

# 大數據改良電腦化適性測驗系統



# 研究動機

台灣城鄉教育資源不均的問題一直存在，偏遠地區的學童往往缺乏良好的學習機會。現今的電腦化適性測驗的試題參數可能過時，影響測驗效度。為了改善這一現況，本研究旨在探討如何利改良電腦化適性測驗（Computerized Adaptive Testing, CAT）來提升偏遠地區學童的英語學習成效。

## 研究目的

- 1. 填補現有電腦化適性測驗在考生母群變化方面的缺口。
- 2. 提供個人化學習體驗與即時回饋。
- 3. 讓電腦化適性測驗能自動調整題目難度。

## 研究設備與器材

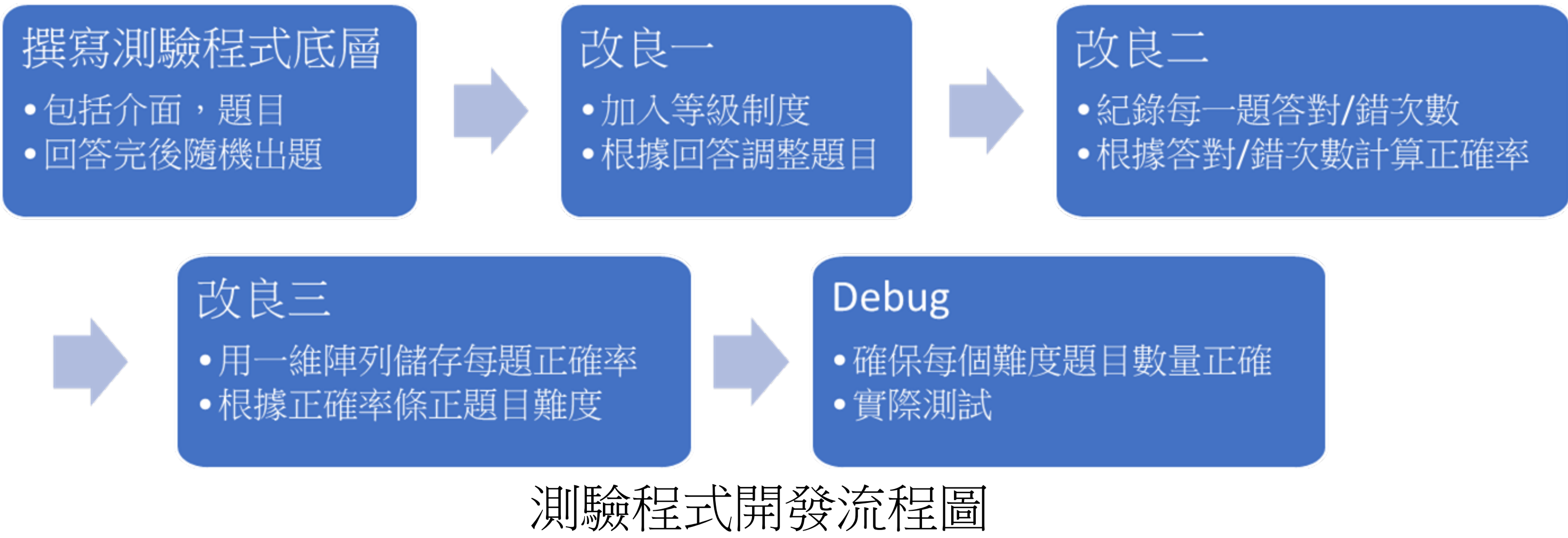
硬體： 筆記型電腦	軟體： 程式語言Dart(3.8.0) Flutter框架 Visual Studio Code	程式語言Python 版本控制系統Git
--------------	---	-------------------------

