

科技部補助
大專學生研究計畫研究成果報告

計 畫 ： App Inventor 2 程式之自動批改系統 名 稱
--

執行計畫學生：林珈均

學生計畫編號：MOST 109-2813-C-110-012-E

研究期間：109年07月01日至110年02月28日止，計8個月

指導教授：楊昌彪

處理方式：本計畫可公開查詢

執行單位：國立中山大學資訊工程學系（所）

中華民國 110年03月22日

摘要

程式設計的自動批改系統相較於傳統的人工批改，擁有更高的效率、更低的錯誤率及人力成本。針對 C/C++、JAVA 的程式設計，已有不少檢測系統，進行電腦自動批改，其運作方式幾乎都是將考生輸出的答案，與標準答案進行比較，採用的方法，就是文字檔案的比對。

針對視窗程式，卻很少有自動批改系統。App Inventor 2 是 MIT 所發展出來的一套視窗程式設計系統，其程式可以放入手機執行，實用性很高。本計畫欲發展一套針對 App Inventor 2 的自動批改系統，我們將利用 App Inventor 2 內建的檔案輸入輸出功能去做功能的檢查。

目錄

壹、研究動機與研究問題	2
貳、文獻回顧與探討	4
參、研究方法及步驟	11
肆、成果展示	18
伍、結果與討論	22
陸、參考文獻	23

壹、研究動機與研究問題

教育部 108 課綱已於 108 年 8 月正式上路，核心素養共有九項，其中「系統思考與解決問題」、「科技資訊與媒體素養」兩項與資訊科技息息相關。其落實課程則為運算思維、程式設計。換言之，程式設計不再僅是資工系或工學院學生的專利，而是國中生、高中生均需要學習的全民運動。然而不同領域與不同對象，對於程式能力的需求不盡相同。資訊相關領域的學生需要知道如何用程式解決問題，可能需要設計複雜的演算法，撰寫困難的程式；而工學院其他科系則注重 C 程式或 C++ 與硬體結合應用在工程上。欲檢測其程式設計的基本邏輯，可以簡化地檢驗其輸出的答案是否與標準答案相同，因此全世界已有不少線上程式測驗或比賽，範例如下：

一、 全國大學程式能力檢定(CPE)

<https://cpe.cse.nsysu.edu.tw/>

二、 University of Valladolid Online Judge(UVA) (1700+ problems):

<https://uhunt.onlinejudge.org/>

三、 Peking University Online Judge (3000+ problems):

<http://poj.org/>

四、 Caribbean Online Judge(3400+ problems)

<http://coj.uci.cu/>

五、 SPOJ Online Judge(6800+ problems):

<http://www.spoj.com/>

六、 Timus Online Judge(1100+ problems):

<http://acm.timus.ru/>

七、 Code Forces(5600+ problems):

<https://codeforces.com/>

其他領域如管理學院、社會科學院的學生或許只是需要設計出一個簡單的介面與功能來作為課程或工作上的輔助，使用上述的線上檢測工具，可能並非相當合適。此時，可以用目前興盛的 App Inventor 2 來實現。但由於 App Inventor 2 是新興的程式設計管道，目前並沒有如 CPE 一樣的電腦批改系統，必須使用人工批改這樣耗時耗力又容易出錯的批改方法，因此本計畫的目標是，製作一套 App Inventor 2 的自動批改系統，未來可應用於各大院校的程式能力檢定上。

App Inventor 2 程式設計完成後，會產生 apk 檔案。我們可以透過安裝並執行去檢查是否達成題目要求的功能。為了使用檔案輸入輸出來進行功能的批改，我們將規定考生要把通常以點擊按鈕執行的程式，功能的部分獨立開以副程式撰寫，點擊按鈕則會去呼叫該副程式，執行的結果相同，卻能讓我們以呼叫該副程式的方式進行檔案輸入輸出，並規定考生某些特定元件或變數的名稱，以達成自動批改的功能。

一般的線上程式批改（如 CPE），無法評判互動式程式，因此我們的 App Inventor 2 自動批改系統也必須將考題類型限制在靜態考題上。以 CPE 題目範例說明，CPE 的每一個題目的輸入資料格式均有明確說明，透過讀檔一次讀取所有輸入，再檢查全部輸出是否正確。CPE 不會有類似「請輸入一個整數」的交談式提示，或者根據前一個答案而輸入不同的資料，因為這樣的逐行輸入的程式無法使用自動批改系統。此外，自動批改系統也無法進行繪圖板、語音辨識、需要使用網路、需要動態時間的考題等。CPE 網站有歷次的考試題 <https://cpe.cse.nsysu.edu.tw/history.php>，可供參考。

貳、文獻回顧與探討

一、 App Inventor 2

App Inventor 2 是一個視覺化的寫程式環境，操作方式類似於 scratch，都是利用拖曳積木的方式實現程式應用，使用者不用先學習程式語法，更容易上手。因此 App Inventor 2 容易讓人寫出功能完整的應用程式，且能實際在智慧型手機或平板電腦上操作。

在 App Inventor 2 之前，MIT 曾發行 App Inventor（現更名為 App Inventor classic，並於 2015 年終止服務）。

二、 AI2Offline

使用 App Inventor 2 網頁版進行 App 開發需要連接網際網路並使用 Google 帳號登入，但我們的考試環境僅有教室內部的區域網路，因此以 AI2Offline 作為替代的撰寫環境，同時也能避免考生事先將程式撰寫完成放置於他的個人帳號中。AI2Offline 為離線版的 App Inventor 2，可以一鍵化的啟動本機中 App Inventor 2 伺服器(Dev Server)、Build 伺服器與 AI Starter。相較於其他的離線版 App Inventor 2 更為直觀好上手。

三、 App Inventor 2 基礎指令與元件指令

App Inventor 2 的指令，分成兩種：基礎指令與元件指令。分述如下：

(一) 基礎指令（共 9 種）

流程控制指令（共 16 項）	if-else、for 迴圈、while 迴圈、條件運算子、螢幕切換指令
邏輯指令（共 6 項）	true、false、not、==、!=、&&、

數學指令（共 27 項）	常數、數字比較、四則運算、指數、二進制運算（and、or、xor）、隨機數字指令、最大最小值、平方根、絕對值、相反數、四捨五入、無條件進位或捨去、模數（module）、三角函數、角度弧度轉換、設定小數位數、判斷是否為數字、進位轉換
文字指令（共 15 項）	字串、字串合併、字串長度、判斷是否為空字串、字串比較、刪除空格、大小寫轉換、子字串起始位置、判斷是否包含子字串、以特定符號分解字串、以空格分解字串、擷取字串、子字串取代、文字加密、判斷是否為字串
清單（起始 index 為 1 的陣列）（共 22 項）	建立陣列、增加元素、搜尋陣列、求陣列長度、隨機選取元素、求目標在陣列中的 index、求陣列中特定 index 的元素、插入元素、取代元素、刪除元素、合併陣列、複製陣列、判斷是否為陣列、反轉陣列、CSV 轉換指令、查找相關元素、合併元素為字串
顏色（共 3 項）	RGB 合成與分解
變數（共 5 項）	全域變數與區域變數
程序（共 3 項）	定義與呼叫副程式
Dictionaries（共 17 項）	建立字典、增加字典項、變更字典項值、移除字典項、取得值、取得索引、判斷索引是否存在、取得字典大小、清單和字典的雙向轉換、複製字典、合併字典、判斷是否為字典

(二) 元件指令

App Inventor 2 的畫面用拖曳元件的方式編排，常用的元件如按鈕 (Button)、複選盒 (CheckBox)、標籤 (Label)、文字輸入盒 (TextBox) 等。在程式指令中，除了上述基礎指令，由於 app 要根據使用者的行為作出回應，例如按下按鈕、複選盒被選取等，因此不同的元件又有不同的指令。

