



[PROPOSAL]

Data Structure Project - Smart Ranking

◎主題：

一、主題
美食

二、想要 rank 的網頁：

<http://jolihi.pixnet.net/blog>

<http://hito-bbq.imagine-7.com/>

<http://www.yayan.com.tw/yayan/index.html>

<http://www.just-eat489.com/>

三、要用來 rank 的關鍵字

燒烤

火鍋

吃到飽

啤酒

◎系統規格：

一、使用者是誰？

任何想鍵入關鍵字進行搜尋的人。

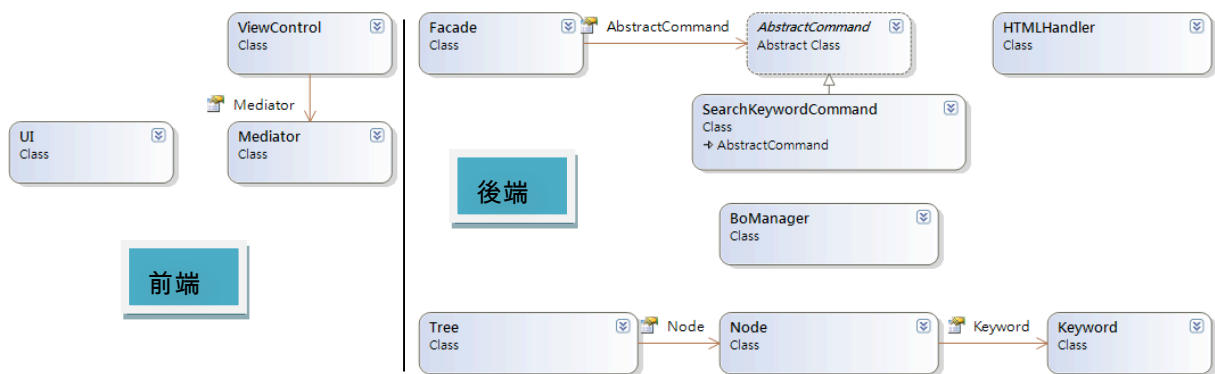
二、使用系統的目的？

希望利用搜尋引擎所查詢出來的結果正為使用者所需。

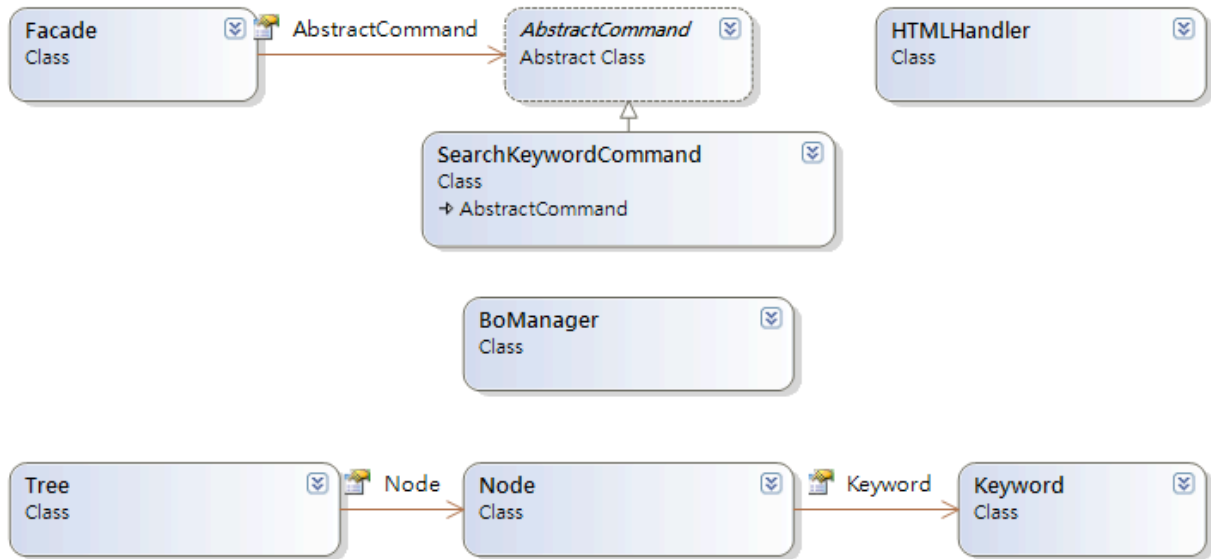
三、系統功能有哪些，如何使用這個系統？

運用此系統以及搜尋引擎，計算出使用者輸入所查詢的關鍵字出現在各個網頁中幾次再乘上使用者所希望的權重比例，以辨別出重要性。只要輸入關鍵字即可使用。

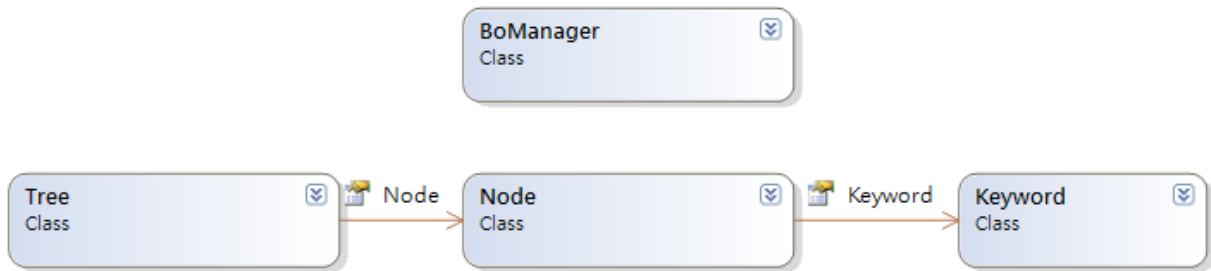
四、Class Diagram (Stage 4)



