

网络编程课程实验

(实验报告见模板，最终以.pdf 形式提交)

第一次实验

基于 UDP 的简易聊天工具

- 要求：编写一个程序，能够获取键盘数据，并将数据发送给指定对象；同时，在接收到数据时，能够将其显示。
- 提示：可提供一个功能菜单，让用户选择发送数据或者接收数据。若选择发送，则调用发送数据函数；若选择接收，则调用接收数据函数。

基于 TCP 的文件下载

- 客户端：向服务器发送一个用户输入的文件名，若文件存在，则下载文件，若不存在，则提示用户。
- 服务端：接收客户端数据，识别并显示客户端所请求下载的文件名，若文件存在，则发送给客户端，若不存在，则提示用户。

基于 TCP 的网页请求

- 需求：模拟一个 TCP Client，完成基于 Https 的网页请求；
- 能够正确构造请求(包含加密)，并能够解析网页；
- 网页最终以 html 的形式保存
- 目标网站不少于三个

第二次实验

具备同时收发功能的 udp 聊天工具

- 通过多线程的方式实现一个在发送数据的同时也能够自动接收数据 UDP 聊天器。

编写一个包含多个售票窗口的售票程序。

- 假设有票的编号为 1-100，有 4 个窗口可以同时售票，车票按顺序依次售出。1 号和 2 号窗口有 20 人排队，3 号和 4 号窗口有 30 人排队，每售出一张票，显示售票的窗口号、票号以及余票数量。

文件拷贝（多进程方式）

- 通过多进程的方式实现拷贝文件夹下的文件(文件夹嵌套不考虑)；在拷贝过程中显示拷贝进度。
- 提示：通过 os 模块的 listdir 获取到要拷贝的文件夹内的文件名列表，并以此来创建相应的子进程，每个文件一个进程。
- 父进程通过获取子进程的消息并计算已经完成的文件拷贝数来计算进度并显示。

第三次实验

半自动爬虫

- 利用正则表达式，提取百度贴吧某一热门帖(回帖数不小于 200)的发站人、发帖时间、发帖内容、楼层数、点赞数等。

解析工具的使用

- 利用 BeautifulSoup、XPath 等解析工具，解析并提取任一小说网站中某一长篇小说的内容，并按章节存储。

豆瓣电影数据提取

- 提取豆瓣电影(或图书)中 Top300 的相关内容，并将电影(或图书)的名称、出版时间、宣传图、热门评论等必要的信息提取出来，并存储到本地。

第四次实验

新浪微博数据提取 (AJAX)

- 利用 AJAX 爬取新浪微博中任一大 V 的微博，需要提取如点赞数、评论数、评论内容，以及对评论的回复等相关数据。

网购平台数据提取 (Selenium)

- 利用 Selenium 爬取京东或者淘宝网站中，任一大类商品(如手机、食品)的商品相关信息(如型号、价格、颜色、具体参数、商品图片等信息)

网易云音乐数据提取 (Scrapy 框架)

- 利用 Scrapy 框架，爬取网易云音乐所有热门华语男歌手的歌手名、专辑名、专辑信息、歌曲名等信息。