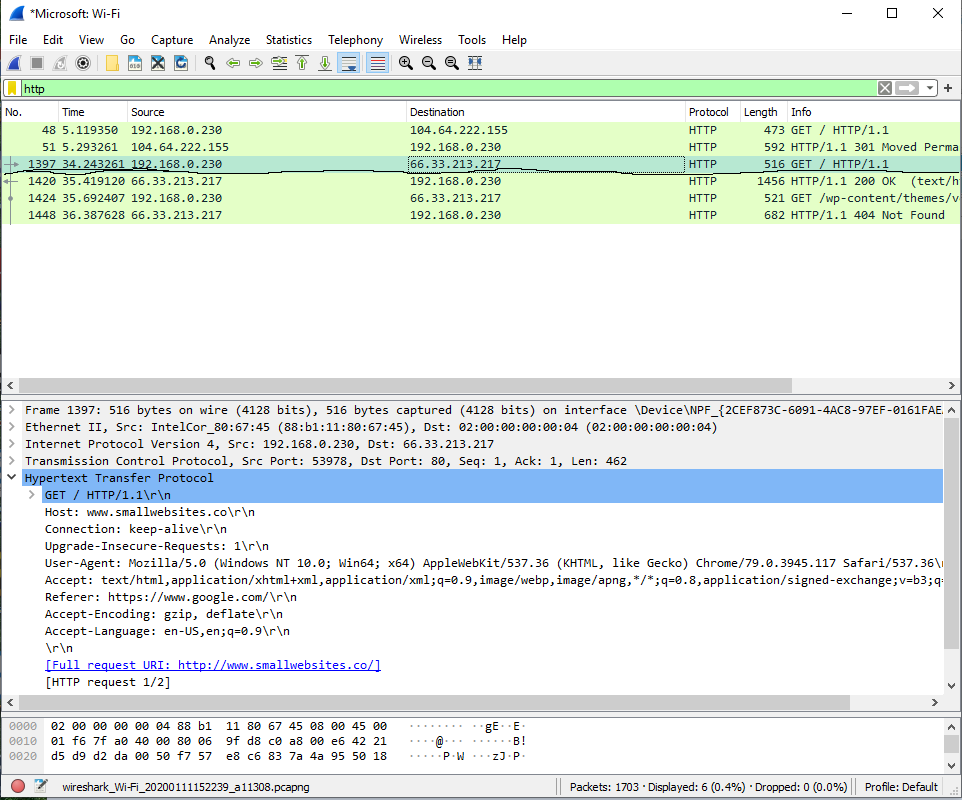
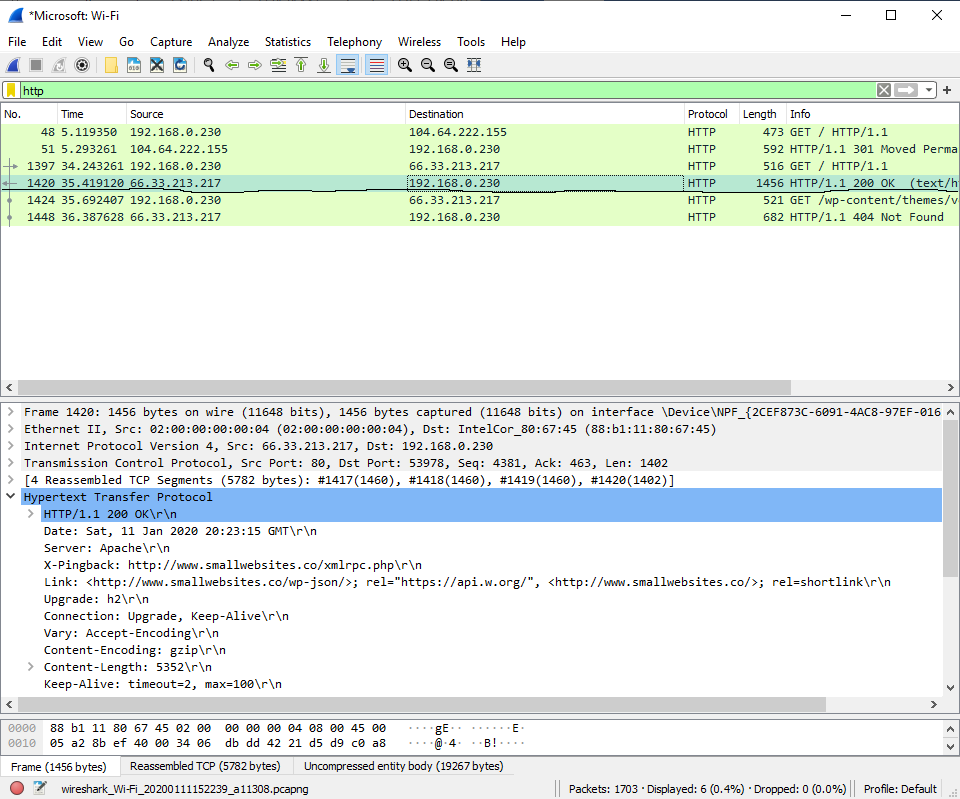
**COMPSCI 4C03 Assignment 1**

**Question 1: Capturing an HTTP Message**

**1-A: Add the Wireshark screenshot and highlight the two messages in it.**

**HTTP GET message:**

****

**HTTP OK Message:**

**1-B: Write the complete URL of the webpage referred in GET message. Which fields/lines of GET message can be used to acquire the complete URL?**

The complete URL of the webpage is: <http://www.smallwebsites.co/> and can be found under the “Hypertext Transfer Protocol section of the message body, under “Full request URI”.

**1-C: Assume the content of a GET message is provided to you in the form of a string. Write the pseudocode or algorithm to parse this string to extract the complete URL out of this string.**

We can easily use regex libraries to parse out the URL, since it is always encapsulated in

[Full request URI: URL GOES HERE] within the message body. The pseudocode using pythons re library is as follows:

Import re

text = GET\_message\_text

*# returns [Full request URI: URL GOES HERE]*

encased\_URL = re.search(“\[Full request URI: .\*\]”, text).group(0)

*# will parse out URI from above*

URL=encased\_URL[19:-1]