will consider in our module.

### An example

Let us start with a function that prints out an instruction for volunteers giving out t-shirts at an event. The function prints out the size and colour that needs to be handed out, given the age and height of an event-goer. It is shown in code snippet [[CS-1]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_6c_design.html#py-quality1).

##### Exercise [CS-1] - Code quality example - version 1

def print\_instruction(height, age):

if height >= 180 and age >= 18:

print("This person needs a dark t-shirt, size large.")

if height < 180 and height >= 150 and age >= 18:

print("This person needs a dark t-shirt, size medium.")

if height < 150 and age >= 18:

print("This person needs a dark t-shirt, size small.")

if height >= 180 and age < 18:

print("This person needs a bright t-shirt, size large.")

if height < 180 and height >= 150 and age < 18:

print("This person needs a bright t-shirt, size medium.")

if height < 150 and age < 18:

print("This person needs a bright t-shirt, size small.")

The Python function above is **logically correct**, fulfilling its functional requirements, which are

* to choose a t-shirt size based on height (with particular height boundaries for sizes small, medium and large)
* to choose a t-shirt colour based on age (one colour for adults and another for young people)
* to print a message with the size and colour information

However, there are various problems with this function, which we will examine and correct one by one.

#### Code that causes superfluous execution

This is bad practice because:

* it is a drain on CPU resources
* it confuses the reader by obscuring the code purpose

Instances in the example:

1. There are six *disjunctive* (non-overlapping) combinations of age and height. Each is checked for with a new if-statement, meaning that even if- an **if-**statement condition is true and its block is executed, further **if-**statements are also executed, unnecessarily, since they definitely cannot be true due to the disjunctiveness. An ***if-elif*** chain should be used here.
2. The six combinations are also *exhaustive* (are all the combinations that exist), meaning that the last **elif** condition check is superfluous. An **else** clause should be used here instead of the **elif** one.
3. Finally, the combinations are *contiguous* (share boundaries), which means that a check for a value's position with respect to a boundary need to be carried out only once (the check height **< 180** in the first **elif** clause is superfluous because the if clause has the check height **>== 180** etc.). These superfluous checks can be left out. We have two height boundaries (180 and 150) and one ege boundary (18). The code, when running, should carry out at most 3 checks.

The unnecessary **if-**statements cause the processing power needed for this function to double on average, affecting *optimality* and *environmental friendliness*. The other two problems also have an efficiency reduction effect, if a lesser one. All the unnecessary conditionals affect the *understandability* and *maintainability* because the functional intention is not expressed clearly by the code.

Version 2 of the function, including the three improvements discussed above, is shown in code snippet [[CS-2]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_6c_design.html#py-quality2). Notice that even though five condition checks are found in the code, at run time no more than three will be executed, which is the minimum possible for this problem.

##### Example [ CS-2] - Code quality example - version 2

def print\_instruction(height, age):

if age >= 18:

if height >= 180:

print("This person needs a dark t-shirt, size large.")

elif height >= 150:

print("This person needs a dark t-shirt, size medium.")

elif:

print("This person needs a dark t-shirt, size small.")

else:

if height >= 180:

print("This person needs a bright t-shirt, size large.")

elif height >= 150:

print("This person needs a bright t-shirt, size medium.")

else:

print("This person needs a bright t-shirt, size small.")

#### Repetition of code

This is bad practice because:

* it is a drain on storage and network resources
* it is hard to maintain, given that many-target changes are prone to error

Instances in the example:

1. In the example the **print** call is almost identical in the six conditional blocks. There are a couple of words that are different, but this can be easily parameterised.
2. The height boundary checks are written into the code twice. There will be cases where this is necessary, if there is a dependency between the two variables that are being checked (age and height). However, in our example there is no dependency (the height and age categories are always the same, regardless of the other variable's value), which means that we can categorise the age and height separately and avoid this repetition.