## 研究過程與方法

一、需求分析，我們會通過大量的討論及查找資料，確定大致方向

二、設計與開發，我們運用Flutter製作APP，包括其中的算法及測驗的架構，同時會開始運用Python製作答題機器人

三、測試及改良，則是會進行大量測試，透過測試來發現Bug並開始Debug。



### • 題庫準備

從歷屆會考以及學測英文測驗中挑選44題關於文法的選擇題。將題目轉換成右圖格式。

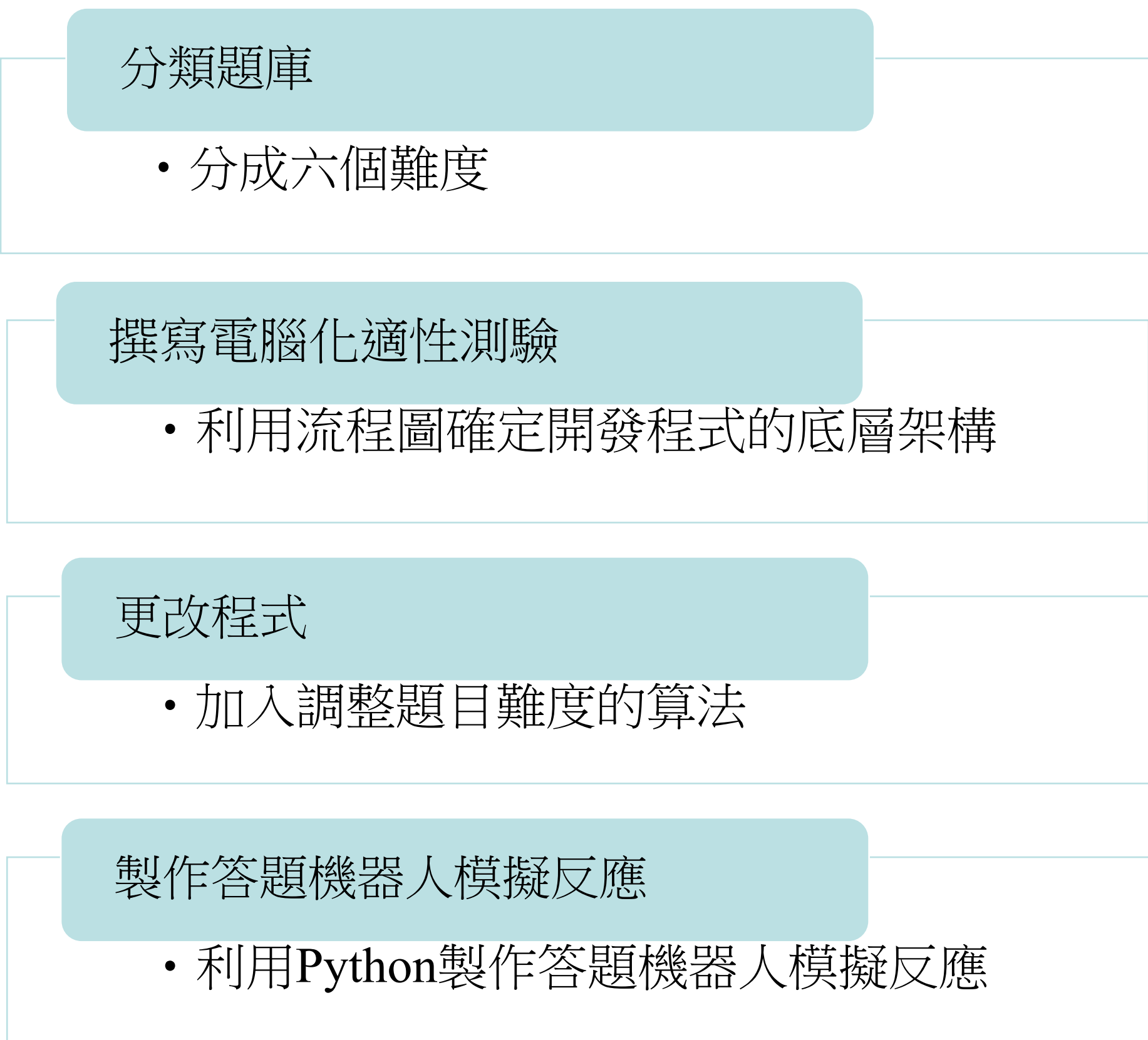
### • 定義初始難度

將題目整理後做成前測交給受試者，將受試者回答的內容放入Excel進行分析，透過各題P值，將題目排名，並且將題目分為六個難度，每個難度7題左右。

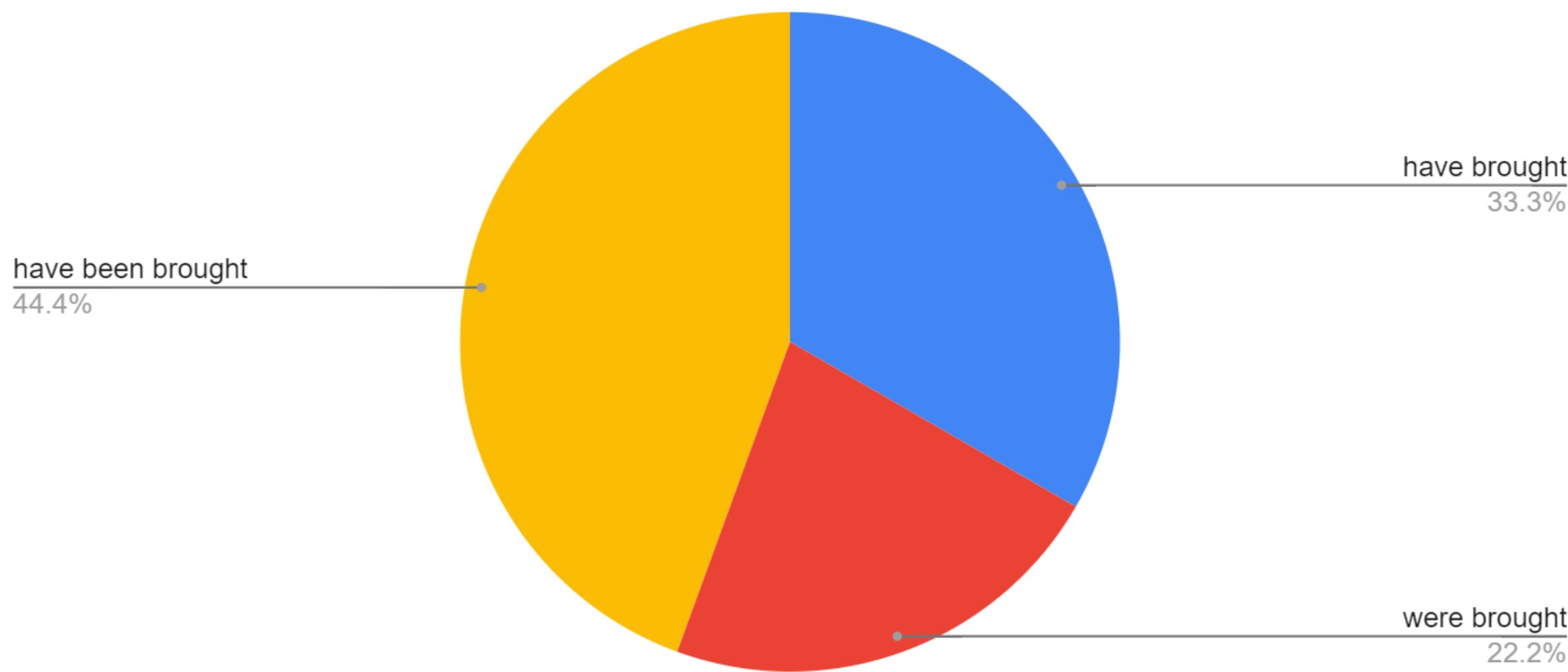
```
Question(  
    questionText:  
        " If the number ( ) is in the same area as the number making the call, an area code  
usually doesn't need to be dialed.",  
    question_number: "4-3",  
    question_appear: "改寫自 102 學測(16)",  
    candidates: [  
        Candidate(label: "calling", type: "(A)現在進行式"),  
        Candidate(label: "being called", type: "(B)現在簡單式", isCorrect: true),  
        Candidate(label: "have called", type: "(C)現在完成式"),  
        Candidate(label: "has been calling", type: "(D)現在完成進行式"),  
    ],  
    correctType: "(B)現在簡單式",  
    difficulty: 4.),
```

