

# CS2006 — A Social News Navigator in *Python*

Chengyi Lin, Ross Apted and Ben Lovell

December 12, 2011

# Contents

<b>1</b>	<b>Running instructions</b>	<b>2</b>
<b>2</b>	<b>Introduction</b>	<b>3</b>
2.1	Aim . . . . .	3
2.2	Requirements: . . . . .	3
<b>3</b>	<b>Code description</b>	<b>4</b>
<b>4</b>	<b>Difficulties</b>	<b>5</b>
<b>5</b>	<b>Testing</b>	<b>6</b>
5.1	How we Tested . . . . .	6
<b>6</b>	<b>Conclusion</b>	<b>8</b>

## Part 1

# Running instructions

In order to run, run the executable `./main.py` file. If for some reason it is not executable, try using `chmod +x` or running it using python (`python main.py`). Whilst more comprehensive usage commands can be found by reading the readme, found in the README file, here are the basics:

- You can quit at any time whilst navigating with the ‘q’ key.
- You can navigate between stories with the arrow keys and press return to open and fold a story or comments
- If you wish to login to comment or upvote something, go into command mode with escape or ‘:’ (colon), and type “login”.
- A comment can be added using the ‘c’ key or by typing “comment” in to the command prompt.
- Over a story element the command “open” launches a web browser to view the link.

## Part 2

# Introduction

### 2.1 Aim

The aim of this Project is to build text-based browser that allows the user to navigate and use all the functionality of a social news site. We chose to use reddit, and have successfully implemented a solution that allows us to navigate and use the website.

### 2.2 Requirements:

#### Primary requirements

- First of all, the program does do what it should do, by being able to browse reddit.
- It retrieves new headlines and displays the relevant information in an informative manner (the story's title, votes, etc).
- More information can be retrieved about a story, in the form of its comments, by pressing the return key on an item and loading the comments on it, and so on with the subcomments.
- It is possible to login to the site, and to then post comments or replies to a story or comment. However, it is not possible to post a new submission to Reddit — mainly because the ability to do this now (since the creation of the practical) now requires the ability to pass a captcha. This is a little difficult, due to the fact that the program is a terminal program, and captchas require the ability to parse an image or audio file, and as such is no longer a trivial addition.

Other things for the project were also tackled:

#### Other Requirements

- Structuring the code to allow other news sites was accounted for with code structure accounting for it.
- Technically, the program can be used offline as long as it is not shutdown after downloading the data. However, you can only see loaded links offline. No serialisation takes places, and as such, will only work in a persistent session (`man screen`).
- Keyboard shortcuts were implemented to make the program easier to use and navigate.

All additional requirements were achieved.

## Part 3

# Code description

The program was separated into logically separate parts. This was to allow modularity in the scraping part and the user interface part. In theory, the user interface or scraper could be changed at any point to allow different systems for the interface and for the scraper.

On the one hand, there was the scraping part. This was modularised by using abstract classes and functions that a concrete scraper would implement. These abstract parts were split into two parts: the objects of which the scraper would populate (such as a comment and story and so on), and the functions modifying this data (such as get comments on a story, comment on a story and so on). These were implemented in **base.py** and **session.py**.

The concrete implementation for Reddit scraping can be found in **reddit\_api.py**. This implementation uses an external library which wraps the reddit api in a Python interface. This can be found here: [https://github.com/mellort/reddit\\_api](https://github.com/mellort/reddit_api)

On the other hand there was the user interface part. The UI was implemented using an ncurses based Python library called “Urwid”, which we used to create a navigatable terminal interface to present the different stories and so on. This referenced the reddit api interface, but only in ways that used the predefined functions — it would have been perfectly possible to switch out the code with another interface. This is coded in the **main.py**

## Part 4

# Difficulties

Certain issues were had when writing the code. For the most part, these were all caused (directly and indirectly) by our dealings with the Urwid user interface library for Python. As one might expect from quite a few open source (and sometimes proprietary) libraries, documentation was less than it could have been. The tutorial explains very little, and resorts to small one line comments explaining very little (of the kind oft found in first year practical reports). The reference documentation had some doc strings explaining init methods of the different objects. All of these explanations appeared to assume knowledge of how the library worked in the first place to understand.

