

INTRODUCCIÓN A PYTHON PARA LINGÜISTAS

...

ALEJANDRO ARIZA

CENTRE DE LLENGUATGE I COMPUTACIÓ

UNIVERSITAT DE BARCELONA

INFORMACIÓN GENERAL

Aula: 4.4

Duración: 30 horas

Horario: Miércoles, 15:00-18:00, del 02/02/2022 al 06/04/2022

Instructor: Alejandro Ariza ([alejandro.ariza14 \[at\] ub.edu](mailto:alejandro.ariza14@ub.edu))

Estudiante de doctorado en UBICS

* Todo el material se encuentra disponible en un repositorio de github

¿A QUIÉN ESTÁ DIRIGIDO?

Curso de introducción a la programación y el procesamiento de lenguaje natural (texto) para lingüistas en Python. El curso también está pensado para otros perfiles con interés en programación (nivel de iniciación) para su aplicación a análisis de texto.

El curso está diseñado asumiendo que los participantes parten de una posición con cero experiencia en programación.

Se explicarán conceptos de programación como tipos de datos o estructuración de ideas conforme vayan presentándose en los ejemplos.

OBJETIVOS DEL CURSO

Los objetivos principales son:

- Entender los conceptos básicos de la programación

- Adquirir conocimientos de estructurado de programas

- Aprender a traducir un problema lingüístico en un problema de programación

En las sesiones prácticas, trabajaremos con:

- Estructuras de control, bucles, funciones y ficheros de texto

- Librerías Python externas pensadas para el procesamiento de texto (NLTK, Spacy)

- Diferentes corpus (Inglés y Español)

AGRADECIMIENTOS Y BIBLIOGRAFÍA

Las diapositivas del curso son adaptaciones del material impartido por Venelin Kovatchev en el curso 2020

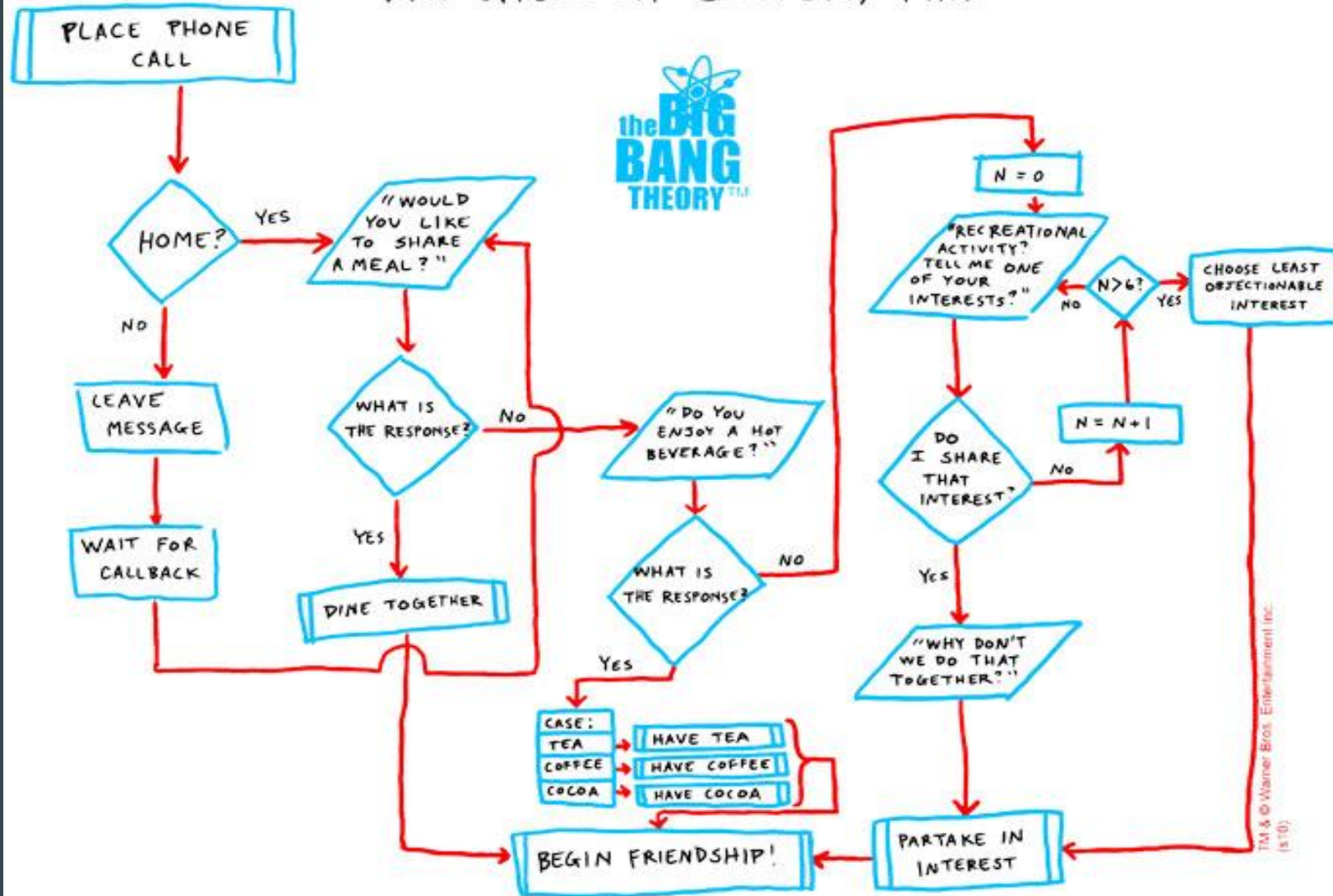
Algunos ejemplos provienen de “Introducción a la programación para humanistas” (Leticia Martín-Fuertes Moreno)