題目轉換後格式

### 研究方法



7.Apples have been grown for thousands of years in Asia and Europe, and ( ) to North America by European colonists in the 17th century.

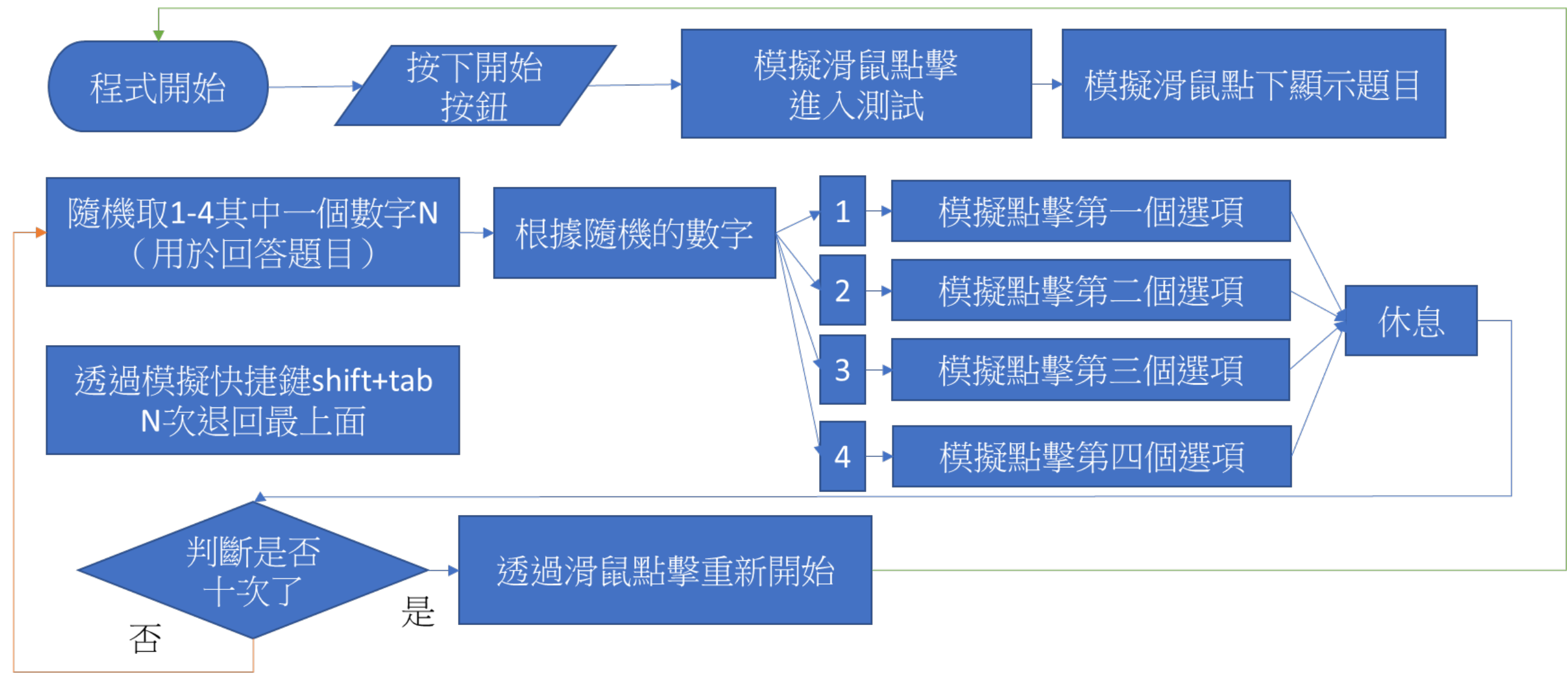