The best documentation relevant to our needs was from a french blog, in which a haphazard explanation showed how to create a selectable list and to put an edit field inside of a footer element. This series of online tutorials can be found here: <http://www.nicosphere.net/series/urwid/>

Some of the ideas from this code were utilised in order to get a selectable list and to get an editable footer. It was soon found that to create stuff outside of the ecosystem of widgets provided in Urwid's default library then additional complexity was caused creating new widgets within Urwid. It often seemed that we were fighting *against* the library, rather than utilising it to create big and wonderful Pythonian things. Because of this, the code we have written using it is convoluted, and difficult to comprehend at first. Long lines passing many arguments to an arbitrary widget caused long lines that could not easily be broken down to under 80 characters without making the code messier and the fact that one function has to send a signal to another in order to string together different inputs was rather annoying. This could be improved with another, higher order function that recursively creates more functions, but time limits did not permit these unpragmatic pursuits for programming beauty.

As we severely underestimated the complexity of the Urwid system, originally only one person was assigned to doing the UI for Python. It was found out later that this code would take more than that to tackle it, and as such significant amount of man hours were spent simply writing/rewriting each other's code regarding it.

## Part 5

# Testing

### 5.1 How we Tested

- Login — Passed

---

```
< ↑ >    HP: WebOS, Enyo app framework goes open source
  4
< ↓ >
```

---

---

```
< ↑ >    HP's WebOS plan modeled after Red Hat's Fedora
  3
< ↓ >
```

---

Login Successfull!

- Login with incorrect username or password — Passed

---

```
< ↑ >    HP: WebOS, Enyo app framework goes open source
  4
< ↓ >
```

---

---

```
< ↑ >    HP's WebOS plan modeled after Red Hat's Fedora
  3
< ↓ >
```

---

Wrong user name or password!

- Open comment page of a story — Passed

< ↑ > 86 < ↓ >	Plan 9 – A distributed OS with a unified communication protocol and I/O driver model. Wow.
< ↑ > Comment < ↓ >	Plan9 is more or less dead. However, there is a very much alive fork called 9front. By user: Sukiari
< ↑ > Comment < ↓ >	see also Inferno, an Open-free/commercial OS inspired by Plan 9: <a href="http://www.vitanuova.com/inferno/index.html">http://www.vitanuova.com/inferno/index.html</a> By user: shavenwarthog
< ↑ > Comment < ↓ >	I find it fascinating to note how much of my current OS came from Plan 9:  Union FS /proc file system Namespaces and what's this... UTF8?!
	By user: railmaniac
< ↑ > Comments	As my OS professor told me:

- Close comment page of a story — Passed

< ↑ > 86 < ↓ >	Plan 9 – A distributed OS with a unified communication protocol and I/O driver model. Wow.
< ↑ > 19 < ↓ >	HP donates WebOS system code to open source developers
< ↑ > 203 < ↓ >	WebOS lives! HP decides to open source the platform
< ↑ > 4 < ↓ >	HP: WebOS, Enyo app framework goes open source
< ↑ > 3 < ↓ >	HP's WebOS plan modeled after Red Hat's Fedora
< ↑ > < ↓ >	-----Load More News-----

- Open the subcomments page of a comment — Passed

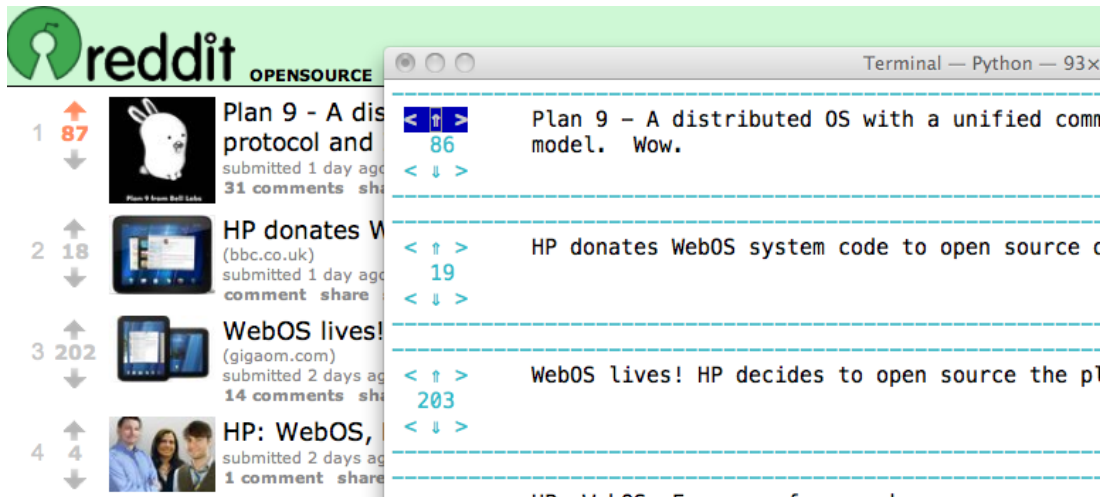


< ↑ > Comments < ↓ >	As my OS professor told me: Unix version 5 Unix version 7 Plan 9 Chapter 11 By user: jmknsd
< ↑ > Comments < ↓ >	i dont get it.. bankruptcy? By user: anthropomorphised
< ↑ > Comment < ↓ >	reply to reply OK!!!! By user: linchengyi
< ↑ > Comments < ↓ >	It's... a joke. By user: jmknsd
< ↑ > Comments < ↓ >	i don't get it. By user: pingu
< ↑ > Comment < ↓ >	bbbbbbbbbbbbbbbbbb By user: linchengyi
< ↑ >	sweet glenda.

