GET /index.html HTTP/1.0 User-Agent: SuperBrowser v1.3

但当客户端向代理发送请求时,请求行中则包含完整的 URI。例如:

GET http://www.marys-antiques.com/index.html HTTP/1.0 User-Agent: SuperBrowser v1.3

为什么会有两种不同的请求格式,一种用于代理,另一种用于服务器呢?在原始的HTTP设计中,客户端会直接与单个服务器进行对话。不存在虚拟主机,也没有为代理制定什么规则。单个的服务器都知道自己的主机名和端口,所以,为了避免发送冗余信息,客户端只需发送部分URI即可,无需发送方案和主机(以及端口)。

代理出现之后,使用部分 URI 就有问题了。代理需要知道目标服务器的名称,这样它们才能建立自己与服务器的连接。基于代理的网关要知道 URI 的方案才能连接到FTP 资源和其他方案上去。HTTP/1.0 要求代理请求发送完整的 URI,解决了这个问题,但它为服务器请求保留部分 URI 的形式(已经有相当多的服务器都改为支持完整 URI 了)。8

因此,我们要将部分 URI 发送给服务器,将完整 URI 发送给代理。在显式地配置客户端代理设置的情况下,客户端就知道要发布哪种类型的请求了。

- (1) 没有设置客户端使用代理时,它会发送部分 URI (参见图 6-15a)。
- (2) 设置客户端使用代理时,它会发送完整 URI (参见图 6-15b)。

6.5.2 与虚拟主机一样的问题

代理"缺少方案/主机/端口"的问题与虚拟主机 Web 服务器面临的问题相同。虚拟主机 Web 服务器会在很多 Web 站点间共享同一个物理 Web 服务器。包含部分 URI(比如/index.html)的请求到达时,虚拟主机 Web 服务器需要知道目的 Web 站点的主机名(更多信息请参见 5.7.1 节和 18.2 节)。

尽管它们出现的问题相似,但解决方法却有所不同:

- 显式的代理要求在请求报文中使用完整 URI 来解决这个问题,
- 虚拟主机 Web 服务器要求使用 Host 首部来承载主机和端口信息。

注 8: 现在,HTTP/1.1 要求服务器为代理请求和服务器请求都提供完整的 URI 处理,但实际上,很多已部署的服务器仍然只接受部分 URI。