前測其中一題選項選擇率



• Python答題機器人--分為兩次迭代

第一版本運行流程圖



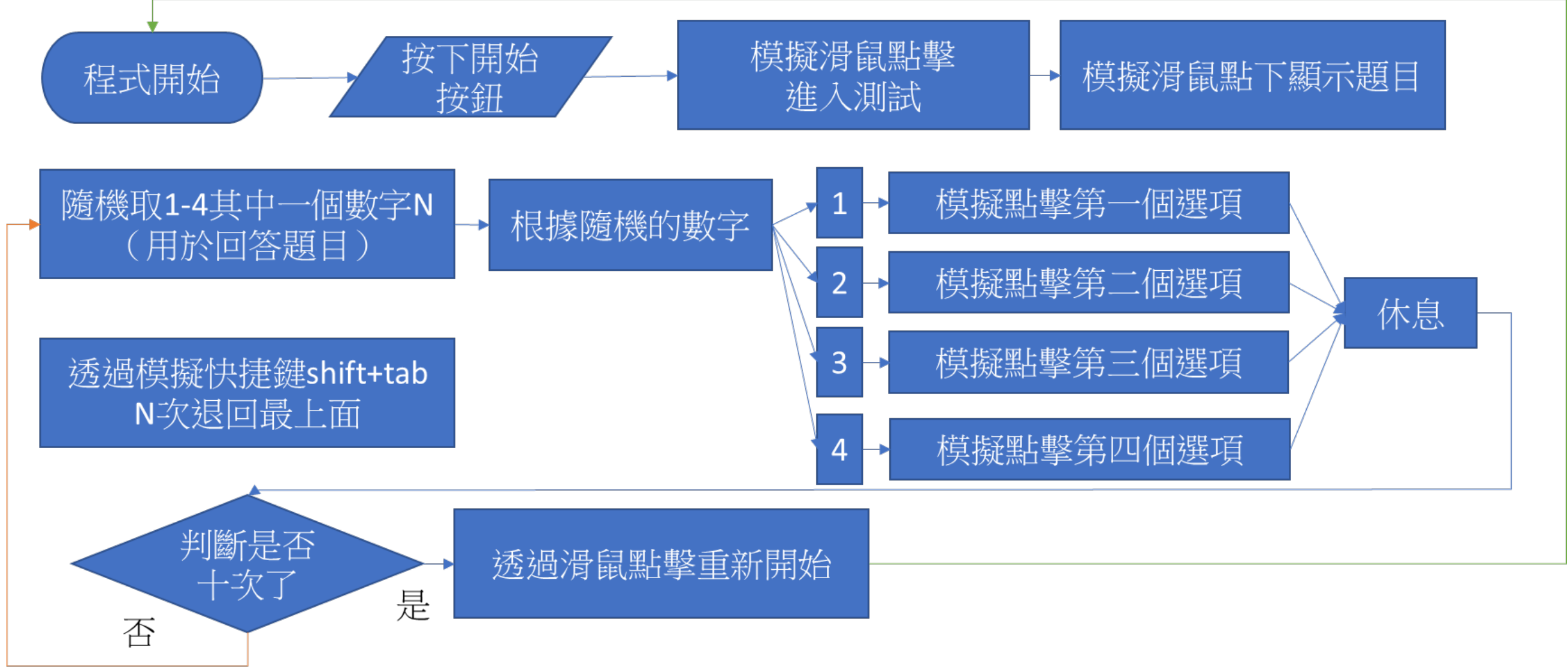
第一版本答題過程中是確認出每一個按鈕在我螢幕上的X,Y軸座標並點擊。

問題點

- 1.沒辦法將視窗縮小，監看終端機輸出
- 2.有時候定位不准會點擊歪掉