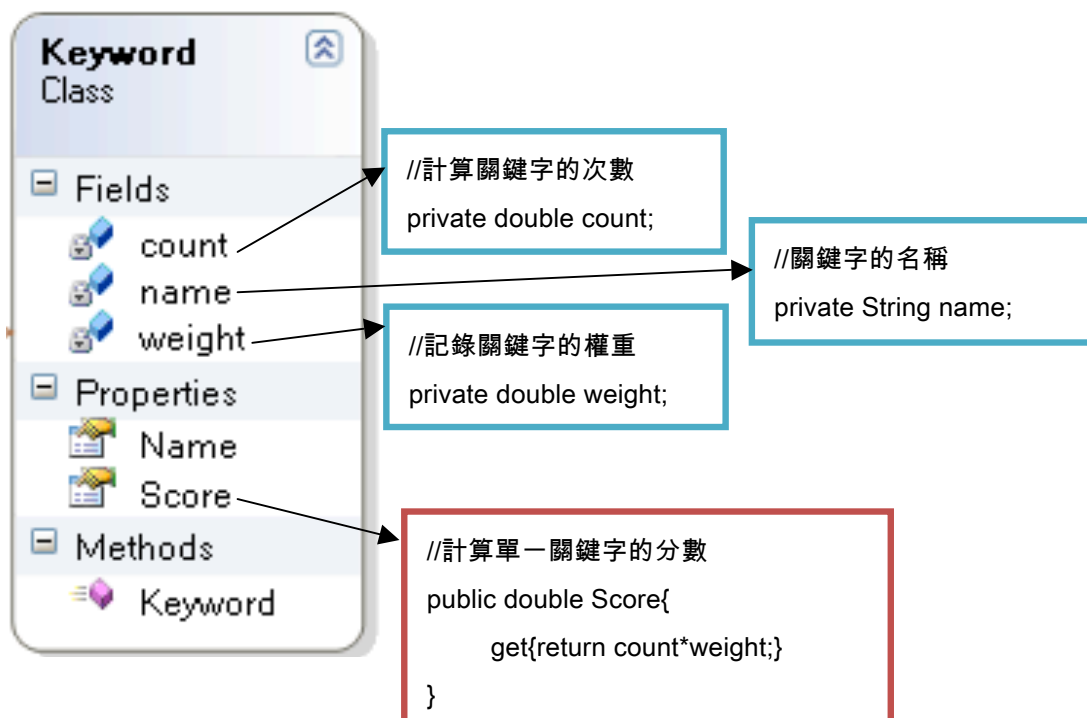
Business Logic Layer



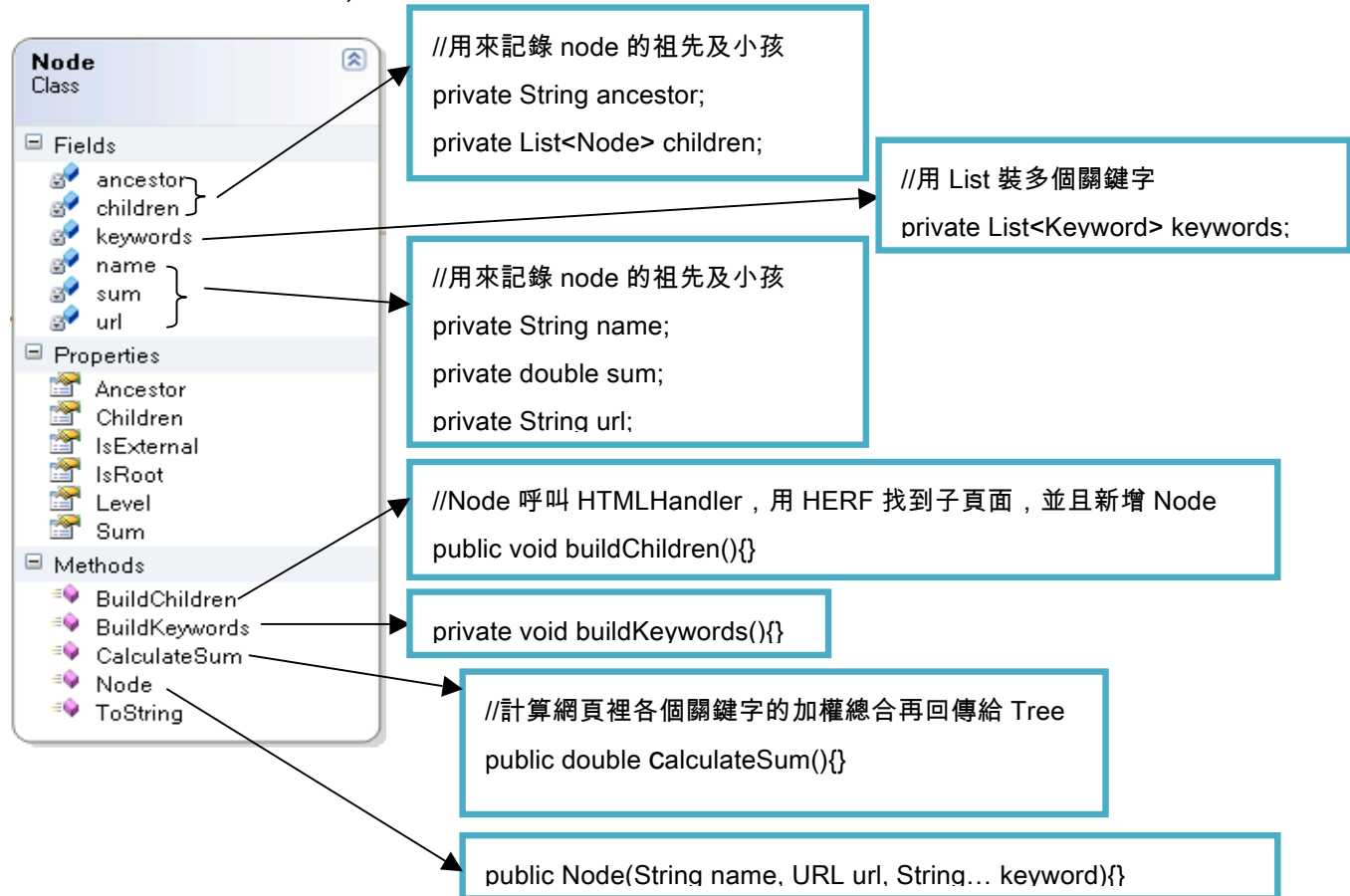
Business Object



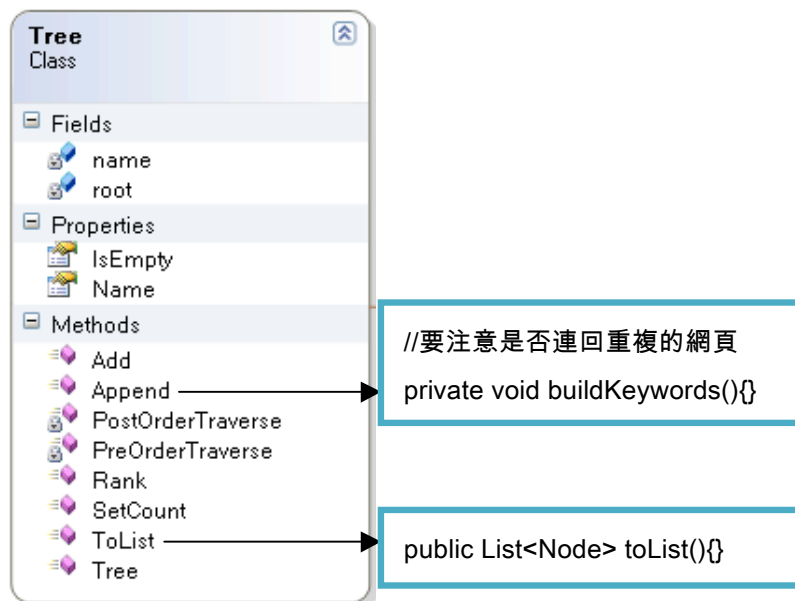
Keyword



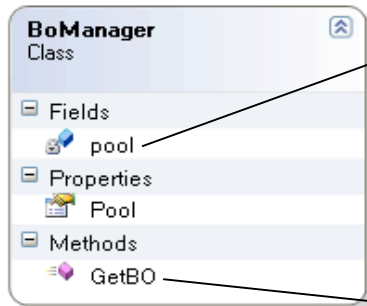
Node //一個 Node , 是一個網頁



Tree //建立整棵樹 , Root



