



# ***PARSING HTTP REQUEST***

## **LENGUAJES DE PROGRAMACIÓN**

Tarik Saghouani Ben-Khalek

Pablo Sánchez Gómez

Eva María Hoyo de la Cruz

José Ignacio Manso LLanes

## Contenido

Introducción .....	3
Entrega .....	4
<i>GET</i> .....	5
<i>POST</i> .....	6
Captura de salidas .....	7
ANEXO I .....	8

## Introducción

Esta práctica trata de parsear dos peticiones HTTP, una GET y otra POST. El protocolo HTTP es el protocolo utilizado en la Web y es un protocolo de texto. Se pide al alumno que construya un diccionario con los parámetros y valores que vienen en las peticiones.

### ¿Qué es el HTTP?

El Protocolo de transferencia de hipertexto (HTTP) está diseñado para permitir las comunicaciones entre clientes y servidores.

HTTP funciona como un protocolo de solicitud-respuesta entre un cliente y un servidor.

Un navegador web puede ser el cliente, y una aplicación en una computadora que aloja un sitio web puede ser el servidor.

Ejemplo: un cliente (navegador) envía una solicitud HTTP al servidor; entonces el servidor devuelve una respuesta al cliente. La respuesta contiene información de estado sobre la solicitud y también puede contener el contenido solicitado.

Nota que en el método GET el query string va en forma de pares nombre/valor en la url, mientras que el método POST la query string aparece en el cuerpo.

A partir de este string, correspondiente a una petición GET, construir el diccionario a continuación.

```
'GET /echo?message=%22Hello%20World!%22 HTTP/1.1\r\nHost:
127.0.0.1:5005\r\nConnection:
keep-alive\r\nUpgrade-Insecure-Requests: 1\r\nUser-Agent: Mozilla/5.0 (X11; Linux x86_64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/80.0.3987.116 Safari/537.36\r\nSec-
Fetch-Dest:
document\r\nAccept:
text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8,
application/signedexchange;
v=b3;q=0.9\r\nSec-Fetch-Site: none\r\nSec-Fetch-Mode: navigate\r\nSec-Fetch-User:
?1\r\nAccept-Encoding: gzip, deflate, br\r\nAccept-Language: es-ES,es;q=0.9,en-
GB;q=0.8,en;q=0.7\r\n\r\n'
{
'Method': 'GET',
'Url': '/echo?message=%22Hello%20World!%22',
'Version': HTTP/1.1,
'Host': '127.0.0.1:5005',
'Connection': 'keep-alive',
'Upgrade-Insecure-Requests': '1',
'User-Agent': 'Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/80.0.3987.116 Safari/537.36',
'Sec-Fetch-Dest': 'document',
'Accept': 'text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,
/*/*;q=0.8,applicatio
n/signed-exchange;v=b3;q=0.9',
'Sec-Fetch-Site': 'none',
'Sec-Fetch-Mode': 'navigate',
```

```
'Sec-Fetch-User': '?1',
'Accept-Encoding': 'gzip, deflate, br',
'Accept-Language': 'es-ES,es;q=0.9,en-GB;q=0.8,en;q=0.7'
'Params' : {
'message':%22Hello%20World!%22
}}
```

A partir de la siguiente petición POST construir el diccionario a continuación.

```
'POST /echo HTTP/1.1\r\nHOST: 127.0.0.1:5005\r\ncontent-type:
application/x-www-form-urlencoded\r\ncontent-length: 23\r\n\r\nmessage=Hello
World!!!!'
{'method':'POST',
'Url': '/echo',
'Version': HTTP/1.1,
'HOST': '127.0.0.1:5005',
'content-type': 'application/x-www-form-urlencoded',
'content-length': '23',
'params': {
'message': 'Hello World!!!!'
}
}
```

En esta práctica se nos pedía dividir dos solicitudes http una POST y otra GET mediante el uso de los diccionarios.

## Entrega

Esta entrega consta de varias partes, por un lado, el código creado a partir de las indicaciones que se encuentran en el enunciado/introducción de la práctica y que se puede encontrar en este mismo documento (ANEXO I), un archivo con dicho código y esta memoria explicando lo que hemos hecho.

