

1.0 请举出其他加速比定律，并且分析比较

弱缩放 (Weak Scaling)：任务总量随处理器数量增加而线性增长，目标是在增加处理器的情况下保持工作负载相同的处理器效率。**强缩放 (Strong Scaling)**：任务总量固定，随着处理器数量的增加加速计算

1.1 請舉例說明結構化數據、半結構化數據、非結構化數據的區別？

結構化數據可以使用關係型數據庫表示和存儲的數據，半結構化數據雖非可使用關係數據庫存儲，但任然可以利用結構特徵獲取和發現，非結構化數據則為沒有或難以發現固定的數據結構用於存儲資料。

1.2 請在生活中舉出一個基於“數字映像”探索或研究現實世界的實體或現象的例子，如何才能獲得更準確的數字映像？

在智能製造系統中有數字孿生的概念，例如美國國防部最早提出在數字空間建立真實飛機的模型，並透過精巧的設計新的數據獲取方式，構建一個能夠精確反映客觀世界的實體，若數據可以被有效的利用和保存，才可以在無序的數據中獲取價值，創造更準確的數字映像。

1.3 什么是大数据的 5v4v 特征？这一特征对大数据带来什么样的变化？

4v 为 volume variety velocity veracity；

5v 为 volume variety velocity veracity value。

从获取的数据质量，将真实性或准确性作为大数据的特征，对于发现事实、解释规律并预测未来提出新的挑战。

1.4 请分析相对于传统统计学而言，大数据在思维方式上的主要变化？

准确>近似 数量少>数量非常大 准确的答案>归纳的答案

1.5 结合一个具体例子，说明数据分析的一般过程？

第谷搜集了大量数量的数据给开普勒，开普勒进行分析得到天体运行的规律，求出轨道运行的规律，并推广预测了未来的走向

1.6 如何理解数据科学？

计算机技术结合数学统计的知识，将大量的数据进行科学化的理解和应用