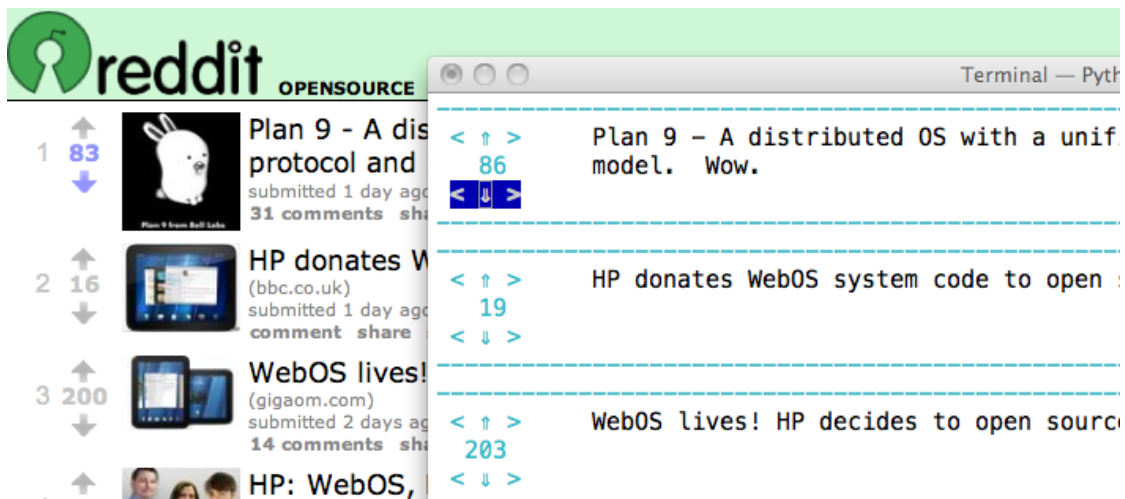
- Close the subcomments of comment — Passed

< ↑ > Comments < ↓ >	As my OS professor told me: Unix version 5 Unix version 7 Plan 9 Chapter 11 By user: jmknsd
< ↑ > Comment < ↓ >	sweet glenda. By user: moonhead
< ↑ > Comment < ↓ >	Bunny wabbit By user: koodough

- Up vote a story — Passed



- Down vote a story — Passed



- Up vote a comment or reply — Passed

**[+] jmknsd** 9分/分 1 day ago  
As my OS professor told me:  
Unix version 5  
Unix version 7  
Plan 9  
Chapter 11  
[permalink](#) [report](#) [reply](#)

---

**[+] anthropomorphised**  
i dont get it.. bankruptcy?  
[permalink](#) [parent report](#)

---

**[+] jmknsd** 0分/分 1 day ago  
It's... a joke.  
[permalink](#) [parent report](#)

---

**[+] linchengyi**  
bbbbbbbbbbbbb

---

< ↑ >  
**Comments**  
< ↓ >

As my OS professor told me:

Unix version 5

Unix version 7

Plan 9

Chapter 11  
By user: jmknsd

---

< ↑ >  
**Comments**  
< ↓ >

i dont get it.. bankruptcy?

By user: anthropomorphised

---

< ↑ >  
**Comments**  
< ↓ >

It's... a joke.

By user: jmknsd

---

- Down vote a comment or reply — Passed

[illegible]

- Vote without login — Passed

```
< 85 > Plan 9 - A distributed OS with a unified communication protocol and I/O
< ↓ > driver model. Wow.

-----

< ↑ > Plan9 is more or less dead. However, there is a very much alive fork
Comments called 9front.
< ↓ > By user: Sukiari

-----

< ↑ > see also Inferno, an Open-free/commercial OS inspired by Plan 9:
Comment
< ↓ > http://www.vitanuova.com/inferno/index.html
By user: shavenwarthog

-----

< ↑ > I find it fascinating to note how much of my current OS came from Plan
Comment 9:
< ↓ >
Error, please login first!
```