Para que la explicación del código sea más comprensible la hemos dividido en dos apartados, uno para cada parte importante.

## GET

Primero introducimos el mensaje GET en una variable:

```
textoget = 'GET /echo?message=%22Hello%20World!%22 HTTP/1.1\r\nHost: 127.0.0.1:5005\r\nConnection: keep-alive\r\nUpgrade-Insecure-Requests: 1\r\nUser-Agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/80.0.3987.116 Safari/537.36\r\nSec-Fetch-Dest: document\r\nAccept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8,application/signedexchange;v=b3;q=0.9\r\nSec-Fetch-Site: none\r\nSec-Fetch-Mode: navigate\r\nSec-Fetch-User: ?1\r\nAccept-Encoding: gzip, deflate, br\r\nAccept-Language: es-ES;es;q=0.9,en-GB;q=0.8,en;q=0.7\r\n\r\n'
```

Luego hacemos una primera división con la función `.split`, esta separa cada elemento y lo mete en una posición de ella cuando encuentra un espacio:

```
subtextoget = textoget.split()
```

A continuación, volvemos a dividir la posición "1" de la lista con la misma función, pero esta vez solo si se encuentra un "=":

```
mensajeget = subtextoget[1].split('=')
```

Después creamos una variable para poder sumar en ella los string que pertenecen al contenido del mensaje sin incluir el primer elemento de la lista.

```
mensajegetimp = ''
```

```
for i in mensajeget:
```

```
    if i != mensajeget[0]:
```

```
        mensajegetimp = mensajegetimp + i
```

Finalmente creamos el diccionario con todos los campos necesarios y emparejamos cada elemento de la lista que habíamos dividido previamente con su campo del diccionario correspondiente, el contenido del mensaje lo tenemos en la variable `mensajegetimp` por ellos está la asignamos a `"message":`:

```
diccionarioget = {  
    'method': subtextoget[0],  
    'Url': subtextoget[1],  
    'Version': subtextoget[2],  
    'Host': subtextoget[4],  
    'Connection': subtextoget[6],  
    'Upgrade-Insecure-Requests': subtextoget[8],  
    'User-Agent': subtextoget[10] + " " + subtextoget[11] + " " + subtextoget[12] + "  
    " + subtextoget[13] + " " + subtextoget[14] + " " + subtextoget[15] + " " + subtextoget[16] + "  
    " + subtextoget[17] + " " + subtextoget[18] + " " + subtextoget[19],  
    'Sec-Fetch-Dest': subtextoget[21],  
    'Accept': subtextoget[23],  
    'Sec-Fetch-Site': subtextoget[25],  
    'Sec-Fetch-Mode': subtextoget[27],  
    'Sec-Fetch-User': subtextoget[29],
```

```

'Accept-Encoding': subtextoget[31]+" "+subtextoget[32]+" "+subtextoget[33],
'Accept-Language': subtextoget[35],
'Params': {

    'message': [mensajegetimp,]
}

```

## POST

Al igual que en el get introducimos la información del mensaje en una variable:

```

textopost = 'POST /echo HTTP/1.1\r\nHOST: 127.0.0.1:5005\r\ncontent-type:application/x-www-form-urlencoded\r\ncontent-length: 23\r\n\r\nmessage=Hello World!!!'

```

Primero dividimos textopost con la función split() para separarlos por espacios.

```

subtextopost = textopost.split()

```

A continuación, separo el 6 elemento de la lista "subtextopost" a partir del "=":

```

mensajepost = subtextopost[8].split('=')

```

Luego creamos la variable mensajepostimp en la que introducimos el contenido del mensaje que en este caso es hello world!!!

```

mensajepostimp=""
mensajepostimp=mensajepostimp+mensajepost[1] + " "
j=0

```

```

for i in subtextopost:

```

```

    if j > 8:

