# תחביר בסיסי

## הגדרה

פייתון היא שפת תכנות כללית High-Level. נוצרה על ידי גווידו ואן רוסום בשנים 1985 – 1990. נקראת על שם סדרת הילדים "הקרקס המעופף של מונטי-פייתון".

### יתרונות

* קוד פתוח.
* מבוססת על Interpreter - כלומר ה-Interpreter מקמפל ומריץ את התוכנית שורה אחר שורה ולא מקמפל את כל התוכנית לפני ההרצה. היתרון הוא שלא צריך לקמפל את התוכנית ומגיעים לשלב ההרצה יותר מהר. החיסרון הוא שביצוע כל שורה לוקח יותר זמן מאשר תוכנית שקומפלה לפני הרצה. יש אפשרות גם לקמפל בשביל הרצה מהירה.
* אינטראקטיבית - ניתן לכתוב ולהריץ כל שורה בנפרד באמצעות ה-Interpreter.
* תמציתית וקריאה - בדרך כלל דורשת הרבה פחות שורות קוד משפות אחרות.
* דינאמית - אין צורך להכריז על משתנים.
* מונחית עצמים.
* נפוצה מאוד בתחומים רבים, כמו: פיתוח רשת, ניתוח מסדי נתונים.

## Interpreter

ה- Interpreterמקמפל ומריץ את התוכנית שורה אחר שורה. תכונה זו מאפשרת לנו לכתוב קוד בצורה אינטראקטיבית, כלומר לאחר כתיבת כל שורה ניתן ישר לקבל תוצאה. ב-Interpreter הרצה של שם משתנה, פעולה מתמטית, או משפט המוקף בגרשיים ידפיס את הערך של המשתנה או המשפט למסך. בהדפסת משפט צריך להקפיד שהגרשיים מאוזנות.

>>> a = 5

>>> a

5

>>> 'He said:"Are you ok?"'

'He said:"Are you ok?"'

## גרסאות

החל מ-2020 פייתון הסירה את התמיכה שלה מ-python2 ומספקת תמיכה אך ורק ב-python3. ניתן לבדוק את הגרסה של פייתון שאנו עובדים בה על ידי הרצה של הפקודה:

python --version

## כללי עיצוב קוד

### שמות

* שמות משתנים ופונקציות יכתבו באותיות קטנות ועם קו תחתון בין מילים - variables\_and\_functions.
* שמות ערכים קבועים יכתבו באותיות גדולות בלבד ועם קו תחתון בין המילים - UPPERCASE\_CONSTANTS.
* שמות מחלקות יחוברו עם אות גדולה בין מילים - ClassName.
* משתנים שלא משתמשים בהם, כמו לדוגמא בלולאה, נסמן בקו תחתון '\_'. לדוגמה, for \_ in range(10).

### כללים חשובים נוספים

* בפייתון, המבנה של בלוקים של קוד בתוכנית נקבע על ידי ההזחה שלהם. עצם הזחת הבלוק מגדירה אותו כבלוק תכנותי שונה. המטרה היא לשפר את קריאות התוכנית, בכך שהיא כופה על המתכנת כתיבה בסטנדרט אחיד (פחות או יותר), וחוסכת בקוד. כל הזחה גדולה או שווה ל-2 היא חוקית כל עוד אנו עקביים. מומלץ להגדיר את tab להיות 4 רווחים כי אורך ה-tab יכול להיות שונה בין מערכות הפעלה.
* לבלוקים של קוד, בלולאות, מחלקות, פונקציות וכדומה, מקדימים תמיד את התו ':' (נקודתיים).
* פקודות מופרדות לרוב על ידי מעבר שורה. שימוש בתו ';' (נקודה ופסיק) להפרדה בין פקודות הוא אופציונלי, וכמעט שלא בשימוש בפועל.

## הערות

הערות בפייתון מתחילים עם סימון #.

