

資料庫原理

朱克剛

何謂資料庫

何謂「庫」？

- 車庫：放車子的地方
- 倉庫：放雜物的地方
- 彈藥庫：放彈藥的地方
- 資料庫：放資料的地方

何謂「資料」？

- 在電腦中所有可以被儲存起來的東西都是資料
 - 例如：一張圖片、一段聲音、一篇文章...等

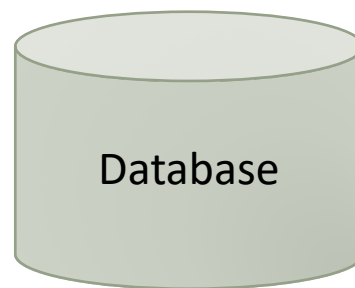
資料庫的「建築材料」

建築材料：

- 車庫、倉庫用的是鋼筋水泥
- 彈藥庫用的可能是鋼板
- 資料庫用的是檔案

檔案：

- SQL Server: .mdf
- Access: .mdb
- Oracle: .dbf
- SQLite: .sqlite



資料庫類型

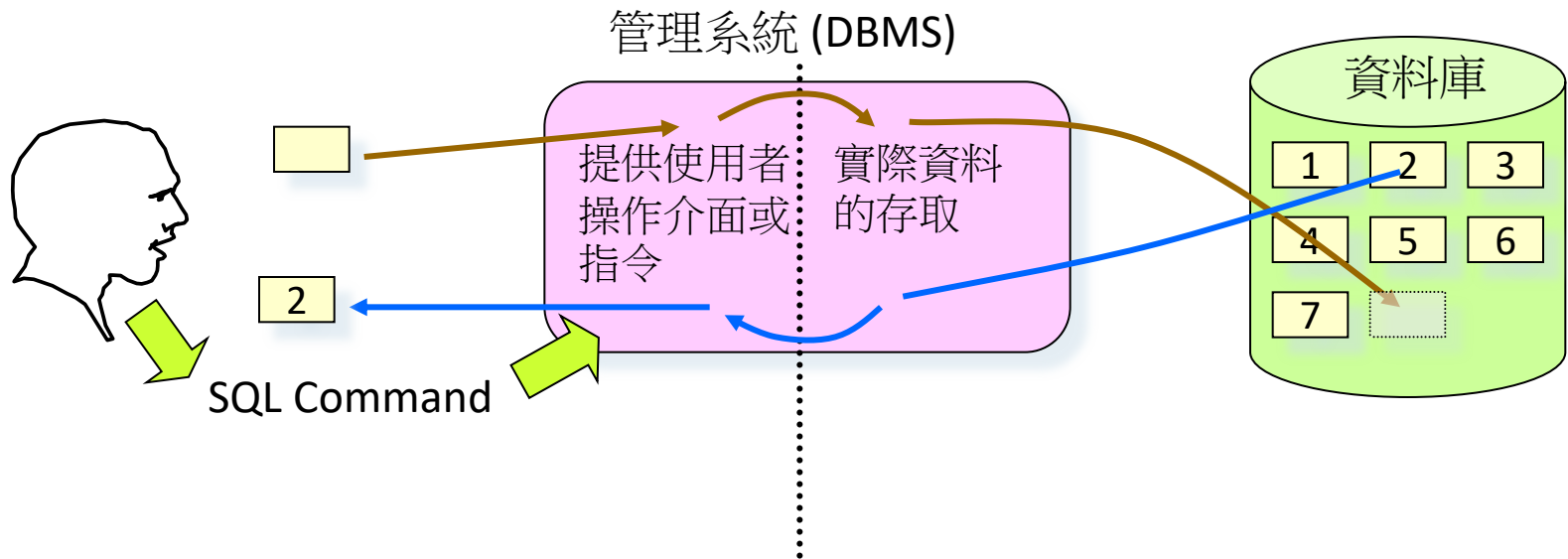
關連式

- 例如：SQL Server、Oracle、Access...等

文本式（Document-based、NoSQL）

- 例如：MongoDB、Apache Cassandra

資料庫管理系統



SQL Command 種類

資料定義語言（DDL, data definition language）

- Create: 建立資料庫物件
- Alter: 變更資料庫物件
- Drop: 刪除資料庫物件

資料操作語言（DML, data manipulation language）

- Insert Into: 插入資料
- Update: 修改資料
- Delete: 刪除資料

資料查詢語言（DQL, data query language）

- