Python Programming for Data Science

Week 40, Monday

Object oriented programming

Regular expressions - popquiz

- What does this regular expression match: ([a-z0-9.%+-]+)@([a-z0-9.-]+)\.([a-z]{2,4})
- How many groups are there in the expression
- What values do these groups correspond to?
- How would we match and extract such information from Python?

Objects and Classes

• Every value in Python is an object. (e.g. "hello", 5, [10,20])

- Every value in Python is an object. (e.g. "hello", 5, [10,20])
- An object has attributes, accessed using the . operator.

```
1 = [1,2,3,4]
1.reverse() # reverse is a method attribute of list
```

- Every value in Python is an object. (e.g. "hello", 5, [10,20])
- An object has attributes, accessed using the . operator.

```
1 = [1,2,3,4]
1.reverse() # reverse is a method attribute of list
```

In short: An object is "a bundle" of data and functionality.

We've seen several built-in types: int, float, list, ...

We've seen several built-in types: int, float, list, ...

Q: What does it mean that an object has a certain type?

We've seen several built-in types: int, float, list, ...

Q: What does it mean that an object has a certain type?

A: The type defines which attributes the object has.

We've seen several built-in types: int, float, list, ...

Q: What does it mean that an object has a certain type?

A: The type defines which attributes the object has.

A different perspective:

We've seen several built-in types: int, float, list, ...

Q: What does it mean that an object has a certain type?

A: The type defines which attributes the object has.

A different perspective:

 A type corresponds to a template for the creation of objects.

We've seen several built-in types: int, float, list, ...

Q: What does it mean that an object has a certain type?

A: The type defines which attributes the object has.

A different perspective:

- A type corresponds to a template for the creation of objects.
- It specifies which attributes an object should have.

We've seen several built-in types: int, float, list, ...

Q: What does it mean that an object has a certain type?

A: The type defines which attributes the object has.

A different perspective:

- A type corresponds to a template for the creation of objects.
- It specifies which attributes an object should have.
- This template is used whenever you create a new object.

Object-oriented programming: terminology In OOP, a different terminology is used

Object-oriented programming: terminology In OOP, a different terminology is used

A template is called a class.

Object-oriented programming: terminology

In OOP, a different terminology is used

- A template is called a class.
- An object created from a class is called an instance of the class

Object-oriented programming: terminology

In OOP, a different terminology is used

- A template is called a class.
- An object created from a class is called an instance of the class
- An object is said to inherit the attributes of its class

Object-oriented programming: terminology

In OOP, a different terminology is used

- A template is called a class.
- An object created from a class is called an instance of the class
- An object is said to inherit the attributes of its class

Example: creating two files:

```
class: file
              data attributes:
                  name
                  closed
              method attributes:
                  close()
                             create instance:
create instance:
                             f2 = open(" file2.txt ")
f1 = open(" file1_txt ")
                             object: file
                            data attributes:
```

object: file

data attributes:
 name: "file1.txt"
 closed: False
 ...
method attributes:
 close()

object: file

data attributes:
 name: "file2.txt "
 closed: False
 ...
method attributes:
 close()
 ...

We have been creating instances all along...

We have been creating instances all along...

```
a = 5  # instance of int
b = "hello"  # instance of str
c = (1,2,3,4)  # instance of tuple
```

We have been creating instances all along...

```
a = 5  # instance of int
b = "hello"  # instance of str
c = (1,2,3,4)  # instance of tuple
```

In the general case, instances are created by calling the class as a function

We have been creating instances all along...

```
a = 5  # instance of int
b = "hello"  # instance of str
c = (1,2,3,4)  # instance of tuple
```

In the general case, instances are created by calling the class as a function

```
b = str(8)  # construct a string by calling str
c = tuple([1,2,3,4]) # constructing tuple by calling tuple
```

For the built-in types, this is mainly used for converting between types.

So far, we have used only the built-in types (int, string, ...)

So far, we have used only the built-in types (int, string, ...) Often, it is convenient to define your own type.

So far, we have used only the built-in types (int, string, ...) Often, it is convenient to define your own type.

Example: DNA sequence.

So far, we have used only the built-in types (int, string, ...) Often, it is convenient to define your own type.

Example: DNA sequence. It could be represented as string, but with a new type, we can:

So far, we have used only the built-in types (int, string, ...) Often, it is convenient to define your own type.

Example: DNA sequence. It could be represented as string, but with a new type, we can:

 ensure that only nucleotides (A,C,G,T) are allowed

So far, we have used only the built-in types (int, string, ...) Often, it is convenient to define your own type.

Example: DNA sequence. It could be represented as string, but with a new type, we can:

- ensure that only nucleotides (A,C,G,T) are allowed
- implement functionality: for instance calculate GC-content

So far, we have used only the built-in types (int, string, ...)

Often, it is convenient to define your own type.

Example: DNA sequence. It could be represented as string, but with a new type, we can:

- ensure that only nucleotides (A,C,G,T) are allowed
- implement functionality: for instance calculate GC-content

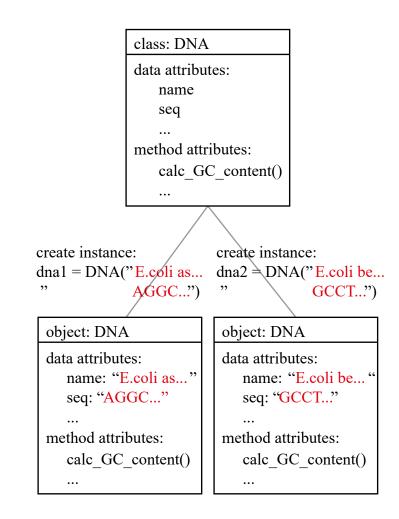
```
class: DNA
              data attributes:
                  name
                  seq
              method attributes:
                  calc GC content()
create instance:
                           create instance:
dna1 = DNA("E.coli as...
                           dna2 \(\)\(\)DNA("E.coli be...
               AGGC...")
                                          GCCT...")
 object: DNA
                            object: DNA
 data attributes:
                            data attributes:
    name: "E.coli as..."
                               name: "E.coli be..."
    seq: "AGGC..."
                               seq: "GCCT..."
 method attributes:
                            method attributes:
    calc GC content()
                               calc GC content()
```

So far, we have used only the built-in types (int, string, ...)

Often, it is convenient to define your own type.

Example: DNA sequence. It could be represented as string, but with a new type, we can:

- ensure that only nucleotides (A,C,G,T) are allowed
- implement functionality: for instance calculate GC-content



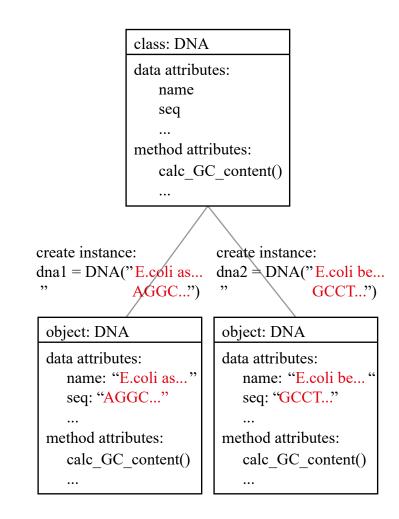
How do we define our own data type?

So far, we have used only the built-in types (int, string, ...)

Often, it is convenient to define your own type.

Example: DNA sequence. It could be represented as string, but with a new type, we can:

- ensure that only nucleotides (A,C,G,T) are allowed
- implement functionality: for instance calculate GC-content



How do we define our own data type? Write a class.