## טיפוסים בסיסיים

השפה תומכת במגוון גדול יחסית של טיפוסי נתונים בסיסיים. המרכזים שבהם:

* int - מאחסן מספר שלם. בגרסה 3.0 אוחד טיפוס זה אל תוך long ואין לו הגבלת גודל (מעבר למה שמאפשר הזיכרון הווירטואלי של המחשב). (בדומה ל־BigInteger בשפת Java).
* float - נקודה צפה. מאפשר אחסון ערך מקורב של מספר ממשי באמצעות ייצוג נקודה צפה. טיפוס זה זהה לטיפוס double בשפת C (ולמעשה ברוב המקרים ממומש על ידיו). בנוסף, ניתן לייבא את המודול decimal המאפשר ייצוג של שברים עשרוניים, ונוח יותר לעבודה במקרים בהם הייצוג הזה מדויק יותר (כגון חשבונאות).
* complex - מספר מרוכב, נכתב בצורה 3+4j. טיפוס זה נוסף לשפה בשל השימוש הרב שעושים בו מדענים ומהנדסים, שהם אחת מהקהילות העיקריות המשתמשות בשפה, ועל מנת למנוע מהם "להמציא את הגלגל" כל פעם מחדש.
* bool - טיפוס בוליאני שיכול להחזיק True או False (עם אות גדולה בהתחלה בניגוד ל-Java ו-C++).
* string - מחרוזות. החל מגרסה 3, המחרוזות הן Unicode, אם כי ישנו טיפוס bytes עבור רצפים של בתים.

אובייקטים מטיפוסי מספרים ומחרוזות הם בלתי ניתנים לשינוי (Immutable), כלומר שינוי ערכו של משתנה מטיפוס כזה הוא בעצם הסרת ההתייחסות אליו והפנייתה לערך חדש.

## משתנים

בפייתון אין צורך להגדיר טיפוס למשתנה, אלא אפשר לרשום ישירות את שם המשתנה ואת הערך ההתחלתי שאנו נותנים לו. הצהרה מסוג זה נקראת dynamic typing. ה-interpreter מבין לבדו לפי ההשמה על המשתנה מהו הטיפוס שלו. לאחר יצירת משתנה ניתן לדרוס את הטיפוס שלו ולתת לו השמה מטיפוס אחר. ניתן לגלות את הטיפוס הנוכחי של משתנה באמצעות פונקציית type.

>>> x = 3

>>> type(x)

<class ‘int’>

>>> x = 3.0

>>> type(x)

<class ‘float’>

ניתן להצהיר על כמה משתנים באותה שורה בדרך הבאה:

t, f = True, False

## אופרטורים

קיימים כל האופרטורים הרגילים שיש בשאר השפות מלבד האופרטורים '++', '--' שאינם חוקיים בפייתון. במקומם אפשר להשתמש ב-‘+=1’. אופרטורים לוגיים בפייתון הם המילים and, or, not. אופרטור xor מפעילים באמצעות !=. ישנם גם מספר אופרטורים חדשים:

\*\* - מייצג חזקה כאשר האופרנד משמאל הוא הבסיס והאופרנד מימין הוא המעריך. אפשר להשתמש גם באופרטור \*\*= שמבצע השמה למשתנה כאשר הוא גם הבסיס בחזקה.

>>> 7\*\*2 # 7^2

49

// - חלוקה ללא שארית. אפשר להשתמש גם באופרטור //= שמבצע השמה למשתנה כאשר הוא גם המונה בחלוקה.

>>> 9//2

4

in - אופרטור השתייכות. בודק האם איבר כלשהו נמצא במשפט, רשימה או מפתחות מילון. יש גם אופרטור הפוך ‘not in’.

>>>"ell" in "Hi and hello"

True

is - אופרטור זהות. בודק אם המשתנים משני צידי האופרטור מצביעים על אותו אובייקט. קיים גם אופרטור ‘is not’.

## תנאים

אותו עקרון של if, if else, else שדומה לשאר השפות, אלא שבפייתון תנאים לא חייבים להיות בתוך סוגריים עגולות, לאחר הצבת התנאי רושמים נקודותיים, והתחום של התנאי הוא עד שמפסיקים את הרווח (טאב) בכל שורה. משתנה שהוגדר בתוך התנאי חי גם מחוץ לתנאי.