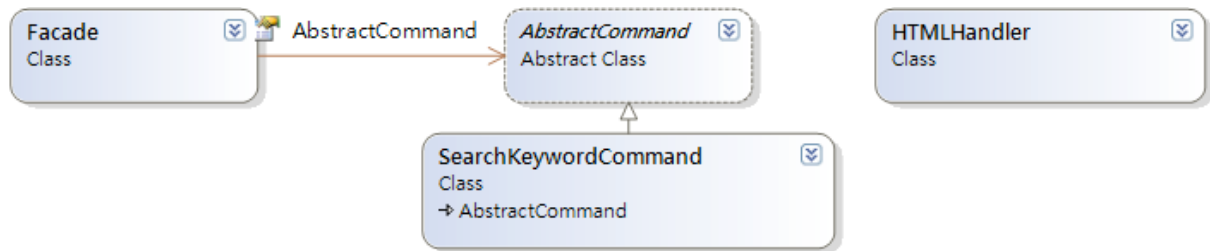
BoManager



```
//用來記錄 node 的祖先及小孩  
private List<Tree> boPool;
```

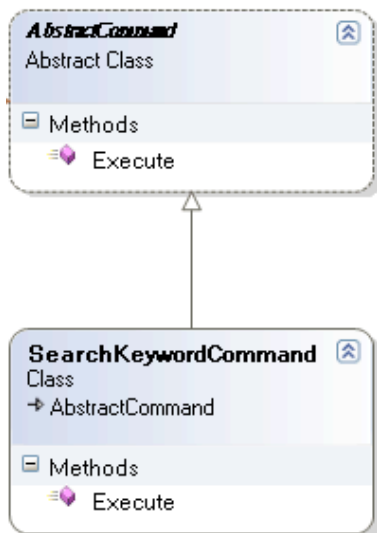
```
//Command 呼叫 BoManager 已取得 Bo  
public List<Tree> getBo(){}
```

Task



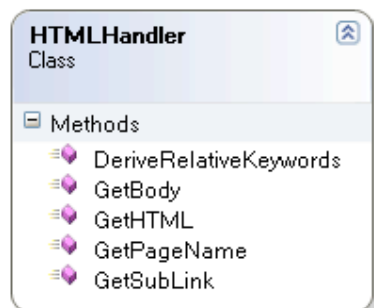
AbstractCommand & SearchKeywordCommand

//使用 Command Pattern :先由 Abstract Command 發布命令給 Concrete Command , 再由 Concrete Command 去執行實際的動作



HTMLHandler

//從 Node 中，拿得 URL，並連上網路該頁面，讀取該頁面之 HTML 檔取得所需的資訊



Facade

//使用 Singleton Pattern : 在程式上表達出物件實體只有一個，使用 Singleton Pattern 後，在外部我們不用去 new 物件實體，只須取得實體。

