

Hash Table & Hash function

05170229 巨資四B 袁嘉謙

Hash function 原理

參考資料: <https://www.inside.com.tw/feature/audi2019/18188-audi-aia-2019-meettaipei>
<https://reurl.cc/72W5L5>

1. Hash function是一個將值利用**數學運算轉換為雜湊值的函式**。
2. 不同長短的訊息所算出的**雜湊值長度會固定**。
3. 輸入**相同的訊息**會產生**完全相同的雜湊值**，但**無法反推**原訊息。
4. 當原訊息更動，雜湊值必有所變動。
5. 當運算方式單調且應用於過多資料，**可能**使不同訊息產生相同雜湊值，**造成Collision**。

Ex. Division Method>>取餘數的方式，許多值的餘數相等



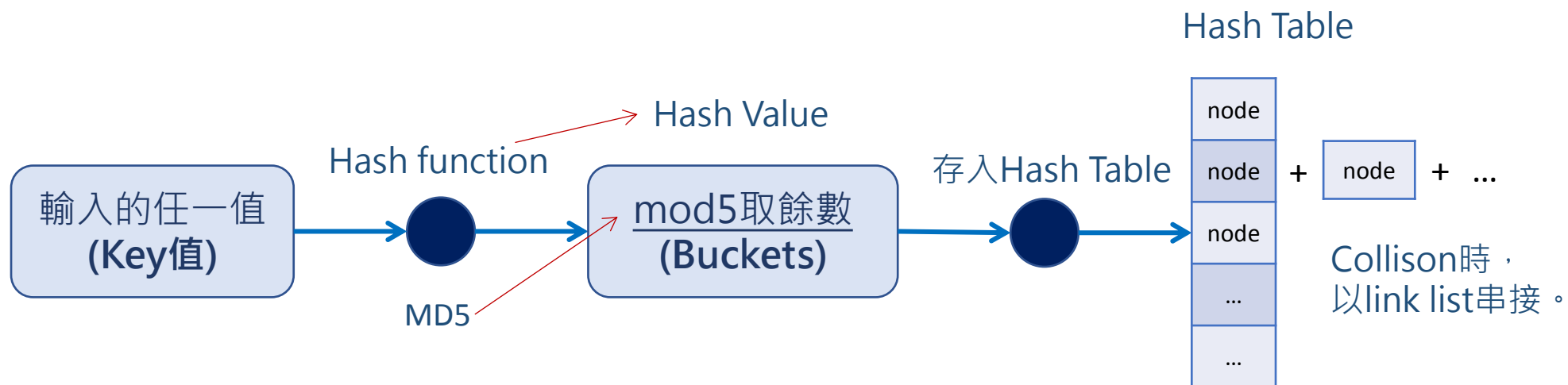
Hash Table 原理

參考資料: <https://blog.techbridge.cc/2017/01/21/simple-hash-table-intro/>

1. Hash Table是**存放Hash function 算出的雜湊值**的數組，也就是儲存 (key, value) 的**一種資料結構**。
2. 步驟：把資料放到雜湊表：給定key值與存放的value > 將key 的字元轉換成數值相加(相加的值即是hash鍵值) > 在table陣列對應到存放的value。
3. 核心概念是要**避免記憶體空間浪費**。
4. 用到的Key數量為n，Table大小為m，**目標就是 $m = \Theta(n)$** 。

Hash Table 流程圖

參考資料: <https://www.inside.com.tw/feature/audi2019/18188-audi-aia-2019-meettaipei>
<https://reurl.cc/72W5L5>



學習歷程

參考資料: <https://www.cs.wcupa.edu/rkline/ds/hash-sets.html>

1. 起初先了解Hash function與Hash Table的關聯，並開始寫程式碼。
2. 卡關一：Hash function與MD5之間的關聯有些卡，後來找到這篇文章<https://reurl.cc/72W5L5>解決疑問，其中參考了PPT的第7頁。

```
def add(self, key):  
    md5 = MD5.new()  
    md5.update(key.encode("utf-8"))  
    md5 = md5.hexdigest()  
    bucket = int(md5)  
    if self.data[bucket] == None:  
        self.data[bucket] = ListNode(md5)  
    else:  
        new = ListNode(md5)  
        now = self.data[bucket]  
        while now.next != None:  
            now = now.next  
        now.next = new
```

3. 測試時發現會出現測資為定義的問題(已解決>>發現是自己在輸入測資時未加' ')
4. 卡關二：contains的撰寫時，初期邏輯有點不順