>>> a = 7

>>> if a < 5:

... b = 10

... else:

... b = 3

>>> b

3

## לולאות for

לולאת for משמשת כדי לעבור על כל האיברים ברשימה מסוימת. למעשה, כל הלולאות בפייתון הם לולאות foreach.

>>> list = [2, 3, "Hello",]

>>> for x in list:

... print(x)

2

3

Hello

כאשר רק רוצים לעבור בלולאה מספר כלשהו של פעמים, אפשר ליצור רשימה של מספרים טבעיים בסדר עולה באמצעות פונקציית range(num).

>>> for i in range(10):

… print(i)

ניתן לקבל בכל איטרציה איבר ברשימה וכן את האינדקס שלו על ידי שנפעיל את הפונקציה enumerate(<list>) על הרשימה, שמחזירה מילון שבו האינדקס הוא המפתח והערך הוא האיבר המתאים ברשימה.

>>> for (i, obj) in enumerate(list):

... print('%d element in list is %s' %(i, str(obj)))

0 element in list is 2

1 element in list is 3

2 element in list is Hello

שימוש נוסף ללולאות for:

>>> a = [x\*x for x in range (1, 10)]

>>> a

[1, 4, 9, 16, 25, 36, 49, 64, 81]

אפשר גם לשלב עם תנאים:

>>> even\_squares = [x\*x for x in range (1, 10) if x % 2 == 0]

>>> print(even\_squares)

[4, 16, 36, 64]

# מחרוזות ומכלים

## מחרוזות

מחרוזת היא אינה מיכל בפייתון אלא טיפוס בסיסי, כיוון שאין בשפה טיפוס מסוג תו, וההתייחסות לתו בודד היא כאל מחרוזת באורך 1. עם זאת, ניתן לגשת אל כל תו במחרוזת באמצעות גישה לפי אינדקס, כמו בטיפוסים הסדרתיים, וניתן להמיר ביניהם בקלות. ניתן לכתוב מחרוזות הן בעזרת גרשיים כפולים: "hello" והן בעזרת גרש יחיד: 'hello'. אין כל הבדל בין צורות אלו, אך מקובל להשתמש בגרש יחיד. מחרוזות הנכתבות בין שלושה גרשיים רצופים - """hello""" או '''hello''' - ניתן לפרוש על פני מספר שורות ברצף.

### Formatting

דרך מאוד נוחה להדפיס מחרוזות עם ערכים משתנים היא באמצעות formatting, שבה ניתן להגדיר מחרוזות עם דגלים {} שאותם נמלא בסוף המחרוזות בתוך השיטה format שתופעל על המחרוזת.

>>> print('{} {} {}'.format(hello, world, 12))  # string formatting

hello world 12

### שיטות מרכזיות

s = hello

print(s.capitalize()) # Capitalize a string. Prints “Hello”

print(s.upper()) # Convert a string to uppercase. Prints "HELLO"

print(s.rjust(7)) # Right-justify a string, so that the total size of the string will be 7. padding with spaces. Prints “ hello”

print(s.center(7)) # Center a string, padding with spaces. Print “ hello ”

print(s.replace('l', '(ell)')) # Replace all instances of one substring with another. Prints “he(ell)(ell)o”

print(' world '.strip()) # Strip leading and trailing whitespace. Prints “world”

## רשימות - Lists

רשימה בפייתון זהו רצף של עצמים, לא חייב שיהיו מאותו סוג. מצהירים על רשימה באמצעות סוגריים מרובעות שבתוכן האיברים מופרדים בפסיקים.

>>> a = [4, 3.0, 'Hello']

אפשר גם להגדיר רשימה שהיא איבר פנימי בתוך רשימה אחרת.