從 ViewControl 取得關鍵字，傳給 Command

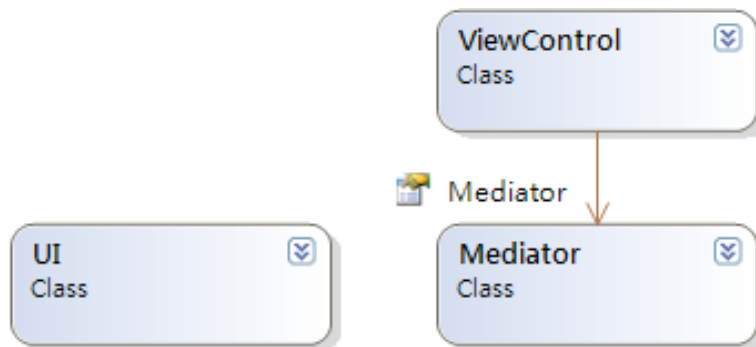
Facade
Class

Methods

- Execute
- GetInstance

```
public void distribute(String signal, String... keyword)
{
    //將 signal 轉成 char , 丟入 switch 中判斷
    switch(new signal)
        ...
}
```


Presentation Layer



ViewControl

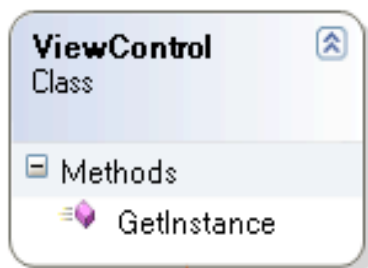
//使用 Singleton Pattern。

抓到 UI 使用者輸入的關鍵字，傳給 Façade

呼叫 Façade 的 excute()，丟入關鍵字及 Signal

Signal 是用來判斷要執行哪個 Command 的依據

P.S.加入""視為同一個關鍵字，需在這裡作 Signal 的判別



Mediator //使用 Mediator Pattern 以協調多個物件

UI //用 Java Swing 做 Graphic User Interface 請使用者輸入多個關鍵字
(一個 TextBox 和一個 Button)

◎系統開發階段：

一、用關鍵字計算網頁的 rank

這部分為寫死的功能，將以我們所提供的數個網頁，配以我們設定好的權重，來計算網頁的排名。是先用 Node 裡頭的 calculateSum()計算各自多個關鍵字的分數加總，然後在 Tree 裡頭的 Rank 照順序排好名次，再傳回前端 print 出來。

二、網頁內的子連結可以讀取出來，一併列入 rank 的計算

此部分與 HW6 內容相似。我們將使用 Node 裡頭的 calculateSum()計算各自多個關鍵字的分數加總，且在此判斷如為指定之網頁，便將其 children 的分數加總至自己身上，最後在 Tree 裡頭的 Rank 照順序排好名次，再傳回前端 print 出來。

三、使用關鍵字以 Google 或 yahoo 搜尋，然後對搜尋回來的連結做 rank

是以接收使用者所輸入之關鍵字傳至後端，再由後端連上 Google 讀取搜尋頁面的連結，網址是以 Google 搜尋網址的 query 加上關鍵字(中文須轉編碼)。連結到的網頁，將讀取其原始碼，並藉由 HW5 所學，存取搜尋網頁之原始檔傳至後端，由 HtmlHandler 的 GetHTML()...等 function 讀所需資訊後，再由我們的 rank 方法 run 過一次後，最後傳到前端呈現出來結果。

四、從 Google 搜尋回來的頁面，再找出可能會有用的關鍵字

燒烤的相關搜尋

[燒烤 台北](#) [出一張嘴燒烤](#)
[燒烤 新竹](#) [石頭燒烤](#)
[原燒](#) [台中燒烤](#)
[燒烤場](#) [巴西燒烤](#)
[小樽燒烤樂園](#) [白石燒烤場](#)

在原始檔(僅節錄燒烤 台北的 tag 部分)上，原始碼如下：

```
<em>燒烤</em>的相關搜尋</div><div class=brs_col><p><a href="/search?hl=zh-TW&lr=&client=firefox-a&hs=UvZ&rls=org.mozilla:zh-TW:official&channel=s&p;q=%E7%87%92%E7%83%A4+%E5%8F%B0%E5%8C%97&revid=356618742&sa=X&ei=uNbaTKCWFouevgPR-ISgCg&ved=0CKsBENUCKAA">燒烤 台北</a></p></div>
```

我們希望能藉由 HW5 所學讀取 tag 之方式，存取搜尋網頁之原始檔傳至後端，由 HtmlHandler 的 DeriveRelativeKeywords()讀取相關搜尋，再呈現出來。

最後，我們有討論過，如有餘力想再增加之搜尋功能為可以搜尋特定檔案類別或語系的網頁來做排序，並且希望 Tree 下的第一層是以檔案類別來區分(原始碼同樣來自於 Google)。