Writing a class

Defined like functions, but using the class keyword:

class class_name:
 class definition

Writing a class - simple example

```
class DNA:
    """Class representing sequences of Deoxyribonucleic acid"""
    pass  # Remember, pass is a placeholder - does nothing

dna = DNA()  # dna is an instance of DNA
```

Writing a class - defining methods

A method is defined like a function, but takes self as a first argument

Writing a class - defining methods

A method is defined like a function, but takes self as a first argument

ACGT

Writing a class - defining methods

A method is defined like a function, but takes self as a first argument

ACGT

self points to the current instance. In this case, when we call print_nucleotides(), dna_object will automatically be sent along as self.

Writing a class - adding data attributes

Attribute are like variables - but created within self.

```
dna class.py
class DNA:
    """Class representing sequences of Deoxyribonucleic acid"""
    def set name(self, name):
        """Set name attribute"""
        self.name = name
                                   # Here we set an attribute
    def get name(self):
        """Return name attribute"""
       return self.name
dna object = DNA()
                                      # Construct object of type DNA
dna object.set name("E.coli bla")
                                      # Call the set name method
print(dna object.get name())
                                      # Call the get name method
                                      # Access name attribute directly
print(dna object.name)
```

```
$ python dna_class.py
E.coli bla
E.coli bla
```

Exercise 1

- 1. Create an empty class called MyInfo.
- 2. Create a method called initialize in the class, which takes three arguments: firstname, lastname, and age and saves these as attributes.
- 3. Write a method in this class called print that prints out this information as:

```
Name: firstname lastname
Age: age
```

4. Check that you did it correctly, by executing:

```
info = MyInfo()
info.initialize("Barack", "Obama", 58)
info.print()
```

Exercise 1 - solution

1. Create an empty class called MyInfo.

```
class MyInfo:
    """Class representing sequences of personal information"""
    pass
```

2. Create a method called initialize in the class, which takes three arguments: firstname, lastname, and age and saves these as attributes.

```
class MyInfo:
    """Class representing sequences of personal information"""

def initialize(self, firstname, lastname, age):
    """Set name and age attributes"""
    self.firstname = firstname
    self.lastname = lastname
    self.age = age
```

Exercise 1 - solution (2)

3. Write a method in this class called print...

```
class MyInfo:
    """Class representing sequences of personal information"""
    ...

def print(self):
    output = "Name: " + self.firstname + " " + self.lastname + "\nAge: " + str(self.age)
    print(output)
```

4. Check that you did it correctly, by executing:

```
info = MyInfo()
info.initialize("Barack", "Obama", 58)
info.print()
```

Our example from before

```
class MyInfo:
   pass
info = MyInfo() # an empty object is pretty useless
```

Our example from before

```
class MyInfo:
   pass

info = MyInfo()  # an empty object is pretty useless
```

We would like to initialize our info object like this:

```
info = MyInfo("Barack", "Obama", 58)
```

Our example from before

```
class MyInfo:
   pass

info = MyInfo() # an empty object is pretty useless
```

We would like to initialize our info object like this:

```
info = MyInfo("Barack", "Obama", 58)
```

Q: How do we do this?

Our example from before

```
class MyInfo:
   pass

info = MyInfo() # an empty object is pretty useless
```

We would like to initialize our info object like this:

```
info = MyInfo("Barack", "Obama", 58)
```

Q: How do we do this?

A: We define a *constructor* in our class

Writing a class - constructor

A *constructor* is a special method that is called automatically when an object is created

Writing a class - constructor

A constructor is a special method that is called automatically when an object is created

In python, you can add a constructor to a class by defining an __init__ method

```
class MyInfo:
    def __init__(self, firstname):
        """Constructor"""
        self.firstname = firstname

info = MyInfo("Barack")
print(info.firstname)
```

Writing a class - constructor

A constructor is a special method that is called automatically when an object is created

In python, you can add a constructor to a class by defining an __init__ method

```
class MyInfo:
    def __init__(self, firstname):
        """Constructor"""
        self.firstname = firstname

info = MyInfo("Barack")
print(info.firstname)
```

Note that like all other methods, __init__ takes self as its first argument.

Exercise 2

- 1. Modify your MyInfo class from before, so that the initialize function becomes a constructor instead.
- 2. Check that you did it correctly, by executing:

```
info = MyInfo("Barack", "Obama", 58)
info.print()
```

Exercise 2 - solution

1. Modify your MyInfo class from before, so that the initialize function becomes a constructor instead.

```
class MyInfo:
    """Class representing sequences of personal information"""

def __init__(self, firstname, lastname, age):
    """Set name and age attributes"""
    self.firstname = firstname
    self.lastname = lastname
    self.age = age
...
```

Initializing objects: a recap

```
class MyInfo:  # Class definition
  pass  # this is just a placeholder

info = MyInfo()  # Instantiating our new class
```

Now we have an object. How do we add data to it?

1. Set attribute outside class

```
class MyInfo:
    pass

info = MyInfo()
info.name = "Barack"
```

1. Set attribute outside class

2. Function outside class

```
class MyInfo:
    pass

info = MyInfo()
info.name = "Barack"
```

```
class MyInfo:
    pass

def initializer(object, name_arg):
    object.name = name_arg

info = MyInfo()
initializer(info, "Barack")
```

1. Set attribute outside class

```
class MyInfo:
    pass
info = MyInfo()
info.name = "Barack"
```

2. Function outside class

```
class MyInfo:
    pass
def initializer (object, name arg):
    object.name = name arg
info = MyInfo()
initializer(info, "Barack")
```

3. Initializer method in class

```
class MyInfo:
    def initialize(self, name arg):
        self.name = name arg
info = MyInfo()
info.initialize("Barack")
```

1. Set attribute outside class

```
class MyInfo:
    pass

info = MyInfo()
info.name = "Barack"
```

2. Function outside class

```
class MyInfo:
    pass

def initializer(object, name_arg):
    object.name = name_arg

info = MyInfo()
initializer(info, "Barack")
```

3. Initializer method in class

```
class MyInfo:
    def initialize(self, name_arg):
        self.name = name_arg

info = MyInfo()
info.initialize("Barack")
```

```
class MyInfo:
    def __init__(self, name_arg):
        self.name = name_arg

info = MyInfo("Barack")
```

1. Set attribute outside class

class MyInfo:
 pass

info = MyInfo()
info.name = "Barack"
ATTRIBUTES DEFINED
OUTSIDE CLASS

2. Function outside class

```
class MyInfo:
    pass

def initializer(object, name_arg):
    object.name = name_arg

info = MyInfo()
initializer(info, "Barack")
```

3. Initializer method in class

```
class MyInfo:
    def initialize(self, name_arg):
        self.name = name_arg

info = MyInfo()
info.initialize("Barack")
```

```
class MyInfo:
    def __init__(self, name_arg):
        self.name = name_arg

info = MyInfo("Barack")
```

1. Set attribute outside class class MyInfo: pass info = MyInfo() info.name = "Barack" ATTRIBUTES DEFINED OUTSIDE CLASS 2. Function outside class class MyInfo: pass def initializer(object, name_arg): object.name = name_arg INITIALIZER DEFINED info = OUTSIDE_BCLASS

3. Initializer method in class

```
class MyInfo:
    def initialize(self, name_arg):
        self.name = name_arg

info = MyInfo()
info.initialize("Barack")
```

```
class MyInfo:
    def __init__(self, name_arg):
        self.name = name_arg

info = MyInfo("Barack")
```

1. Set attribute outside class

```
class MyInfo:
    pass

info = MyInfo()

info.name = "Barack"

ATTRIBUTES DEFINED

OUTSIDE CLASS
```

2. Function outside class

```
class MyInfo:
    pass

def initializer(object, name_arg):
    object.name = name arg
    INITIALIZER DEFINED
    info = OVINESIDE CLASS
    initializer(SIDE Backet)
```