Repetition of code affects the code's *maintainability*, because of the multiple places in which changes would have to be made to modify it. Again, *understandability* can be affected simply because there is more to read in the listing. Repetition means more text in the program files and the unnecessary use of storage space (which in scenarios such as a mobile application can be scarce) and network resources (e.g. when client-side code of a web application is being transferred over the Internet).

Version 3 of the example function, with repetitions removed, is shown in code snippet [[CS-3]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_6c_design.html#py-quality3).

##### Example [CS-3] - Code quality example - version 3

def write\_message\_to\_supplier(height, age):

if age >= 18:

colour="dark"

else:

colour="bright"

if height >= 180:

size="small"

elif height > 150:

size="medium"

else:

size="small"

print("This person needs a {} tee-shirt, size {}.".format(colour, size))

#### Hard-coded constants

This is bad practice because:

* it makes code harder to maintain, with literals that might need to change embeded deep in the program
* it prohibits program configuration in terms of the hard-coded constants
* literals are much harder to interpret than named constants

Instances in the example:

1. The age and height boundaries can be declared with the use of variables at the beginning of the program.

The final version (4) of the example function is shown in code snippet [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_6c_design.html#py-quality4).

##### Example [CS-4Code quality example - version 4

LM\_SIZE\_BOUNDARY = 180

MS\_SIZE\_BOUNDARY = 150

AY\_AGE\_BOUNDARY = 18

def write\_message\_to\_supplier(height, age):

if age > LM\_SIZE\_BOUNDARY:

colour="dark"

else:

colour="bright"

if height > LM\_SIZE\_BOUNDARY:

size="small"

elif height > MS\_SIZE\_BOUNDARY:

size="medium"

else:

size="small"

print("This person needs a {} tee-shirt, size {}.".format(colour, size))

### Coding principles

* Books, articles and other sources of advice on good programming practice abound in **software design principles**.
* These principles all have the aim of helping the software developer comply with common quality requirements as a matter of course, rather than in a separate add-on effort.
* We will look at some that are widely accepted as standard for good program design.

#### The principles

* DRY (Don't Repeat Yourself)
* KISS (Keep it Short and Simple)
* YAGNI (You Ain't Gonna Need It) (don’t build it if you don’t’ need it now – related to agifle)
* SOLID
  + Single responsibility principle
  + Open/closed principle
  + Liskov substitution principle
  + Interface segregation principle
  + Dependency inversion principle

# Week 9 - Files

<https://mockaroo.com/> - Mockaroo - Random Data Generator and API Mocking Tool | JSON / CSV / SQL / Excel

## Reading from text files

* To read from a file, we must first **open** it.
* After reading, we must close the file to free it for any other users or programs that might want to use it.
* A text file can be read in different ways:
  + whole contents in one go
  + line-by-line
  + any number of characters at a time

#### Example [CS-1] Reading a text file in one go

file = open("test.txt", "r")

contents = file.read()

file.close()

|  |  |
| --- | --- |
| file | a file object returned from the call to the function open |
| "test.txt" | the name of the file to be opened |
| "r" | the open mode (see [here](https://docs.python.org/3/library/functions.html#open) for full list) |
| read() | the file method that reads the entire file in one go |
| close() | the file method frees the file resource |

Create a file called *test.txt* in your PyCharm project directory and write a few lines of text into it. Then copy code snippet [[CS-1]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_F_files.html#py-read-txt-file) into a script file. At the end add an instruction to print contents. Run to see if the printed contents matches what was in the file.

#### Example [CS-2] - Reading a text file line-by-line

# reading one line

file = open("test.txt", "r")

print(file.readline())

file.close()

# reading all lines in a while loop

file = open("test.txt", "r")

l = file.readline()

while l:

print(l)

l = file.readline()

file.close()

# reading all lines in a for loop

file = open("test.txt", "r")

for l in file:

print(l)

file.close()

* Notice that the newline character is read out of the file, just like any other character.
* Once the contents have been exhausted, read methods return an empty string. The file must be reopened to be read from the beginning again.

