Python Caesar Cipher

D. Leeuw

27 mei 2025

1 Over dit Document

1.1 Leerdoelen

1. Het doel is om vertrouwd te raken met de built in functies van Python

1.2 Voorkennis

2 ASCII-tabel

Computers zijn rekenaars, ze werken alleen met getallen. Mensen werken graag met tekst. Om een computer om te laten gaan met tekst moet tekst voor de computer omgezet worden in getallen. Als elk programma andere getallen gebruikt voor deze omzetting, dan wordt het een zooitje. We hebben een standaard vertaal tabel nodig om tekst om te zetten in getallen. Eén van deze tabellen is de ASCII-tabel. ASCII staat voor American Standard Code for Information Interchange en dat is precies wat het doet. Het is een standaard om informatie uitwisseling te vergemakkelijken.

	000	001	010	011	100	101	110	111
...0000	NUL	DLE	SPC	0	@	P	‘	p
...0001	SOH	DC1	!	1	A	Q	a	q
...0010	STX	DC2	"	2	B	R	b	r
...0011	ETX	DC3	#	3	C	S	c	s
...0100	EOT	DC4	$	4	D	T	d	t
...0101	ENQ	NAK	%	5	E	U	e	u
...0110	ACK	SYN	&	6	F	V	f	v
...0111	BEL	ETB	’	7	G	W	g	w
...1000	BS	CAN	(	8	H	X	h	x
...1001	HT	EM	)	9	I	Y	i	y
...1010	LF	SUB	*	:	J	Z	j	z
...1011	VT	ESC	+	;	K	[	k	{
...1100	FF	FS	,	<	L	\	l	|
...1101	CR	GS	-	=	M	]	m	}
...1110	SO	RS	.	>	N	^	n	~
...1111	SI	US	/	?	O	_	o	DEL

Om characters om te zetten naar een binaire code nemen we een character, zoeken deze op in te tabel. De kolom geeft de eerste 3 bits en de rij geeft de tweede 4 bits, zo komen we op 7-bits:

A 100 0001 = 65

a 110 0001 = 97

0 011 0000 = 48+0 = 48

9 011 1001 = 48+9 = 57

Het alfabet is op volgorde, dus B is 66 en b is 98 etc.

3 Andere vertaaltabellen

De ASCII-tabel is één van de oudste en was een van de meest gebruikte conversie-tabellen, zolang als computers voornamelijk door Engelstaligen werden gebruikt. Met de karakters van bijvoorbeeld de Europese of Slavische talen werd het allemaal wat lastiger en was het niet mogelijk om al deze karakters te vangen in 7-bits. Er ontstonden dan ook nadere conversie tabellen zoals UTF (Unicode Transformation Format) in 8-bits, 16-bits of zelfs 32-bits. Andere standaarden zijn de ISO 8859-1 en de Windows-1251 (ANSI Latin 1 of ANSI) standaard. Als we ook de Aziatische karakters willen kunnen weergeven dan redden we het ook niet met 8-bits, daar hebben we de 16 en 32-bits varianten voor.

4 Python conversie functies

Python heeft builtin functies die nummers kunnen omzetten naar characters en characters naar nummers. Deze functies kan je vinden op o.a. de site van W3Schools: https://www.w3schools.com/python/python_ref_functions.asp. Zoek de functie op die je nodig hebt om een character om te zetten naar een number en de functie die het omgekeerde doet. Test de functies in een script.

5 Caesar Cipher

Tijdens de security lessen hebben jullie in hoofdstuk 8 encryptie behandeld gekregen. Daarin werd de Caesar Cipher behandeld. Het verschuiven van tekst in het alfabet om zo de tekst onleesbaar te maken. Je zou de letters één plaat kunnen opschuiven zodat een a een b wordt en een c een d etc. Stel bijvoorbeeld dat we het bericht hebben 'Dit is een geheim bericht' en we schuiven alle letters één plaats op dan wordt het ge-encrypte bericht 'Eju jt

ffo hfjn cfsjdiu'. In Caesar zijn tijd werd dit met de hand gedaan, wij hebben computers en Python, dus gaan we dit automatiseren.

We hebben gezien dat de reeks nummers voor uppercase letters een aparte reeks is van de reeks met lowercase letters. Daar moeten we rekening mee houden. Ook als we bijvoorbeeld de letter z één plaats opschuiven dan moet het de a worden en niet { zoals dat volgens de ASCII-tabel zou worden. Om te weten te komen of een variabele uppercase of lowercase is kunnen we `islower` en `isupper` gebruiken. Dat ziet er in Python code zo uit:

```
chars = [ 'a', 'A', 'B', 'b' ]
for char in chars:
    if char.islower():
        print(f"{char} is lowercase")
    elif char.isupper():
        print(f"{char} is uppercase")
```

Voor getallen is er `isnumeric`.

Met deze voorkennis en de informatie uit het boek Security van Boris Sondagh maak je een script dat een substitutie uitvoert (encryptie). Gebruik een variabele voor het aantal keer dat er geschoven moet worden in het alfabet. Kies voor deze variabele een waarde tussen 1 en 10 om de code niet te complex te maken. Denk erom dat je na de Z weer bij A moet beginnen, na z weer bij a en na 9 komt weer 0. Leestekens veranderen niet.