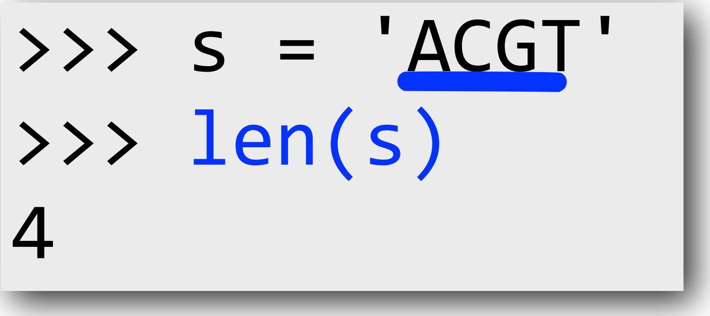
String definitions

String S is a finite sequence of characters

Characters are drawn from alphabet Σ

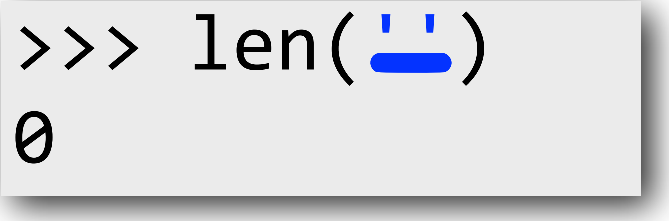
Usually, $\Sigma = \{\, A, C, G, T \,\}$

$|S|$ = number of characters in S



```
>>> s = 'ACGT'
>>> len(s)
4
```

ε is "empty string" $|\varepsilon| = 0$



```
>>> len('')
0
```

String definitions

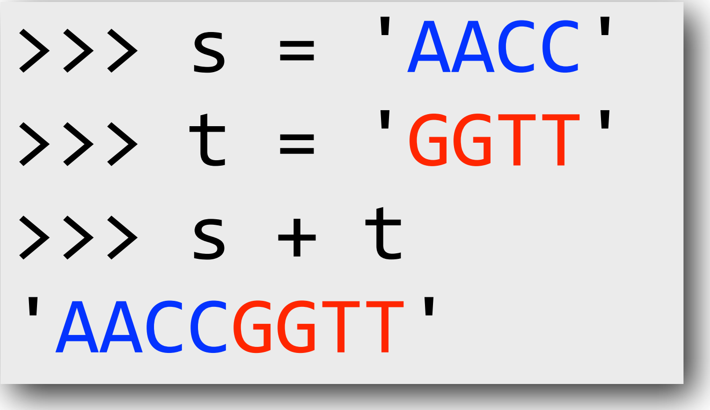
Positions within a string *S* are referred to with *offsets*

```
>>> s = 'ACGT'
          123
>>> s[0]
'A'
>>> s[2]
'G'
```

Leftmost offset = 0

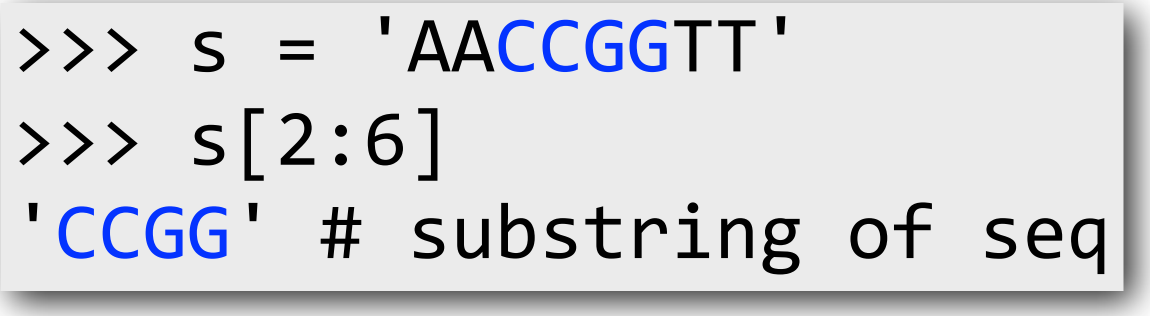
String definitions

Concatenation of *S* and *T* = characters of *S* followed by characters of *T*



```
>>> s = 'AACC'
>>> t = 'GGTT'
>>> s + t
'AACCGGTT'
```

Substring of *S* is a string occuring inside *S*



```
>>> s = 'AACCGGTT'
>>> s[2:6]
'CCGG' # substring of seq
```

String definitions

Prefix of *S* is a substring starting at the beginning of *S*

```
>>> s = 'AACCGGTT'
>>> s[0:6]
'AACCGG' # prefix
>>> s[:6] # same as above
'AACCGG'
```

Suffix is substring ending at end of *S*

```
>>> s = 'AACCGGTT'
>>> s[4:8]
'GGTT' # suffix
>>> s[4:] # like s[4:len(s)]
'GGTT'
>>> s[-4:] # like s[len(s)-4:len(s)]
'GGTT'
```