>>>a = [4, [5,4, [5.2, 9.3]]]

>>>a[1][2][0]

5.2

### גישה לאיברים

הגישה לאיבר ברשימה היא בדיוק כמו במערך ב-Java, כלומר רושמים את שם הרשימה ואז בסוגריים מרובעות את אינדקס האיבר הרצוי. גישה לאינדקס לא מוגדר יוביל לשגיאה. כדי לגשת לכל האיברים מאינדקס i עד אינדקס j (לא כולל) יש להוסיף בתוך הסוגריים [i:j]. כאשר i חסר הכוונה זה מאינדקס 0, וכאשר j חסר הכוונה עד האינדקס האחרון (כולל). גישה לאינדקס שלילי חוזר אחורה מהסוף להתחלה ברשימה. שימוש בתוך הסוגריים ב-"::" ולאחר מספר טבעי יחזיר את הרשימה המתקבלת עם קפיצות של n בין איבר לאיבר.

>>> a[1]

3.0

>>> a[1:2]

3.0

>>> a[-1]

'Hello'

>>> a[::2]

[4, 'Hello']

### שיטות מרכזיות

יש המון שיטות שימושיות שניתן להפעיל על רשימות.

ls.append('bar') # Add a new element to the end of the list

ls.pop() # Remove and return the last element of the list

## Tuples

Tuples הם אוסף מסודר של אובייקטים. הם זהים בדיוק לרשימות אלא ש-tuple הוא מוגן, כלומר לאחר שיוצרים אותו לא ניתן לשנות את האיברים. ניסיון לשנות איבר בתוך tuple יוביל לשגיאת TypeError. הבדל נוסף הוא שעל tuple מצהירים באמצעות סוגריים עגולות ולא מרובעות. הגישה לאיבר ב-tuple זהה בדיוק לגישה לאיבר במערך. כדי להגדיר tuple עם איבר אחד יש לרשום בתוך הסוגריים את הערך ולאחריו פסיק.

>>> a = (4, 3.0, "Hello")

>>> a[1:2]

3.0

>>>t = (1,2,3)

>>>(a, b, c) = t

>>>a

1

Tuple, להבדיל מרשימה גם יכול להיות מפתח במילון וכן איבר ב-set.

d = {(x, x + 1): x for x in range(10)}  # Create a dictionary with tuple keys

## מילונים - Dictionaries

מילון הוא אוסף של איברים מסודרים בצורה של מפתח-ערך. מגדירים מילון באמצעות סוגריים מסולסלות {}. כל איבר מופרד בפסיק, ובין כל מפתח ערך יש נקודותיים ':'. המפתחות במילון חייבים להיות ייחודיים בעוד שהערכים אינם חייבים להיות ייחודיים. הערכים של מילון יכולים להיות מכל סוג שהוא, אך המפתחות חייבים להיות מסוג נתונים בלתי משתנה (immutable) כגון מחרוזות, מספרים או Tuples.

### גישה לאיברים

כדי לגשת לערך במילון ניתן לרשום בסוגריים מרובעות את המפתח שלו.

>>>dict = {'Name': 'Zara', 'Age': 7, 'Class': 'First'}

>>>dict[‘Name’]

Zara

באמצעות אופרטור in ניתן לבדוק האם קיים מפתח כלשהו במילון.

>>> ‘Class’ in dict

True

אם ניגשים למתח במילון שלא קיים תתקבל שגיאה “not a key”. ניתן להשתמש בשיטה get על המילון המקבלת מפתח וערך ברירת מחדל, כך שבמידה והערך לא נמצא יוחזר ערך ברירת מחדל.

>>> print(dict.get('monkey', 'N/A'))

N/A

### שיטות מרכזיות

ניתן לעדכן ערך על ידי גישה למפתח שלו ולאחר מכן השמה. בשיטה זו ניתן גם להכניס רשומה חדשה למילון. מוחקים ערך באמצעות אופרטור del.