四、 Client judge 與 Server judge

參考現行的 online judge，形式分為 client judge 與 server judge 兩種。client judge (例如 CPE) 的主要工作在客戶端，也就是考生的電腦上執行程式碼，再將執行結果送往伺服器端進行成績的彙整；server judge (例如大多數的線上評審) 則是將程式碼送到伺服器端，由伺服器端做執行後，進行答案的比對。這兩種方法各有優缺點，client judge 的答案可能是考生事先偷到答案後偽造的，而並非由正確的程式碼產生；而 server judge 則是需要伺服器端龐大的計算能力，無法同時應付大量程式的執行。

我們的自動批改系統預計採用 server judge 的形式，由於用途為大學生程式能力檢定，因此設計的目標為能夠同時承受一班約 50 人的規模。

五、 XAMPP

XAMPP 當中包含 Apache、MySQL、PHP、Perl，如果要將網頁上的每個服務，包括網頁伺服器、PHP 直譯器、資料庫伺服器等都自己架設，不僅架設麻煩也不易於管理，XAMPP 整合了這些功能讓使用者可以一鍵啟動，快速架設網站。

我們安裝的是 Windows 版本的 XAMPP，除了上述提到的功能外，另外還有 FileZilla FTP Server 和 Tomcat。本計畫使用到以下 XAMPP 內建的功能：

(一) Apache

全名為 Apache HTTP Server，是一個開放原始碼的網頁伺服器(Web Server)軟體，因為它的快速、易於擴充和跨平台等特性，Apache 是目前最流行的網頁伺服器。

(二) MySQL 和 phpMyAdmin

MySQL 是一種開放原始碼的關聯式資料庫管理系統，使用結構化查詢語言(Structured Query Language, SQL)進行資料庫管理。

phpMyAdmin 是以 PHP 為基礎、架構在網站主機上的 MySQL 資料庫管理工具，讓使用者用 web 介面管理資料庫，相較於使用 SQL 語法更簡單、方便，尤其在資料匯入、匯出支援多種檔案格式（例如：CSV 格式、Word、PDF 和 XML 等）。另外，由於 PHP 程式是在網頁伺服器上執行，產生出 HTML 頁面，因此使用者可以遠端管理 MySQL 資料庫。

(三) PHP

PHP 全名為 Hypertext Preprocessor，融合了 C、JAVA、Perl 語言的特點，PHP 用途十分廣泛，主要用來網頁開發，PHP 是嵌入 HTML 中執行，可以快速執行動態網頁，也可以間接存取 MySQL 資料庫中的數據。

(四) Tomcat

Tomcat 本身也包含 HTTP 伺服器，提供 JAVA 與 JSP 語言的網頁程式支援。

六、 Bluestacks

Bluestacks 是一款可以在 Windows 環境中執行的 Android 模擬器。App Inventor 2 提供多種方式讓開發者測試 App，其中打包 apk 並下載不需要網路連線，將下載的 apk 檔拖曳至 Bluestacks 中即可測試 App。

七、 Anbox – Android in a box

由於 Windows 大多不支援指令控制，我們將使用支援 linux 系統的 Anbox 作為執行考生繳交的程式，它的概念是將一整個 Android 系統包裝成盒子的概念，可以在筆記型電腦或手機等裝置上使用。

Anbox 的初始版本於 2017 年 4 月 11 日正式發布，旨在將 Android 上的應用程式移植到 Linux 上。Anbox 的程式碼是開源的，理論上能夠執行任何的 App，再加上它不會直接存取硬體，十分安全。

八、 Android Debug Bridge (ADB)

ADB 是一個可以讓我們用指令去與裝置（在本系統中，裝置即為 Anbox）溝通的工具，我們可以使用 ADB 的指令來安裝 App 或偵錯。在我們的批改系統中，主要將 ADB 運用於指令化 App 在 Anbox 上的安裝、執行及解除安裝。