- 3.不同螢幕尺寸導致點擊定位問題。

基於以上三個問題我們決定更新版本。

第二版本運行流程圖

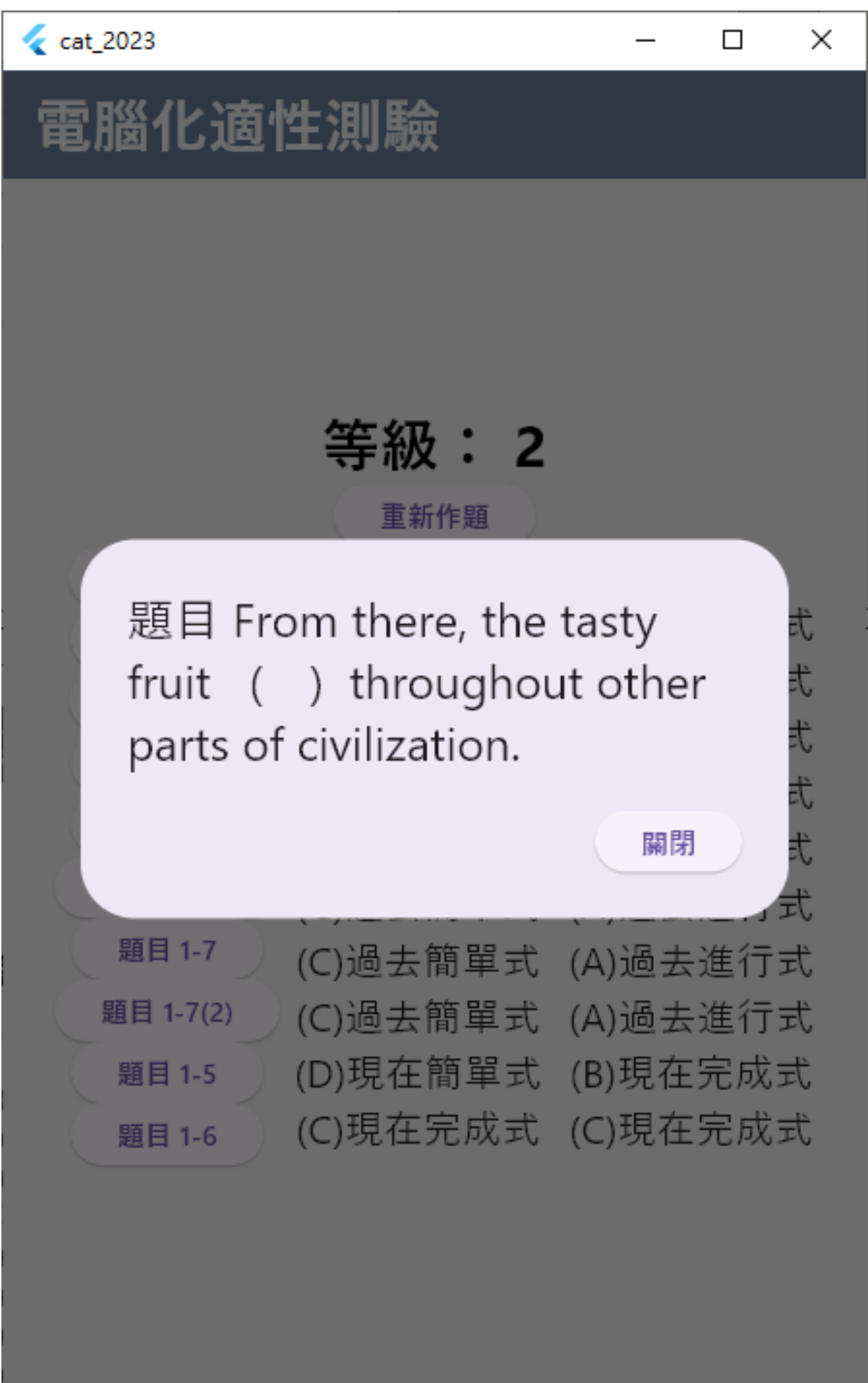
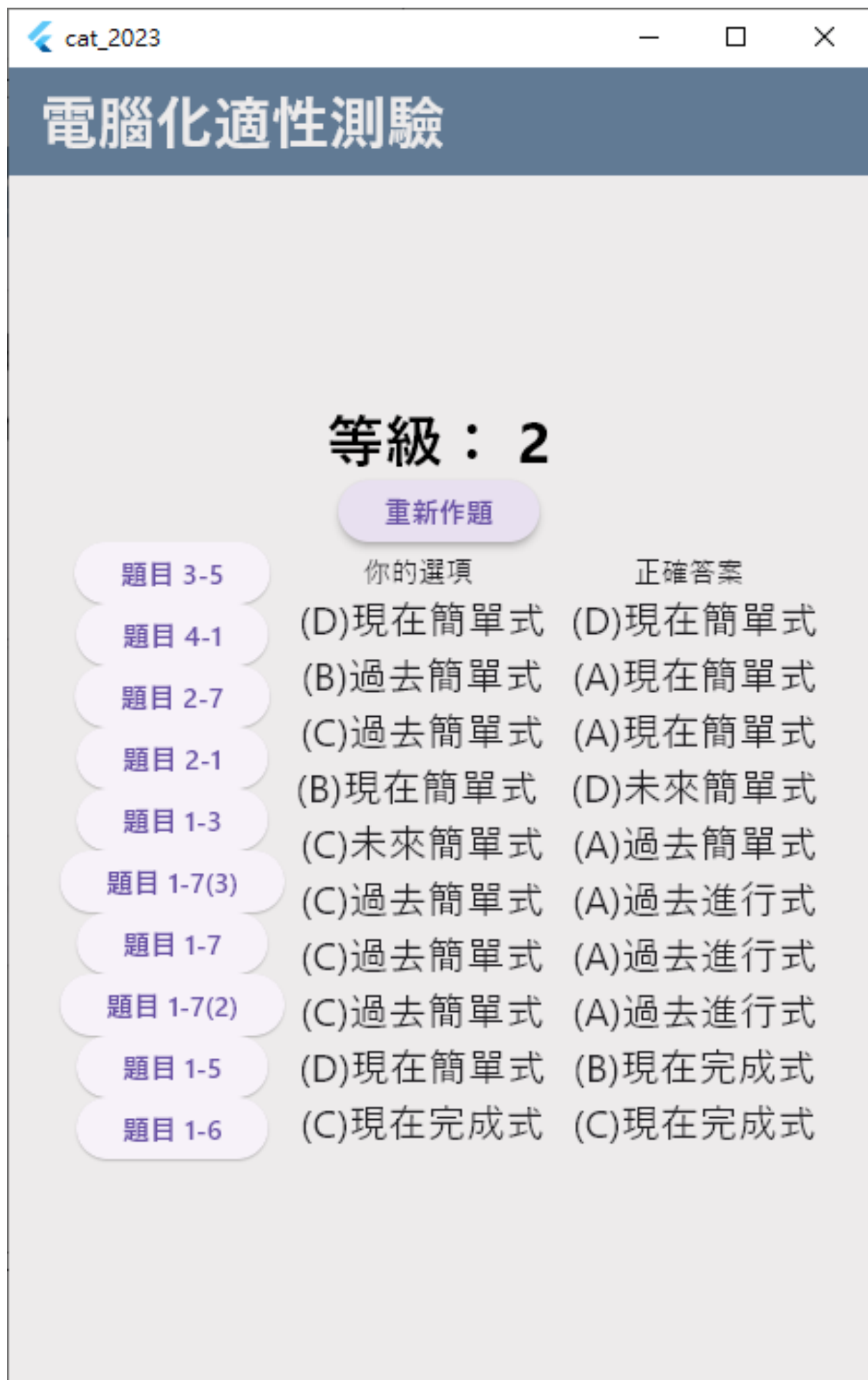
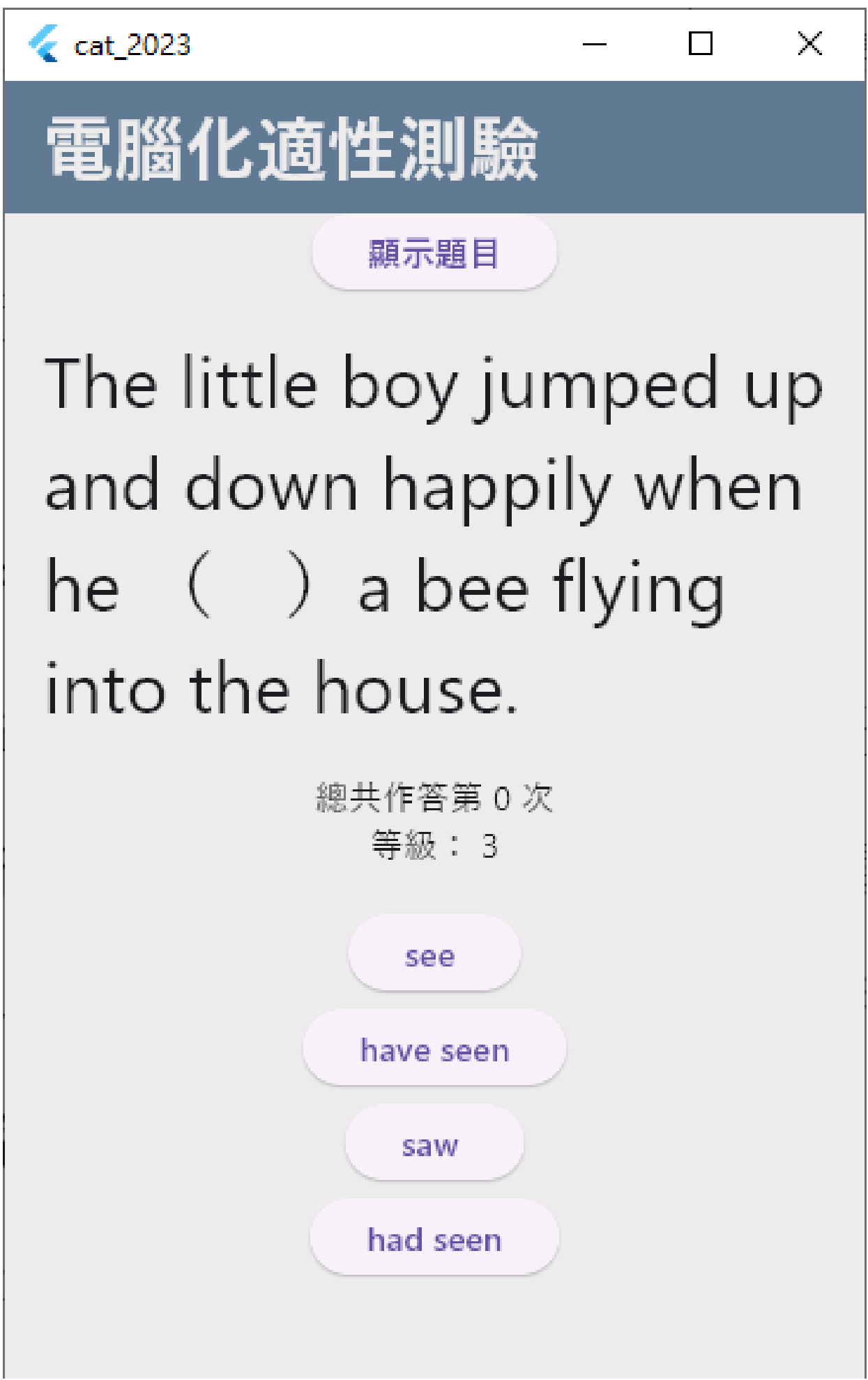


