

Written examination – 14/05/2020

You have 1 hour and 30 minutes for completing the examination. By the final deadline, you should deliver only the original text (i.e. this document) with the definitive answers to the various exercises that must to be written with a pen – pencils are not permitted. You can keep all the draft papers that you may use during the examination for your convenience – blank sheets will be provided to you on request.

Section 1: basic questions

1 – Which of the following programming languages have been developed before 1970:

- Cobol
- Java
- Javascript
- Fortran
- Python

2 – Consider the following snippet of Python code:

```
def ni(s1, s2):  
    if s1 in s2 and s2 in s1:  
        return False  
    else:  
        return True
```

Which value is returned by calling the function above as follows: `ni("27", "42")`?

3 – Write down a small function in Python that takes in input two different positive numbers and returns a list of numbers from 0 to the previous number of the one calculated as the absolute value “| |” (e.g. $|4| = 4$, $|-4| = 4$) of the difference between the first input number and the second one.

4 – Write down which is the set of tools that compose a Turing Machine.

Section 2: understanding

Consider the following functions written in Python:

```
def f(email):
    user = email.split("@")[0]
    vowel = "aeiou"

    i = 0
    j = 0
    for c in user:
        if c not in ".0123456789":
            if c in vowel:
                i = i + 1
            else:
                j = j + 1

    if i < j:
        t = (i, j)
    else:
        t = (j, i)

    d = {"a": 0, "b": 0}
    for c in user.split(".")[1]:
        if c in vowel:
            d["a"] = d["a"] + t[1]
        else:
            d["b"] = d["b"] + t[0]

    return (d["a"], d["b"])
```

Consider the variable `my_email` containing the string of your institutional email address. What is the value returned by calling the function `f` as shown as follows:

```
f(my_email)
```

Section 3: development

Bibliographic coupling is a similarity measure that uses citation analysis to establish a similarity relationship between documents. Two documents are *bibliographically coupled* if they both cite one or more documents in common. The *coupling strength* of two given documents is higher the more citations to other documents they share. For instance, suppose that document A cites documents C, D, E, F, G, and B cites documents C, F, H, I. Thus, documents A and B have a bibliographic coupling strength of 2, since both cite documents C and F.

Write an algorithm in Python – `def bib_coupling(list_of_docs)` – which returns the average of the bibliographic coupling strength values calculated among all the pairs of documents in the input list. Each item in the input list, representing a document, is defined as the set of the documents (identified by strings) it cites. For instance, if we call the function with input `[{"A", "B", "C"}, {"B", "D"}, {"A", "C", "E"}]`, it will return 1 as result, since the bibliographic coupling strength between the 1st document and the 2nd document is 1, the bibliographic coupling strength between the 1st document and the 3rd document is 2, and the bibliographic coupling strength between the 2nd document and the 3rd document is 0 – thus the average of the bibliographic coupling strength values is $(1 + 2 + 0) / 3$, which is equal to 1.