“The NLTK book” y “Python Programming: An Introduction to Computer Science” son buenos recursos adicionales para familiarizarse con Python y NLTK

INTRODUCCIÓN A PENSAMIENTO ALGORÍTMICO

THE FRIENDSHIP ALGORITHM

DR. SHELDON COOPER, Ph.D



COMPUTACIÓN = DATOS + ALGORITMOS

Algoritmo [informática]: *“es un conjunto de instrucciones o reglas definidas y no-ambiguas, ordenadas y finitas que permite, típicamente, solucionar un problema, realizar un cómputo, procesar datos y llevar a cabo otras tareas o actividades.”*,

Wikipedia

Datos [informática]: *“Cifras, letras o palabras que se suministran a la computadora como entrada y la máquina almacena en un determinado formato.”*

Oxford Languages

¿POR QUÉ ES ÚTIL LA COMPUTACIÓN?

La computación es práctica

La cantidad de datos disponibles (tanto estructurados como no estructurados) crece de forma exponencial a una velocidad que los humanos no pueden procesar únicamente a mano

De esta forma y, con instrucciones apropiadas, los ordenadores pueden ejecutar numerosas tareas con mayor rapidez que los humanos

La computación es científica

La automatización permite a los investigadores avanzar a velocidades muy superiores tanto en la preparación y ejecución de experimentos como en el análisis de los resultados

El objetivo es la Inteligencia Artificial no asistida

Todavía queda un largo camino

LENGUAJES DE PROGRAMACIÓN NUESTRA COMUNICACIÓN CON LOS ORDENADORES

CAPAS DE ABSTRACCIÓN: DESDE 0 Y 1 AL LENGUAJE HUMANO

Un ordenador formado por circuitos solo atiende a impulso eléctricos

Estos impulsos eléctricos es lo que identificamos como 0s (ausencia) y 1s (impulso)