九、 共用資料夾

本系統設計方向為供學校使用，考量現今學校多使用 Windows 作為主要的作業系統，因此我們使用虛擬機來讓兩種作業系統同時運作於一部電腦上，在實體機與虛擬機的溝通上，我們使用虛擬機內建的共用資料夾，讓兩個作業系統同時可以利用此功能來傳遞檔案。

十、 大學程式能力檢定

我們的 App Inventor 2 自動批改系統參考了 CPE 當中的「測試資料與程式執行」、「評判伺服器訊息」、「計分方式與排名」。

(一) 測試資料與程式執行

1. 每個題目，除了題目原有的輸入、輸出範例外，考生可利用「題目預設測試資料」、「使用自訂測試資料」、「人工公開測試資料」，進行程式之偵錯。此部分不登錄為「送繳」次數與「送繳」時間（不列入考試

成績)。

2. 「人工公開測試資料」的難易度與評判系統所使用的「隱藏測試資料」(不公開, 固定的) 大約相同, 建議考生通過「人工公開測試資料」之後, 再進行「正式評判」。但無法保證通過「人工公開測試資料」者, 必定可以通過「隱藏測試資料」(若通過前者, 而無法通過後者, 代表程式仍有思慮不週之處)。

(二) 考生「送繳」程式後 (CPE 評判系統的「正式評判」), 評判系統將使用「隱藏測試資料」進行程式的檢視, 此部分將登錄為「送繳」次數與「送繳」時間 (列入考試成績)。

(三) 評判伺服器訊息

COMPILER-ERROR	程式碼未通過編譯。(點入連結可以查閱編譯器所產生的錯誤訊息。)
CORRECT	程式已經正確, 並通過測試。
NO-OUTPUT	程式沒有輸出任何資料。
PENDING	送出的程式碼仍在處理中。
PRESENTATION-ERROR	輸出的結果正確, 但格式錯誤, 例如未依規定空格或換行 (多空格或少空格, 多換行或少換行)。
RUN-ERROR	無法順利將程式執行完畢, 亦即程式執行過程發生錯誤, 例如記憶體存取錯誤。
TIMELIMIT	程式執行所花費的時間超過題目限制。程式可能落入無窮迴圈, 或是必須改進解題方法。

WRONG-ANSWER	輸出的結果錯誤。(若輸出的格式產生過大錯誤，也可能造成此結果。)
--------------	----------------------------------

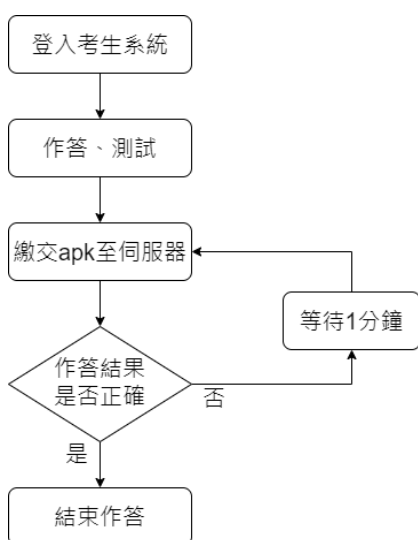
(四) 計分方式與排名

採用 ACM ICPC 評分方式，送繳程式之結果，只有「對」與「錯」。答對題數較多者，排名較前。答對題數相同者，以解題時間總和決定排名。解題時間總和，係指考試開始至解題正確所經過的時間，再加上罰扣時間（每送出題解錯誤一次罰加 20 分鐘）。答錯的題目不計時間及罰扣時間。

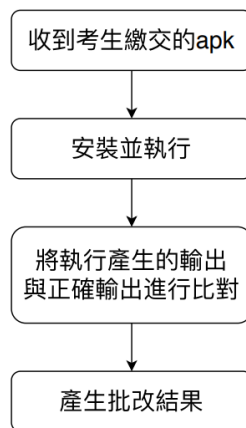
參、研究方法及步驟

本計畫研究進行步驟可分為：取得 apk 檔案、安裝 apk 檔並執行、比對輸出檔、產生批改結果。在我們的批改系統中，考生會先經由測試題目熟悉流程後，再統一發放正式題目的 aia 檔並開始作答。

考生端流程如圖二所示，由於 Anbox 無法同時開啟多個 App，因此限制考生每隔 1 分鐘才能再次繳交 apk 檔。批改端流程則如圖三所示。



圖二：考生端流程圖



圖三：批改端流程圖

一、功能批改

考生程式碼撰寫完成後欲送繳批改，伺服器端需要取得考生的 apk 檔用於功能批改。

在 App Inventor 2 中有一個「檔案管理」元件，該元件支援程式進行檔案的讀取以及寫入，為了實現自動化批改，我們將使用讀檔與寫檔的方式批改 App 是否正確執行。

在考生下載的題目 aia 檔中，存有寫好檔案輸入輸出的程式，並規定考生的作答區在特定副程式中，因此考生作答完畢後，不僅能以手動輸入

測試 App，伺服器端也可以讓 App 讀取輸入檔並產生輸出檔，依據輸出檔與正確答案做比對，即可知道功能是否正確。

二、 回傳結果與評分

在本系統中，公開測資會讓考生於自己電腦上的 Bluestacks 安卓模擬器中執行，僅有用於正式評判之 apk 檔才會送交至伺服器端執行。進行公開測資測試時，考生須自行將輸入檔存於 Bluestacks 的讀取位置，並在執行完程式後自行到寫入位置查看輸出檔是否與測資相同。正式評判時，伺服器會回傳考生程式是否正確或是錯誤原因，本系統中將回傳結果分列 5 項，標準如下：

正確	程式已經正確，並通過測試。
格式錯誤	輸出的結果正確，但格式錯誤，例如未依規定空格或換行（多空格或少空格，多換行或少換行）。
逾時	程式執行所花費的時間超過題目限制。程式可能落入無窮迴圈，或是必須改進解題方法。
輸出錯誤	輸出的結果錯誤。（若輸出的格式產生過大錯誤，也可能造成此結果。）
執行錯誤	無法順利將程式執行完畢，亦即程式執行過程發生錯誤，例如屬性錯誤。

三、 App Inventor 2 題目範例

我們設計的一個 App Inventor 2 題目，「黃金比例計算器」以用來說明本計劃進行的流程。

黃金比例

問題敘述

黃金比例(golden ratio)又稱為黃金分割，是一個數學常數，其值約為 0.618。這個比值能讓人感受美感，常見於數理、建築和藝術。例如，長方形的寬與長之比值若為 0.618，比較容易被認為是具有美感的長方形。人體的許多部位也可用黃金比例來衡量。例如，頭頂至肚臍、肚臍至腳底的長度比值若為 0.618，就是人體的骨架符合為黃金比例。

本題請你設計程式，計算上述的比值，給予如下的評語。並於最後統計獲得三種評語的各自人數。

頭頂至肚臍、肚臍至腳底之比值	評語
比值 < 0.603	upper body too long
$0.603 \leq \text{比值} \leq 0.633$	golden ratio
比值 > 0.633	legs too long

輸入說明

第一列有個正整數 N ($1 \leq N \leq 100$)，代表有 N 個人。接下來有 N 列資料，每一列有兩個整數 A 與 B ，分別一個人的頭頂至肚臍、肚臍至腳底之長度， $50 \leq A, B \leq 200$ 。長度的單位皆為公分。

輸出說明

對於每一個人，於一列輸出其比值（四捨五入至小數點後 3 位，列印時需輸出至小數點後 3 位），以及其對應的評語。最後一列輸出三個整數，分別代表身體太長的人數、黃金比例的人數、腿太長的人數。資料與資料之間需有一個空白。

輸入範例

4

165 95

165 100

165 105

165 110

輸出範例

0.576 upper body too long

0.606 golden ratio

0.636 legs too long

0.667 legs too long

1 1 2

App Inventor 2 說明：

輸入：

兩個文字輸入盒元件” TextBox_N” 及 ” TextBox_input” 分別讀取總人數 N ($1 \leq N \leq 10$) 與 N 個人各自的身高整數 X 與肚臍到腳底長度整數 Y ，身高與肚臍到腳底長度的單位皆為公分，且以空格區隔，格式為 $X_1 Y_1 X_2 Y_2 \dots X_N Y_N$ 。

輸出：

按下按鈕元件” Button_result” 後計算所有資料的黃金比例，計算結果顯示在標籤元件” Label_output” 中。對於每一個人，以空格作為間隔輸出他的比例（四捨五入至小數點後 3 位）和評語，不同人以換行字元’ \n’ 分開，最後一列輸 3 個整數，分別代表黃金比例的人數、身體太長的人數、腿太長的人數，以空格區。

範例介面：

Screen1

黃金比例計算器

計算



範例輸入輸出：

Screen1

黃金比例計算器

4

165 95 165 100 165 105 165 110

計算

0.576 upper body too long
0.606 golden ratio
0.636 leg too long
0.667 leg too long
1 1 2

四、 App Inventor 2 功能批改範例

題目的專案檔中預設會有檔案管理的積木程式，使程式輸入可以手動輸入或讀取 input.txt 檔案，且程式執行的同時也會有輸出檔 output.txt。圖四為範例的 input.txt，圖五為範例的 output.txt，透過比對 output.txt 的內容與正確答案，即可知道功能是否正確，圖六為讀取 input.txt 程式執行的畫面。



```
4
165 95 165 100 165 105 165 110
```

圖四：input.txt 內容



```
0.576 upper body too long
0.606 golden ratio
0.636 legs too long
0.667 legs too long
1 1 2
```

圖五：output.txt 內容



Screen1

黃金比例計算器

4

165 95 165 100 165 105 165 110

計算

0.576 upper body too long
0.606 golden ratio
0.636 leg too long
0.667 leg too long
1 1 2

