

## 壹、前言

### 一、研究動機

近年來，人工智慧掀起一股新科技風潮，生活中漸漸地有許多儀器被AI取代，最近更是有個能和我們聊天的文本生成技術 "ChatGPT"問世。ChatGPT的普及證明了大數據的強大，而AI的使用頻率只會愈來愈高，對世界的影響也會日漸增加。

在選修課程進行的同時，段考範圍中排列組合的複雜使同學們非常煩惱，在AI趨勢與學校課程的雙重影響下，我們便以此為發想，試著運用在選修課程中學習到的程式編寫來做出能夠解開排列組合的簡易AI。

### 二、研究目的

我們想製作一個排列組合AI，當我們輸入題目後即能產生答案的系統。

#### (一) 製作出解題模組

用python製作一套能解相同類型題目的模組。  
(只有條件不同)

#### (二) 文字判別系統

利用python的斷詞系統分開題目字詞，並藉由取出的關鍵字導到相應的解題模組。

## 貳、研究流程

我們先是用習作的題目測試是否可以用python程式解出來，確定是可行之後，我們在網路上抓十題排列組合的題目，且每題都是來自不同常見的類型。

我們文字辨識的步驟是先用jieba斷詞模組分開每個單字，然後再用擷取is.islower(取小寫的英文字母)拿出我們需要的資訊，將資訊放入解題模組後即能得到該題目的答案。

尋找題型並編寫程式

結合題目識別與運算

用文字辨識判別題目邏輯

研究結果探討

我們主要使用的文字判別系統是"Jieba"斷詞模組，此模組分為三種模式，分別是精確模式、全模式和搜索引擎模式，而我們這次是使用到他的全模式系統。透過 jieba.cut() 來進行斷詞，cut\_all 參數為 True 的話就會為全模式。

我們的研究結果其實和原本的目標不太一樣，有點不太完整，因為我們原本是想做能夠判別多種題型的模型，但最後只做出兩種，因為有些題目的關鍵字太像導致難以判別，其實我們可以試著用其他的斷詞模型判別，也許會更好分辨出我們想要的題目。

精確模式	將句子最精確的切開，適合文本分析
全模式	把句子中所有的可以成詞的詞語都斷出來，速度非常快。
搜索引擎模式	在精確模式的基礎上，對長的詞語再次切分，提高召回率，適合用於搜索引擎分詞。

## 參、研究結果

(1)平面上有10個點，其中任三點均不共線，則由此10點可決定幾條直線，幾個三角形。45、120 相異物組合	(5)渡輪3艘，每艘可載6人，則6人過渡，可安全過渡的方法數有幾種。729 乘法原理
(2)以0、1、2、3、4、5共六個數字組成的四位數奇數(數字不可重複)共有多少個?144 有限制的相異物排列	(6)把3個轉學生分配到現有的12個班上，限制每班最多分配1個，共有幾種分配方式。1320 相異物排列
(3)有9個小朋友要入座排成一列的4張椅子，只有4人可以入座，則入座的方式有幾種。3024 相異物排列	(7)三男兩女排成一列若男生之間不排女生，共有多少種排法?36 相鄰必綁
(4)將"庭院深深深幾許"七個字全取排成一列，則三個"深"字不全連在一起的排法有幾種。720 有限制的不盡相異物排列	(8)將"mhchcm"這些英文字母任意排列，問共有幾種不同的排列方法?90 不盡相異物排列

圖(一)排列組合八大類型題目

```
4.自mathematical一字中，取四個字母。(1)組合數 (2)排列數

import jieba
document = ['自mathematical一字中，取四個字母。(1)組合數 (2)排列數']

#精準模式
for sentence in document:
    seg_list = jieba.lcut(sentence)
    print(seg_list)
for i in seg_list:
    if i.islower():
        print(i)
        q=i
```