第二個版本是在測試程式中意外發現的，測試中發現透過鍵盤上的Tab鍵可以切換選項，按下Enter鍵可以按下選項，於是基於此方式我們開發出了第二代的程式，第二代的不同在於每選一次選項，都會模擬快捷鍵Shift+Tab來回到最上面的選項。

研究結果

- 我們開發了一套基於Flutter框架的大數據適性測驗系統。系統根據答題紀錄，運用算法改變每題個難度值，使題目排序發生改變。
- 經多輪實驗測試後，我們發現以使用者回答過程為基礎校正P值可以更好評估他們的真實能力。
- 與原有使用全體測驗者無法更新P值的傳統系統相比，本系統運用大數據分析，能使題庫隨時調整難易程度，使測驗結果更加適性化。
- 從測試結果看，本系統在不同能力者上測試效果均得到提升。特別是對於能力水平落後的受測者，由於題目設定調整更合適，能更準確反映其水平。

使用者介面

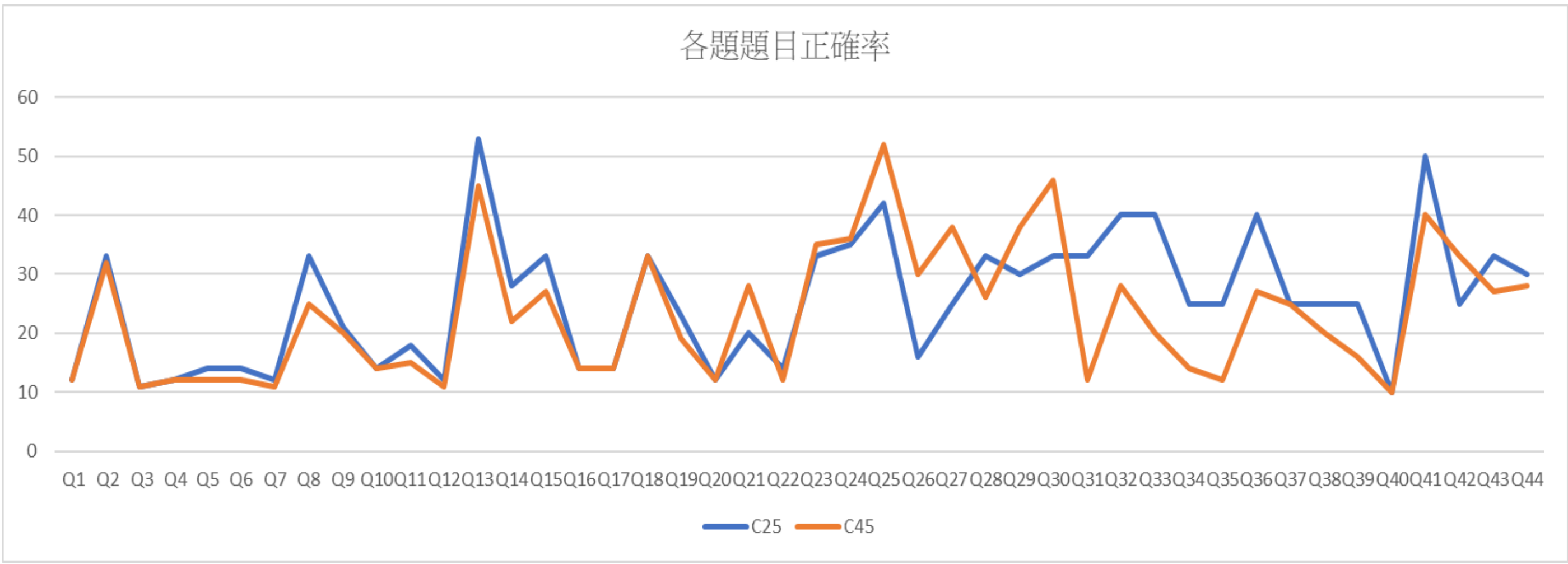


這是實際答題的狀況，答題時會顯示題目及選項。使用者根據選項自行判斷，並選取最正確的選項。同時使用者可以即時看到目前測驗進行到第幾題，以及目前自己的等級，答對等級會加一，答錯則扣一。

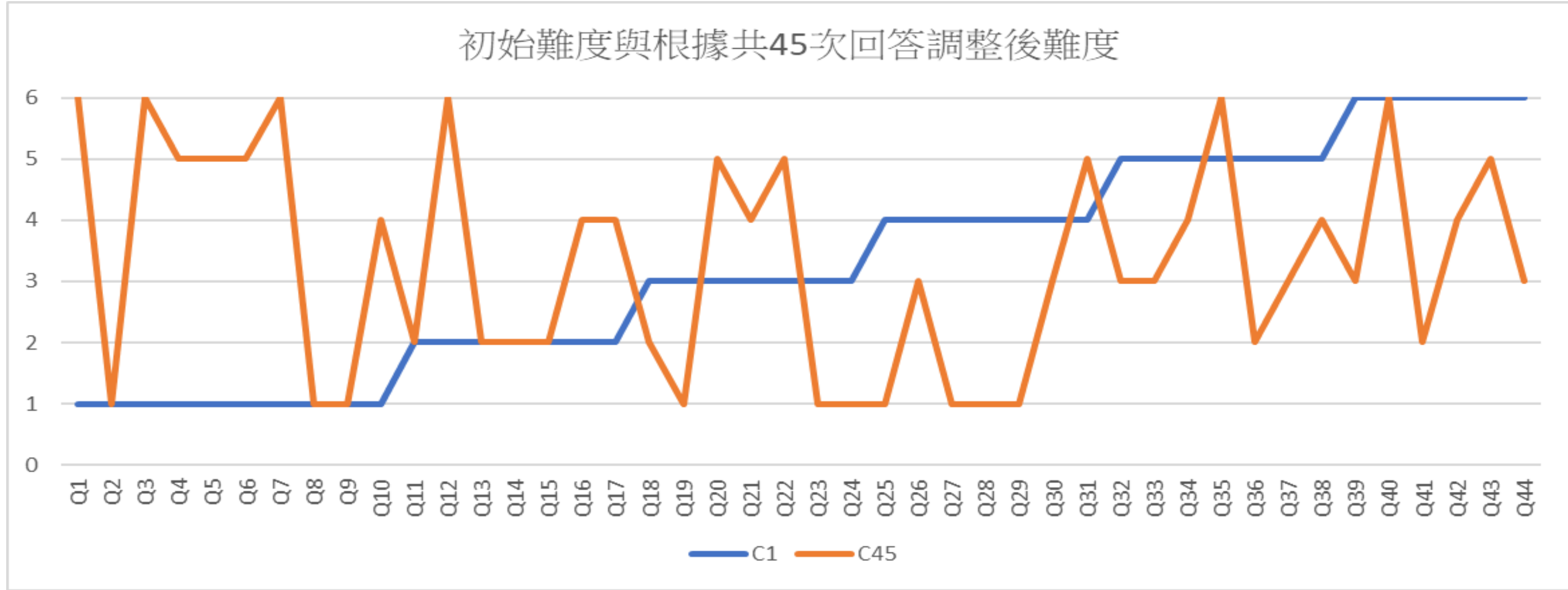
這是作答完的回饋，會顯示目前的等級，以及前10題答道的題目正確的選項及你選的選項，按下題目則會顯示題目內容，按下重新作題則會回到一開始的畫面。



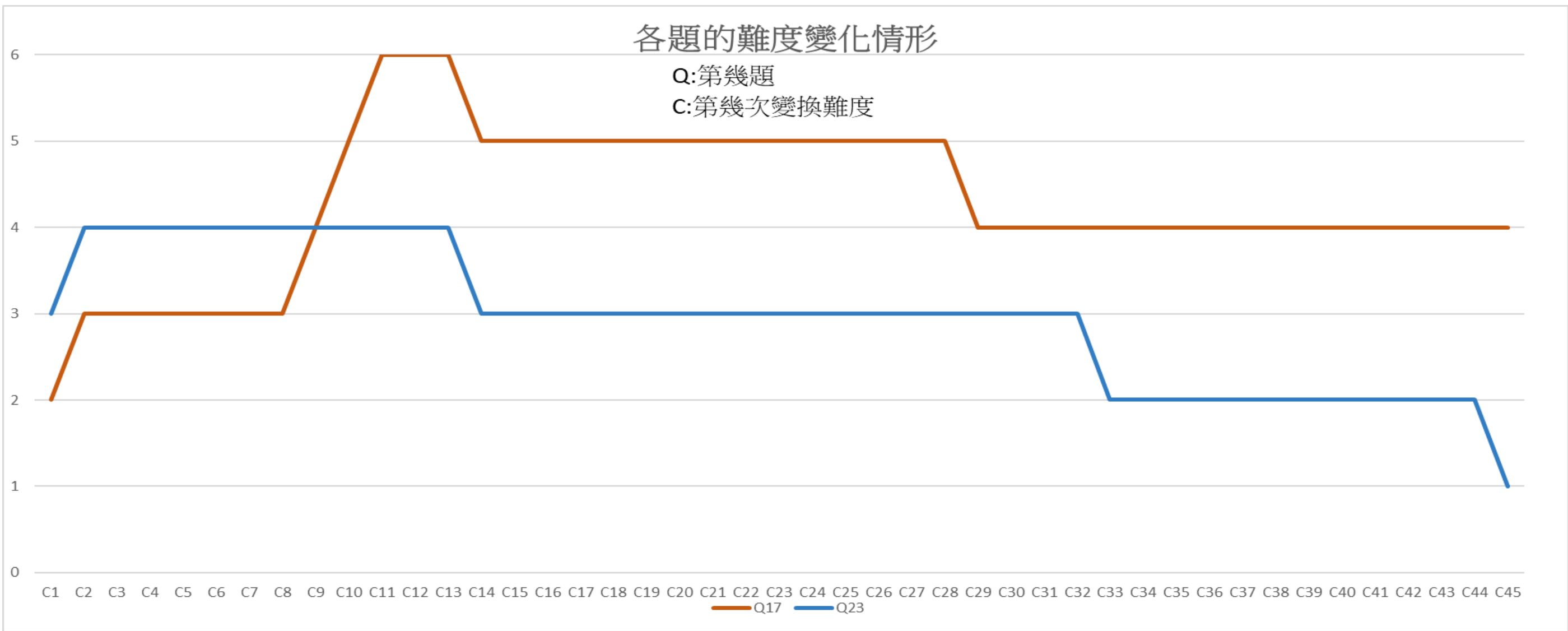
## 數據結果



此圖是在第25次改變難度時每一題的P值（藍色線）以及第45次改變難度時的P值（橘色線）



這是初始難度C1和共45次的回答後調整的難度C45的折線圖，C1初始難度為階梯狀，C45調整後的難度比C1更加起伏。



此圖為答題50次後隨機選取兩題查看其題目難度變化情形，Q為第幾題，C為第幾次變換難度，可以看出題目會隨著該題的P值，進而改變該題難度。

## 討論

通過分析測試結果和運行過程中的終端輸出，我們發現了一些改進和優化的空間。

一、將程式執行時的終端機輸出自動記錄成Excel表格

二、此應用程式須改進之處

（一）程式加入存檔功能

（二）讓不同設備同步題庫難度

（三）優化程式碼

（四）提供個人化回饋

（五）每次答題後將答對狀況自動繪製成折線圖

## 結論

一、大數據電腦化適性測驗系統開發

我們成功開發了一套基於Flutter框架的大數據電腦化適性測驗系統。該系統利用演算法根據使用者答題記錄動態調整題目難度，從而改變題目排序。

二、答題機器人作用

通過答題機器人的回答，驗證了系統背後的難度調整系統的可行性。

三、真實母群題反應及答題機器人模擬反應

系統能夠根據題目的P值進行相應的難度調整。

綜上所述，我們的大數據適性測驗系統在題目難度調整和測試效果上取得了顯著的成果，為適性化教育提供了良好的改進方案。

## 參考資料

鄭育文、陳柏熹、宋曜廷、陳信豪、蕭孟筵（2014）。電腦化適性職涯性向測驗編製研究。**教育心理學報**，**46(2)**，271-288。

陳新豐（2007）。台灣學位電腦化測驗研究的回顧與展望。**教育研究與發展期刊**，**3(4)**，217-248。

余民寧（2009）。**試題反應理論IRT及其應用**（初版）。臺北市：心理。