3. Initializer method in class

```
class MyInfo:
    def initialize(self, name_arg):
        self.name = name_arg

info = MyInfo()
info.initialize("Barack")
```

```
class MyInfo:
    def __init__(self, name_arg):
        self.name = name_arg

info = MyInfo("Barack")
```

1. Set attribute outside class

```
class MyInfo:
    pass

info = MyInfo()
info.name = "Barack"
ATTRIBUTES DEFINED
    OUTSIDE CLASS
```

2. Function outside class

```
class MyInfo:
    pass

def initializer(object, name_arg):
    object.name = name arg
    INITIALIZER DEFINED
    info = OUTSIDE CLASS
    initializer(SIDE Backet)
```

3. Initializer method in class

```
class MyInfo:
    def initialize(self, name_arg):
        self.name = name_arg

info = MyInfo()
info.initialize("Barack")
```

```
class MyInfo:
    def __init__(self, name_arg):
        self.name = name_arg

info = MyInfo("Barack")

MOST ELEGANT
```

```
class MyInfo:
    def initialize(self, name):
        self.name = name

info = Myinfo()
info.initialize("Barack")
```

```
class MyInfo:
    def initialize(self, name):
        self.name = name

info = Myinfo()
    info.initialize("Barack")
```

```
class MyInfo:
    def initialize(self, name):
        self.name = name

info = Myinfo()
info.initialize("Barack")
```

```
class MyInfo:
    def __init__(self, name):
        self.name = name

info = Myinfo("Barack")
```

```
class MyInfo:
    def initialize(self, name):
        self.name = name

info = Myinfo()
info.initialize("Barack")
```

```
class MyInfo:
    def __init__(self, name):
        self name = name

info = Myinfo("Barack")
```

```
class MyInfo:
    def initialize(self, name):
        self.name = name

info = Myinfo()
info.initialize("Barack")
```

```
class MyInfo:
    def __init__(self, name):
        self_name = name

info = Myinfo("Barack")
```

It's really the same as:

```
class MyInfo:
    pass

info = MyInfo()
info.name = "Barack"
```

Object Oriented Programming: Inheritance

Problem:

How do we write a program that registers and keeps track of information about all types of employees and students at the University.

Different challenges:

Problem:

How do we write a program that registers and keeps track of information about all types of employees and students at the University.

Different challenges:

 Input: fancy graphics or simply from the command line. Let's ignore this part for now.

Problem:

How do we write a program that registers and keeps track of information about all types of employees and students at the University.

Different challenges:

- Input: fancy graphics or simply from the command line. Let's ignore this part for now.
- Main challenge: students and different types of employees have different properties. How do we model this?

Possible strategy

Possible strategy

We divide people into different categories (types)

Possible strategy

- We divide people into different categories (types)
- Each type of person has its own class.

Possible strategy

- We divide people into different categories (types)
- Each type of person has its own class.
- Whenever we create a new person object, we save it to a list or dictionary.

Which attributes should it have?

Name

- Name
- CPR number

- Name
- CPR number
- Address

- Name
- CPR number
- Address
- List of passed courses

- Name
- CPR number
- Address
- List of passed courses
- Grades

- Name
- CPR number
- Address
- List of passed courses
- Grades
- Enrollment date

- Name
- CPR number
- Address
- List of passed courses
- Grades
- Enrollment date

```
class Student:
    def init (self, name, cpr,
                 address, courses,
                 grades, enrollment):
        self.name = name
        self.cpr = cpr
        self.address = address
        self.courses = courses
        self.grades = grades
        self.enrollment = enrollment
albert = Student("Albert Einstein",
                 "140379-1235",
```

Which attributes should it have?

Name

- Name
- CPR number

- Name
- CPR number
- Address

- Name
- CPR number
- Address
- Office nr

- Name
- CPR number
- Address
- Office nr
- Job description

- Name
- CPR number
- Address
- Office nr
- Job description

```
class Technical:
    def init (self, name, cpr,
                 address, office,
                 job descr):
        self.name = name
        self.cpr = cpr
        self.address = address
        self.office = office
        self.job descr = job descr
elvis = Technical ("Elvis Presley",
                  "160835-6735",
                  "electrician")
```

Which attributes should it have?

Name

- Name
- CPR number

- Name
- CPR number
- Address

- Name
- CPR number
- Address
- Office nr

- Name
- CPR number
- Address
- Office nr
- List of publications

- Name
- CPR number
- Address
- Office nr
- List of publications
- Research interests

- Name
- CPR number
- Address
- Office nr
- List of publications
- Research interests

Our three classes

class Student		
name		
cpr_number		
address		
courses		
grades		
enrollment		

class Technical		
name		
cpr_number		
address		
office_number		
job_descr		

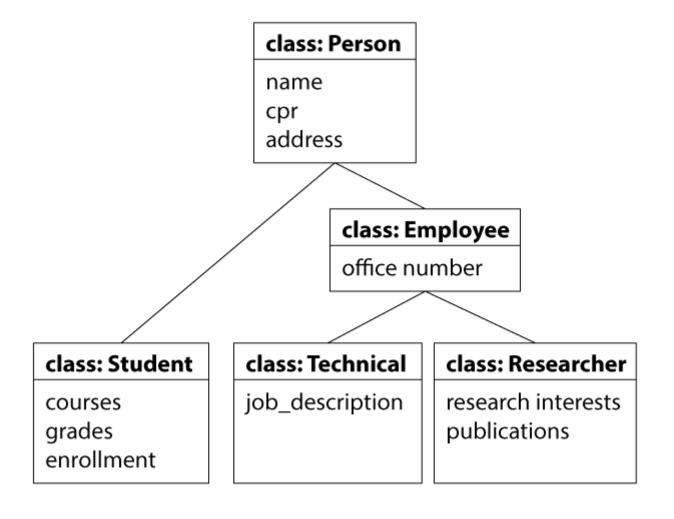
class Researcher
name
cpr_number
address
office_number
publications
interests

Our three classes

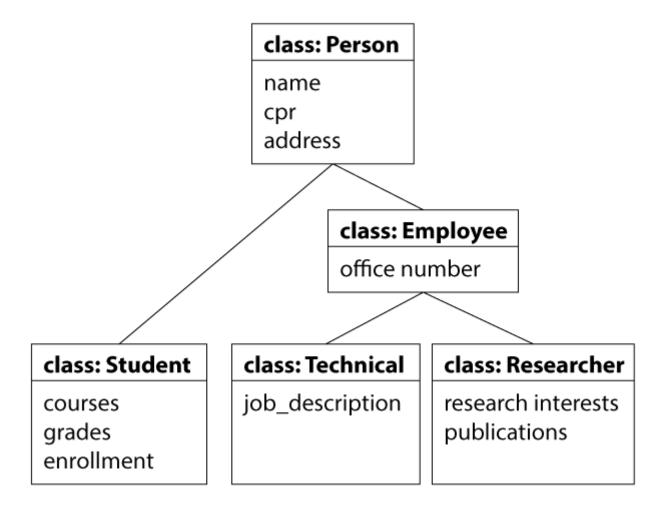
class Student	class Technical	class Researcher
name	name	name
cpr_number	cpr_number	cpr_number
address	address	address
courses	office_number	office_number
grades	job_descr	publications
enrollment		interests

There is quite a lot of overlap between our classes.

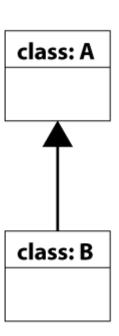
Consider the data as a hierarchy



Consider the data as a hierarchy

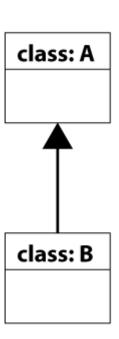


Idea: classes can inherit properties from other classes.

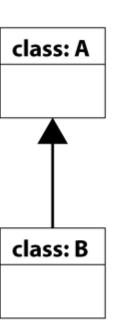


If class B inherits from A

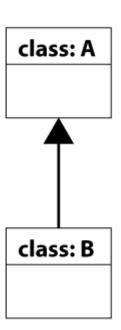
• A is the base-class (or super-class) of B



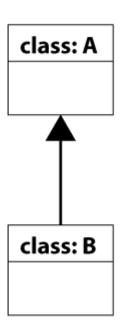
- A is the base-class (or super-class) of B
- B is a derived-class or sub-class of A



- A is the base-class (or super-class) of B
- B is a derived-class(or sub-class) of A
- B is a specialization of A



- A is the base-class (or super-class) of B
- B is a derived-class(or sub-class) of A
- B is a specialization of A
- Inheritance is called an is-a relationship:
 "B is an A"



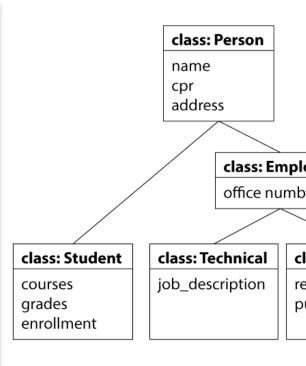
Inheritance in Python

To let a class B inherit from a class A:

```
class B(A):
    class definition
```

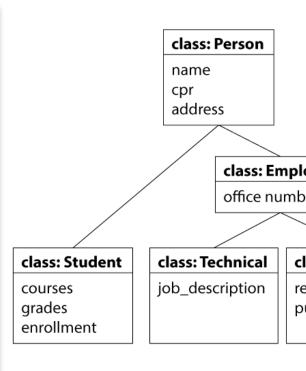
Inheritance in Python - for our data

```
class Person:
   # Attributes: name, cpr, address
class Employee(Person):  # Inherits from Person
    # Attributes: office
class Student(Person):  # Inherits from Person
    # Attributes: courses, grades, enrollment data
class Technical(Employee): # Inherits from Employee
    # Attributes: job description
class Researcher(Employee): # Inherits from Employee
    # Attributes: research interests, publications
```



Inheritance in Python - for our data

```
class Person:
   # Attributes: name, cpr, address
class Employee(Person):  # Inherits from Person
    # Attributes: office
class Student(Person):  # Inherits from Person
    # Attributes: courses, grades, enrollment data
class Technical(Employee): # Inherits from Employee
    # Attributes: job description
class Researcher(Employee): # Inherits from Employee
    # Attributes: research interests, publications
```



Note that I omitted the pass statements for brevity

Exercise 3

- 1. Implement a constructor for the Person class.
- 2. Implement a constructor for the Student class.

Exercise 3 - Partial solution

```
class Person:
   def init (self, name, cpr, address):
        self.name = name
        self.cpr = cpr
        self.address = address
class Student(Person):
    def init (self, name, cpr, address, courses, grades, enrollment):
        self.courses = courses
        self.grades = grades
        self.enrollment = enrollment
        # What about the name, cpr and address parameters?
albert = Student(name="Albert Einstein", cpr="14031879-1235",
                 address="112 Mercer Street, Princeton",
                 courses=["Linear Algebra", "Relativity"],
                 grades=["B","A"], enrollment=1895)
```

Exercise 3 - Partial solution

```
class Person:
    def init (self, name, cpr, address):
        self_name = name
        self.cpr = cpr
        self.address = address
class Student(Person):
    def init (self, name, cpr, address, courses, grades, enrollment):
        self.courses = courses
        self.grades = grades
        self.enrollment = enrollment
        # What about the name, cpr and address parameters?
albert = Student(name="Albert Einstein", cpr="14031879-1235",
                 address="112 Mercer Street, Princeton",
                 courses=["Linear Algebra", "Relativity"],
                 grades=["B", "A"], enrollment=1895)
```

What about the name, cpr, and address parameters in the Student constructor?