圖六：以讀檔方式執行的黃金比例計算器

肆、成果展示

以下將展示本系統實際操作之畫面，分為考生端及批改端。

一、 考生端

進入考場後，考生電腦會顯示登入畫面（如圖七、圖八），考生根據分配到之帳號及密碼登入系統。

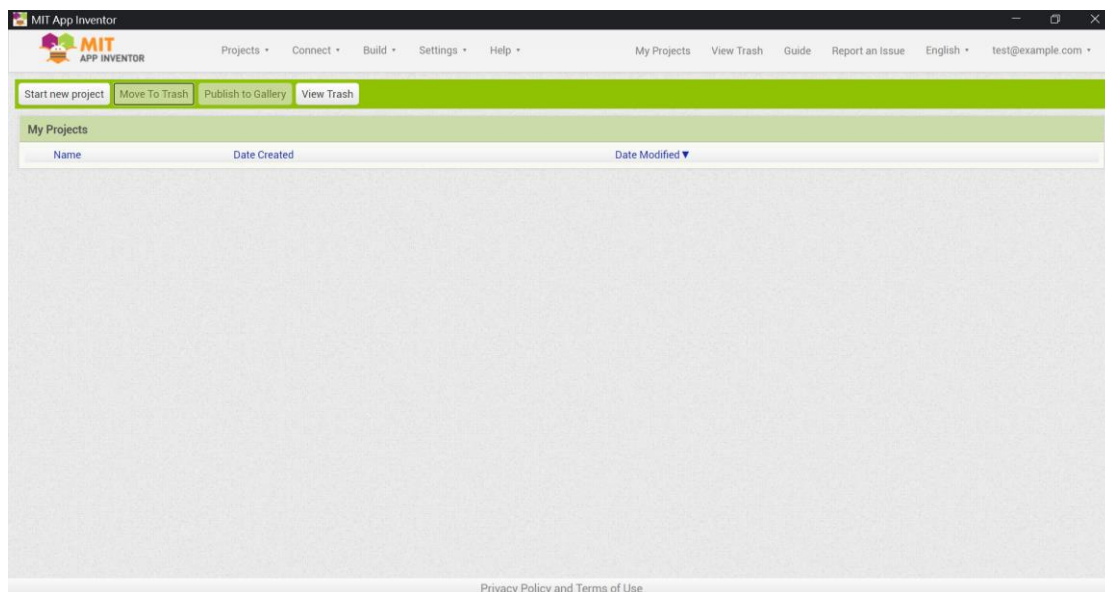


圖七：考生電腦登入前

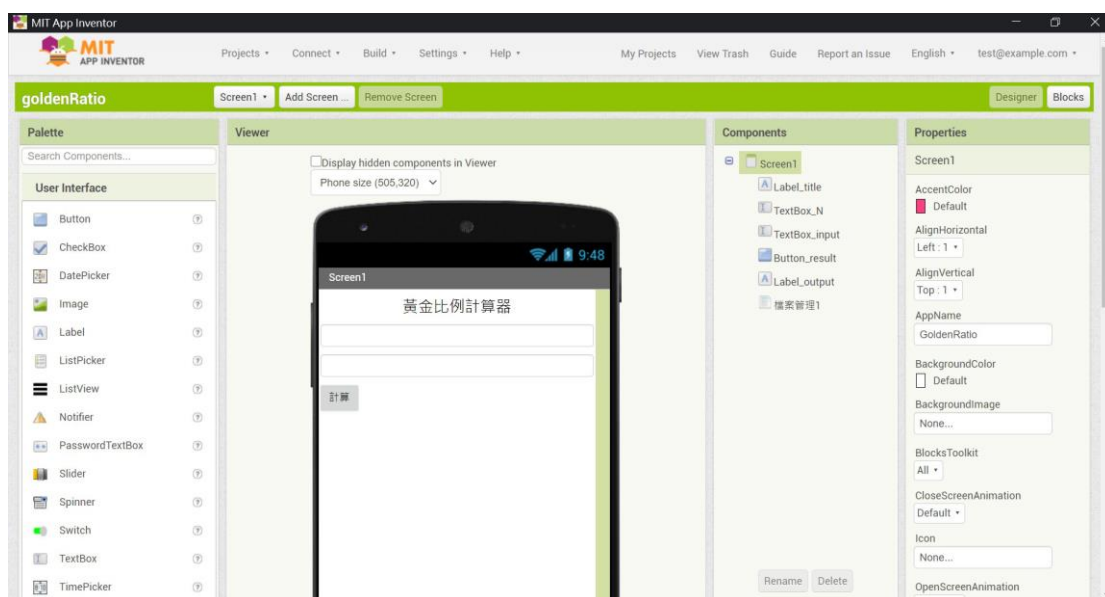


圖八：考生電腦登入後

在作答階段，考生需點擊下載模板以取得我們預先撰寫好之 aia 檔，再前往 AI2Offline 匯入專案並撰寫程式，如圖九、圖十。



圖九：專案檔匯入前



圖十：專案檔匯入後

考生於作答區撰寫好程式後，將該專案檔打包成 apk 檔，可以選擇先將檔案安裝至 BlueStacks 進行公開測資測試，或是直接上傳至批改端進行評判。當考生選擇上傳後，該按鈕會停止作用 1 分鐘，待批改端產生結果後，該題會由尚未繳交改變為批改結果，如圖十一所示。