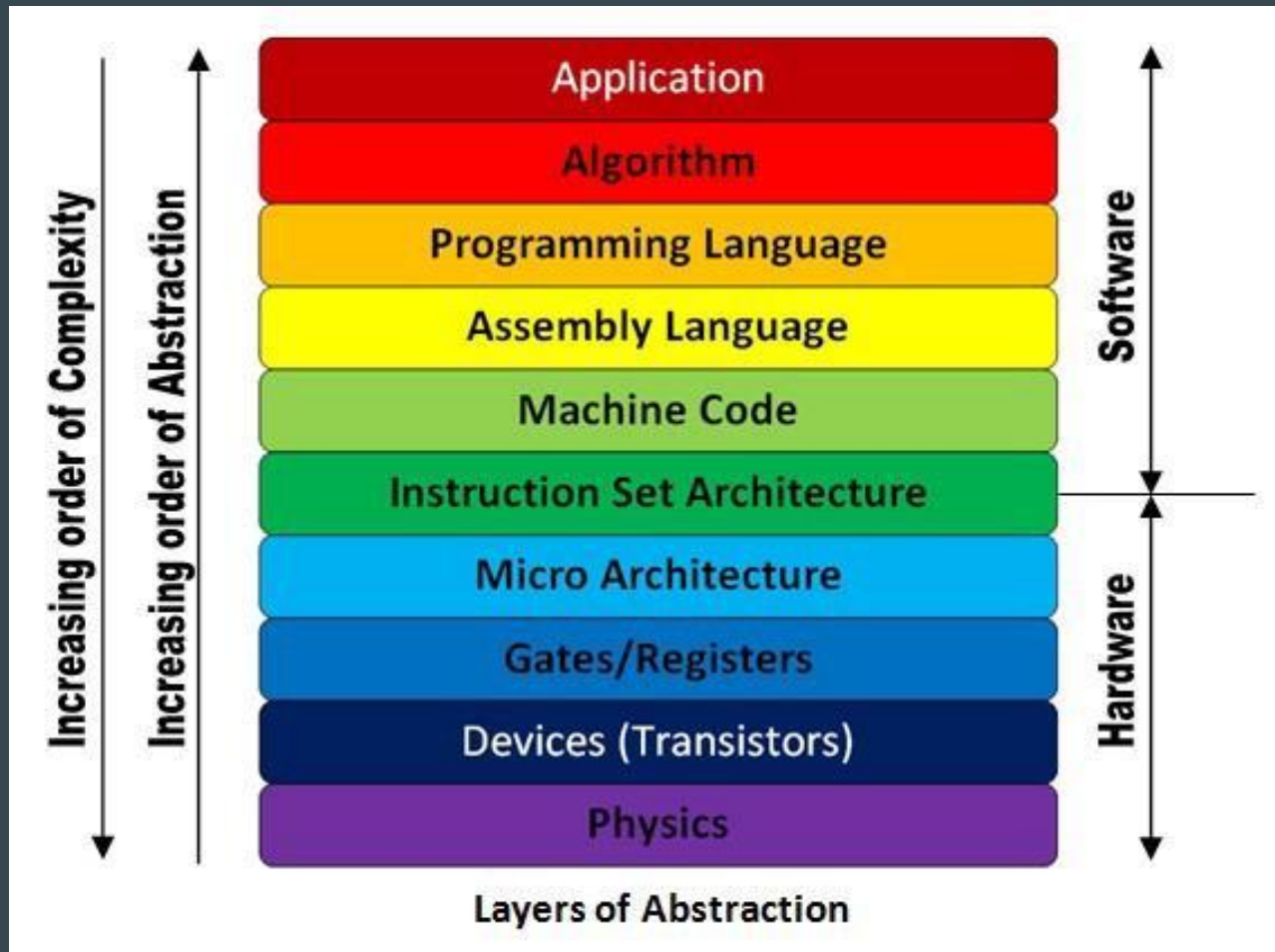
El cerebro humano también funciona por impulsos eléctricos causados por los neurotransmisores

Tras numerosas interacciones entre las neuronas, un humano es capaz de entender tanto lenguajes naturales como lenguajes artificiales, matemáticas, etc

Entonces, ¿Cómo pueden los humanos y los ordenadores entenderse?

Hacer que los humanos aprendan el lenguaje de los 0s y los 1s

“Enseñar” a los ordenadores un lenguaje que sea más parecido al de los humanos



CAPAS DE ABSTRACCIÓN

- Contra más capaz de abstracción se añaden, más “humano” parece
- Los niveles más bajos de la abstracción son más similares al lenguaje máquina
- En este curso, vamos a trabajar con los niveles más altos (no es necesario que aprendáis el lenguaje de 0-1)
- Contra mayor abstracción, mayor interpretabilidad pero menor velocidad de cómputo

LENGUAJES DE PROGRAMACIÓN

Los lenguajes de programación son un compromiso entre 0-1s y el lenguaje natural humano

Para el programador, un lenguaje de programación es un lenguaje:

- Con estructura y reglas predefinidas (Sintaxis)

