

## 2. Laboratorium, TliM, HTTP

### Ćwiczenie 0

Uruchom Firefox/IceWeasel. Zainstaluj rozszerzenie LiveHTTP Headers (menu: Narzędzia/Dodatki).

- Otwórz dowolną stronę.
- Zobacz nagłówki HTTP: Prawy przycisk na zawartości strony, wybierz: Pokaż informacje o stronie, następnie zakładka Headers.
- Otwórz panel wyświetlający nagłówki: Alt-L (albo menu: Widok/Panel Boczny/LiveHTTPHeaders).
- Wczytaj dowolną stronę, prześledź informacje wyświetlane w bocznym panelu.

### Ćwiczenie 1

Używając telnet zobacz jaki jest nagłówek oraz zawartość strony pod adresem: <http://home.agh.edu.pl/~wojnicki/>.

```
telnet home.agh.edu.pl 80
      GET / ~wojnicki/ HTTP/1.0
      'enter'
      'enter'
```

### Ćwiczenie 2

Czy dostęp do zasobu <http://home.agh.edu.pl/~wojnicki/> może być realizowany poprzez protokół HTTP 1.1, czy tylko HTTP 1.0? Przetestuj.

```
telnet home.agh.edu.pl 80
      GET / ~wojnicki/ HTTP/1.0
      HOST: home.agh.edu.pl
      'enter'
      'enter'
```

### Ćwiczenie 3

Sprawdź, jaki typ odpowiedzi zwraca serwer przy zadaniu dostępu do zasobu <http://home.agh.edu.pl/~wojnicki/> (Uwaga: brak '/' na końcu adresu jest ważny!). Co znaczy ten kod?

400 - błąd zapytania

### Ćwiczenie 4

Używając metody GET i wersji protokołu 1.1 zapytaj wyszukiwarkę Google o wyrazy: Informatyka, wiedząc że wyrazy powinny być przekazane jako wartość atrybutu 'q' (np. zapytanie o hasło pies powinno wyglądać następująco: <http://www.google.pl/search?q=pies&nord=1> )

- Użyj telnet.

Jeżeli google zamyka połączenie (timeout) użyj nc zamiast telnet, oraz umieść zapytanie w pliku:

```
nc google.com 80 < plik_z_zapytaniem
```

albo:

```
cat plik_z_zapytaniem | nc google.com 80
```

Podpowiedź: przeczytaj man nc

b) Użyj przeglądarki (np. firefox).

```
telnet www.google.com 80
GET /search?q=Informatyka&nord=1 HTTP/1.1
HOST: www.google.com
'enter'
'enter'
```

## Ćwiczenie 5

Zbadaj co zwraca HEAD dla `http://home.agh.edu.pl/wojnicki/`. Porównaj z HEAD dla `http://home.agh.edu.pl/~wojnicki/`. Zwróć uwagę na nagłówek Location oraz status odpowiedzi serwera.

```
telnet home.agh.edu.pl 80
HEAD / ~wojnicki/ HTTP/1.0
HOST: home.agh.edu.pl
'enter'
'enter'
```

## Ćwiczenie 6

Zbadaj co zwraca GET dla `http://home.agh.edu.pl/~wojnicki/pic/me.jpg`. Jaki jest typ danych zasobu?

TE CHIŃSKIE ZNACZKI

## Ćwiczenie 7

Ile żądań (zapytań) można przesłać za pomocą jednego połączenia w przypadku HTTP/1.0 a ile w przypadku HTTP/1.1? Wypróbuj na dowolnie wybranym serwerze i zasobie.

1.0 tylko jedno

1.1 więcej, ale trzeba się śpieszyć bo okno szybko się zamyka

## Ćwiczenie 7.5

Pobierz zasób `http://paw.agh.edu.pl` za pomocą HTTP/1.0. Operację ponów dla HTTP/1.1. Czy odpowiedzi są takie same? Dlaczego?

Nie w 1.1 jest zawartość body

## Ćwiczenie 8

Używając nc (netcat) i przekierowania standardowego wejścia/wyjścia pobierz zawartość witryny `http://home.agh.edu.pl/~wojnicki/` i umieść w pliku `w1.html`.

Zapytanie HTTP można umieścić w pliku:

```
nc ..... < plik_z_zapytaniem
```

ładź też użyć echo oraz mechanizmu pipe:

```
echo -e 'GET .... \r\n\r\n' | nc ....
```

```
echo -e 'GET/ wojnicki/HTTP/1.0\r\n\r\n' >plik
```

```
nc home.agh.edu.pl 80 < plik
```

## Ćwiczenie 8.5

Nadobowiązkowe.

Netcat może również nasłuchiwać i odbierać połączenia na wskazanym porcie. Dzięki temu w prosty sposób można np. przysyłać dowolne informacje pomiędzy węzłami w sieci. Np. chcąc przesłać prosty komunikat, na węźle docelowym (charon.kis.agh.edu.pl) należy uruchomić netcat w trybie nasłuchiwania, powiedzmy na porcie 1234:

nc -l 1234 Na węźle nadającym, można przesłać np. komunikat: Hello:

echo Hello | nc charon.kis.agh.edu.pl 1234 Oczywiście nic nie stoi na przeszkodzie, aby przesłać tą metodą dowolne dane np. pliki.

## Ćwiczenie 9

Używając np. sed oraz wyrażeń regularnych usuń wszystkie znaczniki HTML z pliku w1.html, a wynik umieść w pliku w2.html.

Podpowiedzi:

- Niezbędny fragment składni edytora liniowego sed: sed 's/zrodlo/cel/g' znajduje ciąg znaków 'zrodlo' oraz zastępuje go ciągiem znaków 'cel'. Ciąg 'zrodlo' może być wyrażeniem regularnym.
- Utwórz wyrażenie regularne (man regex) pasujące do znacznika HTML: znak '<', dowolny ciąg znaków z wyjątkiem '>', znak '>'.
- Użyj w/w wyrażenia jako 'zrodlo' dla sed, 'cel' powinien być pustym łańcuchem znaków.

sed 's/<[^>]\*>/g' w1.html > w2.html

## Ćwiczenie 10

Wykonaj ćwiczenie 9 używając mechanizmu potoków (pipe):

echo .... | nc ... | sed ... | less

Gratulacje, właśnie zbudowałeś(aś) swoją własną przeglądarkę WWW

echo -e 'GET / wojnicki/ HTTP/1.0\r\n\r\n' | nc home.agh.edu.pl 80 | sed 's/<[^>]\*>/g' | less

Uwaga: W przypadku ćwiczenia 9 i 10 można dostać „dodatkowy punkt” - za rozwiązanie na zajęciach, albo nadesłanie rozwiązania mailem na adres prowadzącego nie później niż do północy następnego dnia po zajęciach. Jeden punkt może zaważyć na ocenie końcowej.

## Ćwiczenie 11

Istnieje wiele innych użytecznych narzędzi dla protokołu HTTP np. curl, wget. Poczytaj, sprawdź w działaniu.