圖十一：考生收到批改結果後

此時，考生亦可由查看成績前往排行榜觀看自己與其他考生之成績，如圖十二所示。

排行榜

上一頁

學號	答對題數	答題時間	第一題	第二題
B063040001	0	0	0	0
B063040002	1	31	1(31+0)	2
B063040003	2	105	1(27+0)	2(58+20)
B063040004	1	57	0	2(37+20)
B063040005	0	0	3	2
B063040006	1	73	2(53+20)	1
B063040007	1	44	1(44+0)	0
B063040008	0	0	1	2
B063040009	1	24	0	1(24+0)
B063040010	1	39	2(19+20)	0
B063040011	0	0	3	0

圖十二：查看排行榜

最後，點擊結束測驗即可離開本系統。

二、 批改端

批改端由 shell script 偵測是否收到考生提交之 apk 檔，當有新的檔案提交時，該 script 會安裝該檔案至 Anbox 並進行檢查。apk 的安裝結果可分為以下三種：

(一) 安裝失敗

安裝失敗可能是因為該檔案由改變副檔名而產生等原因造成，會以執行錯誤回傳。

(二) 安裝成功，但參數錯誤

參數錯誤的原因可能為考生提交了非本題之專案或變更了專案名稱，此部分也以執行錯誤回傳。

(三) 安裝成功，且參數正確

批改端將會呼叫並執行此 App。

本系統目前給予每個 App 各 2 秒的安裝時間，4 秒的執行時間，當 App 正常執行並產生輸出檔後，script 會比對輸出是否正確；若無輸出檔產生，程式將會回傳逾時。

以上步驟執行完畢後，除了產生批改結果外，script 還會將排行榜的內容更新，並將所有安裝成功的 apk 檔保留下來以供日後查看。

伍、結果與討論

目前程式功能批改採用與 CPE 相同的方式，已能夠過濾程式功能不完備的檔案，介面設計彈性較高，暫定以人工方式批改。

考試系統中，考生端大部分已經開發完成，後續將著重在批改端功能的介面化，預計將設計管理者的介面，包含能夠建立考試、以 excel 檔案輸入考生資料及輸出考試成績等功能。

陸、參考文獻

一、 App Inventor 2

<http://appinventor.mit.edu>

二、 TQC 舉辦的 App Inventor 2 設計比賽

https://www.tqcplus.org.tw/content_brochure_IAI.asp

三、 App Inventor 2 source

<https://github.com/mit-cml/appinventor-sources>

四、 App Inventor 2 指令說明

http://www.appinventor.tw/ai2_chinese

五、 全國大學程式能力檢定(CPE)

<https://cpe.cse.nsysu.edu.tw/>

六、 University of Valladolid Online Judge(UVA)

<https://uhunt.onlinejudge.org/>

七、 Peking University Online Judge

<http://poj.org/>

八、 Caribbean Online Judge

<http://coj.uci.cu/>

九、 SPOJ Online Judge

<http://www.spoj.com/>

十、 Timus Online Judge

<http://acm.timus.ru/>

十一、 Code Forces

<https://codeforces.com/>

十二、 自己架一個 online judge 系統

<https://blog.techbridge.cc/2020/03/20/build-your-own-online-judge/>

十三、 Anbox

<https://anbox.io/>

十四、 XAMPP 架設網站

<https://www.kocpc.com.tw/archives/165133>

十五、 XAMPP 維基百科

<https://zh.wikipedia.org/wiki/XAMPP>

十六、 Apache 網頁伺服器

<http://sggroup.com.tw/learning/apache/apachedoc.Htm>

十七、 MySQL 資料庫入門

<https://www.itread01.com/content/1550418327.html>

十八、 phpMyAdmin 維基百科

<https://zh.wikipedia.org/wiki/PhpMyAdmin>

十九、 PHP 簡介

<http://www.southmaster.com/article/pub.php?page=red2.php&id=4953>

二十、 Apache Tomcat

https://zh.wikipedia.org/wiki/Apache_Tomcat