

115年專門職業及技術人員高等考試會計師、不動產估價師、
專利師、民間之公證人、植物診療師考試試題

等 別：高等考試

類 科：專利師（選試專業英文及計算機結構）、專利師（選試專業日文及計算機結構）

科 目：計算機結構

考試時間：2小時

座號：_____

※注意：(一)可以使用電子計算器。

(二)不必抄題，作答時請將試題題號及答案依照順序寫在試卷上，於本試題上作答者，不予計分。

(三)本科目得以本國文字或英文作答。

大型語言模型（LLM, Large Language Model）如 GPT、Claude 等，需要龐大的矩陣運算、管線化推論、重複存取大量參數等。以下題目探討 LLM 在計算機組織的運算議題。

- 一、我們在電腦 A 上執行一個 LLM 推論任務需時 25 秒，工作頻率為 2 GHz。若要在電腦 B 上只用 70% 的時間完成推論，並且電腦 B 的架構調整後導致該推論所需要的時鐘週期數是 A 的 1.3 倍，功耗則是 A 的 1.2 倍：
- (一)電腦 B 應具備多少工作頻率才能達到目標？（15 分）
 - (二)電腦 B 執行該 LLM 推論所消耗的能量為電腦 A 的多少倍？（10 分）

二、CPI 與 LLM 運算的指令集效能：

- (一)CPI（cycles per instruction）是如何計算而來？（10 分）
- (二)有那些因素會影響 LLM 的訓練及運算過程的 CPI？（15 分）

三、LLM 在一個管線化（pipelining）的 CPU 裡執行：

- (一)何謂 pipelining？（5 分）
- (二)有那些因素會影響 LLM 在這 CPU 的管線裡的執行效能？（10 分）
- (三)在一個有 10-stage pipeline 的高效能處理器裡執行 LLM 的訓練。若有 10% 的指令產生 hazard，每個 hazard 需多 2 個 clock cycles，則該程式的 CPI 值為何？（10 分）

四、LLM 與記憶體階層元件（DRAM, ROM, flash memory, cache memory, hard disk）：

- (一)在 LLM 的運算中，那一種記憶體裝置最接近 CPU？（5 分）
- (二)那些記憶體在斷電後仍能保存原來的資料？（10 分）
- (三)LLM 裡的某個矩陣運算在電腦 A 上執行，若 cache miss rate = 0 時 CPI = 1.5。當 cache miss rate = 8%，miss penalty = 20 clock cycles，則該矩陣運算的 CPI 值為何？（10 分）