```

```

        mensajepostimp=mensajepostimp+i

```

```

    j=j+1

```

Finalmente creamos el diccionario y le asignamos a cada campo su valor correspondiente:

```

diccionariopost = {
    'method': subtextopost[0],
    'Url': subtextopost[1],
    'Version': subtextopost[2],
    'Host': subtextopost[4],
    'content-type': contenttype[1],
    'content-length': subtextopost[7],

    'Params': {
        'message': [mensajepostimp]
    }
}

```

## Captura de salidas

Esta salida corresponde con la salida en la terminación del entorno de programación, en nuestro caso PyCharm

```
[{"GET", "/echo?message=%22Hello%20World!%22", "HTTP/1.1", "Host:", "127.0.0.1:5005", "Connection:", "keep-alive", "Upgrade-Insecure-Requests:", "1", "User-Agent:", "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/80.0.3987.116 Safari/537.36", "Sec-Fetch-Dest:", "document", "Accept:", "text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8,application/signedexchange;v=b3;q=0.9", "Sec-Fetch-Site:", "none", "Sec-Fetch-Mode:", "navigate", "Sec-Fetch-User:", "?1", "Accept-Encoding:", "gzip, deflate, br", "Accept-Language:", "es-ES;es;q=0.9,en-GB;q=0.8,en;q=0.7"}  
{"method": "GET", "url": "/echo?message=%22Hello%20World!%22", "Version": "HTTP/1.1", "Host": "127.0.0.1:5005", "Connection": "keep-alive", "Upgrade-Insecure-Requests": "1", "User-Agent": "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/80.0.3987.116 Safari/537.36", "Host": "127.0.0.1:5005", "content-type": "application/x-www-form-urlencoded", "content-length": "23", "message": "Hello", "World!!!!"}  
{"method": "POST", "url": "/echo", "Version": "HTTP/1.1", "Host": "127.0.0.1:5005", "content-type": "application/x-www-form-urlencoded", "content-length": "23", "Params": {"message": ["Hello World!!!!"]}}  
Process finished with exit code 0
```

Haciendo zoom para que se aprecie mejor:

```
'GET', '/echo?message=%22Hello%20World!%22', 'HTTP/1.1', 'Host:', '127.0.0.1:5005', 'Connection:',  
'method': 'GET', 'url': '/echo?message=%22Hello%20World!%22', 'Version': 'HTTP/1.1', 'Host': '127  
'POST', '/echo', 'HTTP/1.1', 'HOST:', '127.0.0.1:5005', 'content-type:application/x-www-form-urle  
'method': 'POST', 'Url': '/echo', 'Version': 'HTTP/1.1', 'Host': '127.0.0.1:5005', 'content-type'  
  
rocess finished with exit code 0
```

Como no se apreciaba todo el texto de la salida, adjuntamos a continuación una copia:

```
['GET', '/echo?message=%22Hello%20World!%22', 'HTTP/1.1', 'Host:', '127.0.0.1:5005',  
'Connection:', 'keep-alive', 'Upgrade-Insecure-Requests:', '1', 'User-Agent:', 'Mozilla/5.0', '(X11;',  
'Linux', 'x86_64)', 'AppleWebKit/537.36', '(KHTML, like Gecko)', 'Chrome/80.0.3987.116',  
'Safari/537.36', 'Sec-Fetch-Dest:', 'document', 'Accept:',  
'text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8,a  
pplication/signedexchange;v=b3;q=0.9', 'Sec-Fetch-Site:', 'none', 'Sec-Fetch-Mode:', 'navigate',  
'Sec-Fetch-User:', '?1', 'Accept-Encoding:', 'gzip, deflate, br', 'Accept-Language:', 'es-  
ES;es;q=0.9,en-GB;q=0.8,en;q=0.7']  
{'method': 'GET', 'url': '/echo?message=%22Hello%20World!%22', 'Version': 'HTTP/1.1', 'Host':  
'127.0.0.1:5005', 'Connection': 'keep-alive', 'Upgrade-Insecure-Requests:', '1', 'User-Agent':  
'Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko)  
Chrome/80.0.3987.116 Safari/537.36', 'Sec-Fetch-Dest': 'document', 'Accept':  
'text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8,a  
pplication/signedexchange;v=b3;q=0.9', 'Sec-Fetch-Site': 'none', 'Sec-Fetch-Mode': 'navigate',  
'Sec-Fetch-User': '?1', 'Accept-Encoding': 'gzip, deflate, br', 'Accept-Language': 'es-  
ES;es;q=0.9,en-GB;q=0.8,en;q=0.7', 'Params': {'message': ['%22Hello%20World!%22']}}  
['POST', '/echo', 'HTTP/1.1', 'HOST:', '127.0.0.1:5005', 'content-type:application/x-www-form-  
urlencoded', 'content-length:', '23', 'message=Hello', 'World!!!!']  
{'method': 'POST', 'url': '/echo', 'Version': 'HTTP/1.1', 'Host': '127.0.0.1:5005', 'content-type':  
'application/x-www-form-urlencoded', 'content-length': '23', 'Params': {'message': ['Hello  
World!!!!']}}
```