- Puede ser aprendido con la práctica

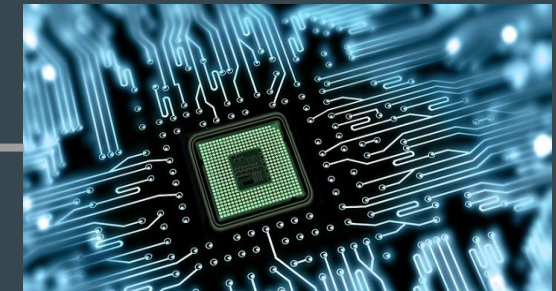
Para el ordenador, un lenguaje de programación es un programa:

- Incluye todas las instrucciones necesarios para “traducir” los comandos a 0s y 1s

LENGUAJES DE PROGRAMACIÓN



```
for i in people.data.users:
    response = client.api.statuses.user_timeline.get(screen_name=i.screen_name)
    print 'Got', len(response.data), 'tweets from', i.screen_name
    if len(response.data) != 0:
        ltdate = response.data[0]['created_at']
        ltdate2 = datetime.strptime(ltdate, '%a %b %d %H:%M:%S +0000 %Y')
        today = datetime.now()
        howlong = (today-ltdate2).days
        if howlong < daywindow:
            print i.screen_name, 'has tweeted in the past', daywindow,
            totaltweets += len(response.data)
            for j in response.data:
                if j.entities.urls:
                    for k in j.entities.urls:
                        newurl = k['expanded_url']
                        urlset.add((newurl, j.user.screen_name))
        else:
            print i.screen_name, 'has not tweeted in the past', daywindow
```



Vosotros

El lenguaje de
programación

0101101100011
0100100010111
01111101100...

PYTHON COMO LENGUAJE DE PROGRAMACIÓN

El lenguaje de este curso

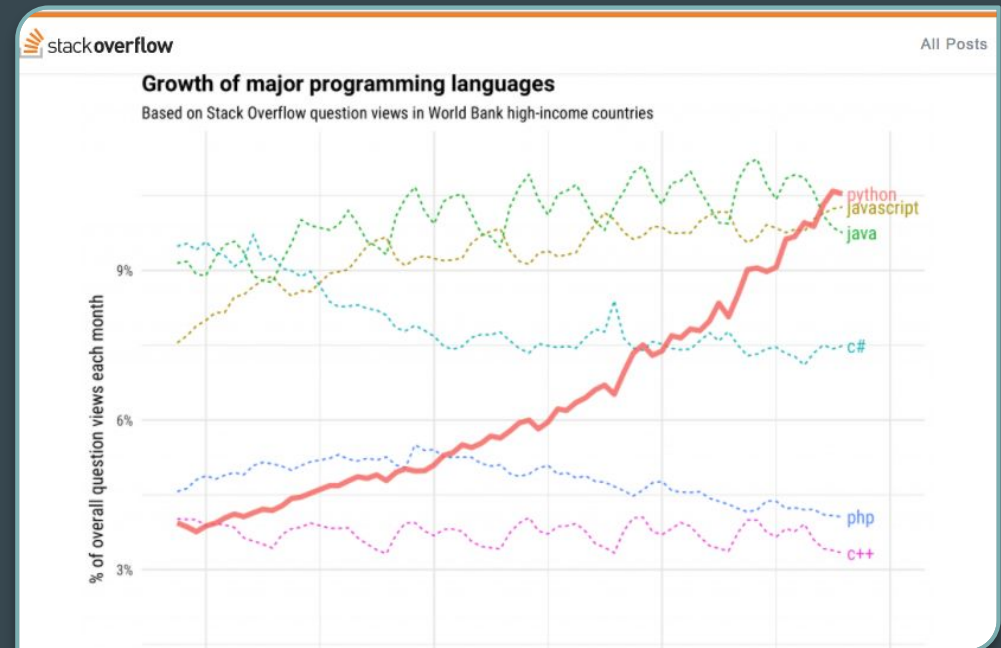
Rápida curva de aprendizaje

Intenta acercarse al lenguaje humano

Lenguaje con gran aplicabilidad (ampliamente adoptado por la comunidad científica)

Extremadamente popular en los últimos años

Muchas librerías externas y programas a nuestra disposición



¡GRACIAS!
¿PREGUNTAS?