- Comment on a story — Passed

Some reasons as to why you...  
 Read his final report for a b...  
 nine.  
 He also has a blog where h...  
 First Week of Plan 9  
 Second Week of Plan 9  
 permalink parent report rep...  
 [-] anthropomorphised 2...  
 Thanks for the links. Sc...  
 Really not sure what to...  
 permalink parent report...  
 [-] linchengyi 1 分/分 5 ho...  
 aaaaaaaaaaaaaaaaaa...  
 permalink parent 編輯 刪...  
 [-] linchengyi 1 分/分 9 hours ago  
 comment on story okokokokokokok...  
 permalink 編輯 刪除 reply  
 [-] Ferrofluid 1 分/分 8 hours ago  
 Supported platforms : x86, x86-6...  
 which is very useful for people run...  
 this is true cloud computing  
 permalink report reply  
 [-] linchengyi 1 分/分 7 seconds ago  
 comment to story OK!!!!!!!!!!  
 permalink 編輯 刪除 reply

Terminal — bash — 93x31  
 Plan 9 - A distributed OS with a unified commu  
 model. Wow.  
 HP donates WebOS system code to open source de  
 WebOS lives! HP decides to open source the pla  
 HP: WebOS, Enyo app framework goes open source  
 HP's WebOS plan modeled after Red Hat's Fedora  
 -----Load More News-----  
 Comment Successfull! Please reload the comment page.

- Reply to a comment or a reply — Passed

[-] anthropomorphised 3 分  
 i dont get it.. bankruptcy...  
 permalink parent report r...  
 [-] linchengyi 0 分/分 9...  
 reply to reply OK!!!!  
 permalink parent 編輯

Comments  
 i dont get it.. bankruptcy?  
 By user: anthropomorphised  
 Comment  
 sweet glenda.  
 By user: moonhead

- Comment or reply without login — Passed

< ↑ > 85 < ↓ >	Plan 9 - A distributed OS with a unified communication protocol and I/O driver model. Wow.
< ↑ > Comments < ↓ >	Plan9 is more or less dead. However, there is a very much alive fork called 9front. By user: Sukiari
< ↑ > Comment < ↓ >	see also Inferno, an Open-free/commercial OS inspired by Plan 9: <a href="http://www.vitanuova.com/inferno/index.html">http://www.vitanuova.com/inferno/index.html</a> By user: shavenwarthog
< ↑ > Comment < ↓ >	I find it fascinating to note how much of my current OS came from Plan 9:
Error, please login first!	

- Load more news — Passed

1		protocol and I/O driver model	submitted 1 day ago by Gaidin	32 comments share save	82	< ↑ > 203 < ↓ >	WebOS lives! HP decides to op
2		HP donates WebOS s	(bbc.co.uk) submitted 1 day ago by tdbos	comment share save hide	21	< ↑ > 4 < ↓ >	HP: WebOS, Enyo app framework
3		WebOS lives! HP dec	(gigaom.com) submitted 2 days ago by fortell	14 comments share save	203	< ↑ > 3 < ↓ >	HP's WebOS plan modeled after
4		HP: WebOS, Enyo ap	submitted 2 days ago by Tangle	1 comment share save hid	4	< ↑ > 20 < ↓ >	Imagine a world with 50% wome
5		HP's WebOS plan mo	submitted 2 days ago by Tangle	comment share save hide	3	< ↑ > 10 < ↓ >	Italian Rock Band Release Alb
6		Imagine a world with	leading open technol	The Ada Initiative ha	18	< ↑ > 40 < ↓ >	Highlights of Blender work fr
7		Italian Rock Band Re	(omgubuntu) submitted 3 days ago by repus	comment share save hide	10	< ↑ > 11 < ↓ >	MSFT's free/OSS (no, really)
8		Highlights of Blender	the world. (Let's see	(youtube.com) submitted 3 days ago by 5	43	< ↑ > 8 < ↓ >	From Open Source to Open Rese
9		MSFT's free/OSS (no	powerful Python IDE	Kinect, ... support (py	13	< ↑ > 6 < ↓ >	-----Load More News-----
10		From Open Source to	(blogs.computerworlduk.com)	submitted 3 days ago by fortell	8	< ↓ >	
11		Paul Allen Is a Propo	a Subscription to Fin	submitted 3 days ago by apoth	6		

## Part 6

# Conclusion

We successfully produced a text-based browser, using python. The browser allows the user to navigate and use all the functionality of a social news site. The browser retrieves and displays headlines, comments, votes, replies and source links And also allows the user to login and post stories and comments — as well as voting them. Keyboard shortcuts are also implemented for easier navigability. We worked together to implement the required functions, and had a working and successful program which went beyond what the spec required.