## ANEXO I

```
textoget = 'GET /echo?message=%22Hello%20World!%22 HTTP/1.1\r\nHost:
127.0.0.1:5005\r\nConnection: keep-alive\r\nUpgrade-Insecure-Requests:
1\r\nUser-Agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/80.0.3987.116 Safari/537.36\r\nSec-Fetch-Dest:
document\r\nAccept:
text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*
/*;q=0.8,application/signedexchange;v=b3;q=0.9\r\nSec-Fetch-Site:
none\r\nSec-Fetch-Mode: navigate\r\nSec-Fetch-User: ?1\r\nAccept-Encoding:
gzip, deflate, br\r\nAccept-Language: es-ES,es;q=0.9,en-
GB;q=0.8,en;q=0.7\r\n\r\n'
textopost = 'POST /echo HTTP/1.1\r\nHOST: 127.0.0.1:5005\r\ncontent-
type:application/x-www-form-urlencoded\r\ncontent-length:
23\r\n\r\nmessage=Hello World!!!!'

subtextoget = textoget.split()
print(subtextoget)

mensajeget = subtextoget[1].split('=')
mensajegetimp=""
for i in mensajeget:
    if i != mensajeget[0]:
        mensajegetimp=mensajegetimp+i
diccionarioget = {
    'method': subtextoget[0],
    'Url': subtextoget[1],
    'Version': subtextoget[2],
    'Host': subtextoget[4],
    'Connection': subtextoget[6],
    'Upgrade-Insecure-Requests': subtextoget[8],
    'User-Agent': subtextoget[10]+" "+subtextoget[11]+" "+subtextoget[12]+"
"+subtextoget[13]+" "+subtextoget[14]+" "+subtextoget[15]+"
"+subtextoget[16]+" "+subtextoget[17]+" "+subtextoget[18]+"
"+subtextoget[19],
    'Sec-Fetch-Dest': subtextoget[21],
    'Accept': subtextoget[23],
    'Sec-Fetch-Site': subtextoget[25],
    'Sec-Fetch-Mode': subtextoget[27],
    'Sec-Fetch-User': subtextoget[29],
    'Accept-Encoding': subtextoget[31]+" "+subtextoget[32]+"
"+subtextoget[33],
    'Accept-Language': subtextoget[35],
    'Params': {

        'message': [mensajegetimp,]
    }
}

print(diccionarioget)

subtextopost = textopost.split()
contenttype = subtextopost[5].split(':')
print(subtextopost)

mensajepost = subtextopost[8].split('=')
```



```

mensajepostimp=""
mensajepostimp=mensajepostimp+mensajepost[1] + " "
j=0
for i in subtextopost:
    if j > 8:
        mensajepostimp=mensajepostimp+i
        j=j+1
diccionariopost = {
    'method': subtextopost[0],
    'Url': subtextopost[1],
    'Version': subtextopost[2],
    'Host': subtextopost[4],
    'content-type': contenttype[1],
    'content-length': subtextopost[7],

    'Params': {
        'message': [mensajepostimp]
    }
}

print(diccionariopost)

```