#### Example [ CS-3] - Reading arbitrary numbers of characters from a text file

# reading characters

file = open("test.txt", "r")

print(file.read(1))

print(file.read(2))

print(file.read(3))

file.close()

Copy the code from [[CS-2]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_F_files.html#py-read-txt-file-lines) and then from [[CS-3]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_F_files.html#py-read-txt-file-chars) into a Python script file and try out the different ways of reading the file.

Create a file with the following contents. Then write a script that reads the file into a list of lists, stripping the double quotes from the individual values and the newline characters from the end of each line.

"Annie","Apple","2","TheOrchard","Atown","AA","ABC789",2010,1,2,"dev"

"Bob","Builder","1","Site","Btown","BB","XYZ123",2000,1,1,"ops"

"Charlie","Chime","5","TheHill","Ctown","CC","MNO000",1990,3,3,"dev"

"Danny","Doom","4","TheCave","Dtown","DD","@!#999",1900,12,31,"lead"

"Emmy","Eel","5","TheLake","Etown","EE","LKE0000",1995,5,5,"dev"

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/files/read_messy.py)

## Writing to a file

* To write to a file, we must open the file and explicitly specify write or append mode (read mode is the default).
* The file must be closed after writing. Like with reading, a file must be freed if it is to be used by other programs.

#### Example [ CS-4] - Writing to a text file

file = open("test.txt", "w")

file.write("A new line!")

file.close()

Try the code in [[CS-4]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_F_files.html#py-write-txt-file) and check the file to see the result. Now repeat the exercise but opening the file in append mode. What is the result now?(

Qs: what was the result… get error?

## Reading and writing CSV files

#### Example [CS-5] - Reading a CSV file

import csv

personFile = open("people.csv")

csvReader = csv.reader(personFile, delimiter=",")

for row in csvReader:

print(row)

personFile.close()

Use the code in [[CS-5]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_F_files.html#py-read-csv) to read [this file](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/files/people.csv).

#### Example [CS-6] - Writing a CSV file

import csv

personFile = open("people.csv", "a")

csvWriter = csv.writer(personFile, delimiter=",")

csvWriter.writerow(["Kit", "Kat", "X00777777", "Stonebatter", "D7"])

personFile.close()

Use the code from [[CS-6]](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/lec_F_files.html#py-write-csv) to write an additional line to the CSV file.

## The with statement

#### Example [CS-7] - Using the with statement

with open("people.txt", "r") as file:

list\_of\_lists = []

for line in file:

list\_of\_lists.append(list(map(lambda s: s.strip('"\n'), line.split(","))))

print(list\_of\_lists)

The with statement cleans up after itself, which means that we do not need to close a file opened as part of this statement.

## Generating data

Data can be obtained from websites like [this one](https://www.mockaroo.com/).

# Week 10 – Containers

## Python built-in container types

There are four types of built-in container types in Python:

* lists
* tuples
* sets
* dictionaries

So far in the module we have used the first two. In this session we will look at **sets** and **dictionaries**.

* The Python [**set**](https://docs.python.org/3/library/stdtypes.html#set) container contains unique elements (like in maths). It can be used, among other things, for **testing membership** and **identifying unique values in lists**.

### In class Exercise – [CE-A] - Create a new Python script and in it save the following list into a variable:

[1, 1, 2, 23, 1, 3, 6, 7, 7, 8, 23, 1, 1, 9, 2, 22, 34]

Write some code that uses a set container to count the number of unique values in the list.

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/set/unique.py)

* The Python [**dict**](https://docs.python.org/3/library/stdtypes.html#dict) container is a mapping object, which maps unique keys to arbitrary objects.

### In class exercise - [CE B]

Use the [shop code](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/lab_solutions/a5.zip) from a previous assignment and modify it to use a dictionary instead of a list for the list of products. Which version is better and why?

[EXAMPLE SOLUTION](https://jelena-vk-itt.github.io/jvk-tudt-notes/ooswd/res/files/python/dict/shop_with_dict.zip)

* Python has additional container data types available in the [**collections**](https://docs.python.org/3/library/collections.html) package (which must be included in code with an import statement).