圖(二):文字斷句程式

```
#不盡相異物排列

import math
def group_duplicates(word):
    char_count = {}
    result_set = set()

    # 計算每個字母的出現次數
    for char in word:
        if char.isalpha():
            char_count[char] = char_count.get(char, 0)

    # 將出現次數超過兩次以上的字母加入結果集合
    for char, count in char_count.items():
        if count >= 2:
            result_set.add(char)

    x = q
    result = group_duplicates(x)

    c=set(result)
    y=len(c)

    a=set(x)
    b=len(a)

    #四異
    k=4
    s=math.comb(b,k)
    print(s)

    #二同二異
    j=1
    f=math.comb(y, j)
    p=2
    o=math.comb(b-2, p)
    print(f*o)

    #二同二同
    d=math.comb(y, p)
    print(d)

    print(s+f*o+d)

    #排列數
    from math import comb, factorial
    e = s*24
    g = f*o*6
    h = d*12
    print(e+g+h)
```

圖(三):英文字母重複次數計算

圖(三)的英文處理模組

組合數數量

排列數數量

圖(四):計算部分

我們首先製作的是文字斷句程式，它是透過詞性、關鍵字的判斷來決定題目類型。就像我們在研究流程講到的，先用"Jieba"分開題目的關鍵字，再用"is.islower"取出我們需要的字詞，而"q"就是我們需要的元素，如圖(二)。

這是一個不盡相異物排列的題目，這張圖主要是在處理英文單字，計算英文字符串中重複兩次以上的字母有幾個、有幾種英文字母，然後把整個過程設成一個模組"group\_duplicates"，如圖(三)。

而最後計算的過程，先把四種可能性分開處理，再用前面圖(二)、圖(三)得到的數據，算出組合數和排列數，如圖(四)。

## 肆、結論

AI的形成總是需要一段長久的時間，絕非一朝一夕便可造就，如同一開始所述，AI真正強大的地方在於大數據的力量，如果沒有大量的數據樣本，AI的回答方式會過於單一，甚至出現無法解讀的情況，而同樣的，我們也不希望遇到此狀況。

我們以製作簡易AI為目標前進時，其實就有想過這個問題，如果成功做出來也會因為數據不夠而無法使用，在做到第五到第六種類題時，這個問題便愈來愈明顯，而原因在於我們沒有足夠的時間去收集大量題目(樣本)、思考並編程，以及在編程方面現有知識背景的不足，以至於數據會遠遠不夠，文字判別系統也間接變得不靠譜，在意識到以我們現在的狀況已無法達到預期時，我們只能降低目標。

目前的研究只達到十種題目的編程，離一開始目標的目標仍有一段距離，因此我們希望後續能持續擴充題目範圍；而若要達到AI的水準，便須以更大量的數據進行訓練。期望未來能搭配介面設計，實作成可應用之系統。

## 資料來源

Python - 知名 Jieba 中文斷詞工具教學, 用pip 安裝

<https://blog.kennycoder.io/2020/02/12/Python-%E7%9F%A5%E5%90%8DJieba%E4%B8%AD%E6%96%87%E6%96%B7%E8%A9%9E%E5%B7%A5%E5%85%B7%E6%95%99%E5%AD%B8/>

排列組合題目(1)

<https://math.ymhs.tyc.edu.tw/chenjt99/ScrClass/1042/HomWok/Book2/%E7%BF%92%E4%BD%9C%E7%94%B2-2%20%E6%8E%92%E5%88%97.pdf>

排列組合題目(2)

<http://www.ycvs.nipc.edu.tw/czfiles/0/1000/img/121/CH10.pdf>

"Jieba"斷詞模組

<https://blog.kennycoder.io/2020/02/12/Python-%E7%9F%A5%E5%90%8DJieba%E4%B8%AD%E6%96%87%E6%96%B7%E8%A9%9E%E5%B7%A5%E5%85%B7%E6%95%99%E5%AD%B8/>