>>> dict['Age'] = 8; # update existing entry

>>> dict['School'] = "DPS School"; # Add new entry

>>> del dict[‘Age’] # delete existing entry

### איטרציה

ניתן לעבור על כל איברי המילון באמצעות שיטת items המחזירה רשימה של זוגות:

>>> d = {'person': 2, 'cat': 4, 'spider': 8}

>>> for animal, legs in d.items():

…   print('A {} has {} legs'.format(animal, legs))

A person has 2 legs

A cat has 4 legs

A spider has 8 legs

## Set

זהו אוסף לא מסודר של איברים **שונים**. מגדירים set באמצעות סוגריים מסולסלות {}. כל איבר מופרד בפסיק. האיברים ב-set חייבים להיות ייחודיים. הערכים יכולים להיות מכל סוג שהוא. איטרציה על איברים ה-set היא בדיוק כמו ברשימה, אך בגלל שהאיברים לא מסודרים אי אפשר לדעת מה יהיה סדר האיברים באיטרציה.

animals = {'cat', 'dog'}

### שיטות מרכזיות

animals.add('fish')    # Add an element to a set

animals.add('cat') # Adding an element that is already in the set does nothing

animals.remove('cat') # Remove an element from a set

## פעולות על מחרוזות ומכלים

* כדי לקבל את אורך המיכל או המחרוזת משתמשים בפונקציה len.
* הפעלת אופרטור + תשרשר בין המכלים או המחרוזות.
* ניתן להשתמש באופרטור in, המחזיר ערך bool, כדי לבדוק אם איבר נמצא במיכל.

# פונקציות

## הגדרה

פונקציות מתחילים עם מילת המפתח def ואחריהם שם הפונקציה וסוגריים. בתוך הסוגריים יש למקם את כל פרמטרים לפונקציה. הבלוק של הפונקציה מתחיל בנקודותיים שבאים לאחר הסוגריים. הבלוק של הפונקציה הם כל השורות המוזחות שבאות אחרי ההצהרה. המשפט הראשון של פונקציה יכול להיות משפט אופציונלי - מחרוזת תיעוד לפונקציה או docstring. פונקציה תמיד תסתיים ב-return. באמצעות return אפשר להחזיר משהו חזרה לפונקציה הקוראת. במידה וה-return ריק, זה שקול ל-return None.

def func\_name(parameters):

"function\_docstring"

function\_suite

return [expression]

כל הארגומנטים בפייתון מעברים By Reference, ולכן כל שינוי בארגומנט בתוך הפונקציה ישפיע גם בחזרה על הפונקציה הקוראת. ניתן להגדיר בפונקציה ערכי ברירת מחדל לפרמטר כך שיהיה ניתן לקרוא לפונקציה גם בלי להכניס את הפרמטרים שהם ברירת מחדל. לאחר שמגדירים פרמטר עם ברירת מחדל לא ניתן להכניס פרמטרים ללא ברירת מחדל לאחריו.

def printinfo( name, age = 35 ):

"This prints a passed info into this function"

print "Name: ", name

print "Age ", age

return;

## פונקציות אנונימיות - Lambda Expressions

פונקציות אלה נקראות אנונימיות מכיוון שהן אינן מוכרזות באופן הסטנדרטי באמצעות המילה השמורה def. אלו פונקציות המוגדרות באמצעות שורה אחת. ההצהרה עליהן מתחילה במילה השמורה lambda לאחר מכן רושמים את הפרמטרים שהפונקציה מקבלת עם פסיק בין כל פרמטר. ניתן להכניס איזה מספר של פרמטרים שנרצה. לאחר הפרמטרים רושמים נקודותיים ':', ואז את הביטוי שהוא גם ערך החזרה מהפונקציה.

lambda [arg1 ,arg2,.....argn]: expression

דוגמאות לפונקציות אנונימיות:

>>> f = lambda x: x\*x

>>> f(5)

25

>>> def o(f,g): # composite functions

... return lambda x: f(g(x))

>>> b = o(lambda x: x\*x, lambda x: x+1)

>>> b(3)

16

# קבצים

## open

כדי להחזיר אובייקט המייצג קובץ משתמשים בפונקציה open. הפונקציה מקבלת:

* path אל הקובץ אותו רוצים לפתוח.
* הרשאת גישה: "r" לקריאה. "w" לכתיבה. "a" להוספה. "r+" הן לקריאה והן לכתיבה.

file = open('file\_path', 'w')