Calling the base-class constructor

To pass parameters to the base class, you will need to call the constructor of the base class explicitly

```
class A:
    def __init__(self):
        print("Initializing A")

class B(A):  # inherits from A
    def __init__(self):
        A.__init__(self)  # Call constructor of base class
        print("Initializing B")

b = B()
```

```
$ python base_class_example.py
Initializing A
Initializing B
```

Exercise 4

• Can you solve the previous exercise now?

Exercise 4 - solution

```
class Person:
   def init (self, name, cpr, address):
        self.name = name
        self.cpr = cpr
        self.address = address
class Student(Person):
    def init (self, name, cpr, address, courses, grades, enrollment):
       Person. init (self, name, cpr, address) # Base class constructor
        self.courses = courses
        self.grades = grades
        self.enrollment = enrollment
albert = Student(name="Albert Einstein", cpr="14031879-1235",
                 address="112 Mercer Street, Princeton",
                 courses=["Linear Algebra", "Relativity"],
                grades=["B","A"], enrollment=1895)
```

Object Oriented Programming: Composition

How do we model a group of students

How do we model a group of students

Idea: create a class that has a list of students as attribute

```
dream_team = StudentGroup([albert, niels])
```

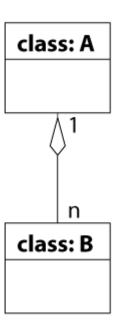
How do we model a group of students

Idea: create a class that has a list of students as attribute

```
dream_team = StudentGroup([albert, niels])
```

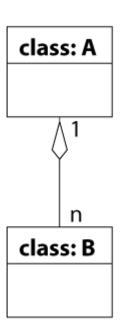
An object that contains other objects is called *composition*.

Consider a class A that is a composition of *n* instances of class B:



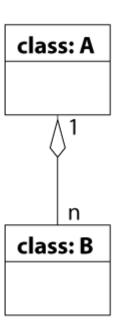
Consider a class A that is a composition of *n* instances of class B:

 Objects of type B are components within an object of type A.



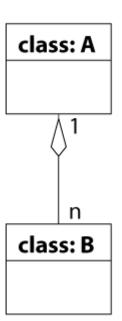
Consider a class A that is a composition of *n* instances of class B:

- Objects of type B are components within an object of type A.
- Think of the object A as consisting of smaller parts - each modelled by a separate class B



Consider a class A that is a composition of *n* instances of class B:

- Objects of type B are components within an object of type A.
- Think of the object A as consisting of smaller parts - each modelled by a separate class B
- Composition is a has-a relationship



Example

```
class StudentGroup:
    def __init__(self, students=[]):
        "''Constructor. The list of students is optional'''
        self.students = students

def add_student(self, student):
        "''Constructor. Add a new student to the group'''
        self.students.append(student)
```

OOP - composition exercise

 Make sure you have a working Student class (you can copy it from the slides), so you can create Student objects like this

 Add a get_names() method to the StudentGroup class example on the previous slide. The method should iterate over all students in the group and return a list of their names.

OOP - composition exercise - solution

```
class StudentGroup:
    # ... constructor and add student() methods omitted
    def get names (self):
        '''Return list of names of students in group'''
        names = []
        for student in self.students:
            names.append(student.name)
        return names
# Create student objects
albert = Student(name="Albert Einstein", cpr="14031879-1235")
niels = Student(name="Niels Bohr", cpr="07101885-7459")
# Create student group object
group = StudentGroup(students=[albert, niels])
# Call get names method
print(group.get names())
```

```
['Albert Einstein', 'Niels Bohr']
```