

等 別：四等考試

類 科：圖書資訊管理

科 目：資訊系統與資訊檢索概要

考試時間：1 小時 30 分

座號：_____

※注意：(一)禁止使用電子計算器。

(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。

- 一、試說明搜尋引擎中採用的 Pagerank 排序的原理為何？相較於相關排序（relevance ranking），採用 Pagerank 排序的優點及缺點為何？（25 分）
- 二、請論述如何利用蒐集使用者的點擊行為（click behavior）來優化包括相關排序、索引系統及緩衝儲存等搜尋引擎的效能？（25 分）
- 三、何謂分區組合檢索法（building block）？試說明若採用此一檢索法所獲得的資料量太多不夠準確，可以分別採取那些方法提升此一檢索法的資料檢索準確率（precision rate）？（25 分）
- 四、試說明文件索引（document indexing）採用字元（character）、字（word）、片語（phrase）及概念（concept）等不同層次語彙進行索引（index），對於資訊檢索的準確率（precision rate）與召回率（recall rate）的影響為何？怎麼從文件中萃取出重要且能代表該文件的關鍵字，作為該文件的索引詞（index term）？（25 分）