### with

נהוג בפתיחת קובץ להשתמש בהצהרת with כדי לטפל בחריגות. חריגה במהלך פעולה על קובץ עלולה למנוע מהקובץ להיסגר כראוי, מה שעלול להכניס מספר באגים בקוד, כלומר שינויים רבים בקבצים לא נכנסים לתוקף עד שהקובץ נסגר כראוי. כשמשתמשים ב-with אין צורך לקרוא ל-file.close() אלא הצהרת with עצמה מבטיחה הקצאה ושחרור של משאבים כאשר הקוד המשתמש במשאב מבוצע לחלוטין. הצהרה זו בעצם שקולה לשימוש ב-try לפני פקודה על קובץ ו-finally כדי לסגור את הקובץ, אלא שהצהרת with עוזרת להפוך את הקוד ליותר נקי וקריא.

with open('file\_path', 'w') as file:

file.write('hello world !')

## פעולות על קובץ

### close

מפסיקה את כל המשאבים הנמצאים בשימוש ומשחררת את המערכת של תוכנית מסוימת זו.

# NumPy

## הגדרה

NumPy היא ספרייה בפייתון המוסיפה תמיכה במערכים ומטריצות רב-ממדיות, יחד עם אוסף גדול של פונקציות מתמטיות ברמה גבוהה להפעלת מערכים אלה. NumPy היא תוכנת קוד פתוח ויש לה תורמים רבים.

לאחר התקנת NumPy, בכל תוכנית צריך לייבא אותו על ידי import. NumPy מיובא בדרך כלל תחת הכינוי np.

import numpy as np

## מערך ב-NumPy

בפייתון יש לנו רשימות המשרתות את מטרת המערכים, אך הן איטיות לעיבוד. NumPy שואפת לספק אובייקט מערך המהיר פי 50 מהרשימה המובנית של פייתון. מערכי NumPy מאוחסנים במקום אחד רציף בזיכרון בניגוד לרשימות, כך שתהליכים יכולים לגשת אליהם ולנהל אותם בצורה יעילה מאוד. כמו כן, הוא מותאם לעבודה עם ארכיטקטורות המעבד העדכניות ביותר. מערכים משמשים לעתים קרובות במדעי הנתונים, שם המהירות והמשאבים חשובים מאוד.

אובייקט המערך ב-NumPy נקרא ndarray, מפני שניתן להגדיר לכל מערך n ממדים (מערכים מקוננים). מערך ב-NumPy מספק שיטות תומכות רבות שהופכות את העבודה עם ndarray לקלה מאוד. הגישה לאיברים במערך היא כמו שלמדנו בגישה לאיבר ברשימה.

>>> arr0d = np.array(42)

>>> arr1d = np.array([1, 2, 3, 4, 5])

>>> arr2d = np.array([[1, 2, 3], [4, 5, 6]])

>>> arr3d = np.array([[[1, 2, 3], [4, 5, 6]], [[1, 2, 3], [4, 5, 6]]])

>>> arr2d.shape()

(2, 3)

>>> arr3d.ndim()

2

>>> arr3d.dtype()

dtype('int32')

>>> arr2d.size()

6

## פעולות על ndarray

ברשימות רגילות, פעולות אריתמטיות, כמו חיבור וכפל, רק מגדילות את הרשימה. לעומת זאת פעולות אריתמטיות על ndarray מפעיל את הפעולה לכל איבר בנפרד בהתאמה.

>>> np.array([2, 4, 5]) + np.array([7, 3, 2])

array([9, 7, 7])

>>> np.array([2, 4, 5]) \* np.array([7, 3, 2])

array([14, 12, 10])

* np.dot(<np.array>, <np.array>) - מקבלת שני ndarray ומחזירה את המכפלה הפנימית שלהם. עבור שני מערכים חד ממדיים פעולה זו שקולה למכפלת וקטורים, ועבור שני מערכים דו-ממדיים פעולה זו שקולה להכפלת מטריצות.
* np.mean(<np.array>) – מקבלת וקטור ומחזירה את הממוצע של כל הרכיבים בו.
* np.linalg.inv(<np.array>) - מקבלת מערך דו-ממדי (מטריצה) ומחזירה את המטריצה ההופכית שלה, כך שאם יכפילו בין שניהם תתקבל מטריצת היחידה.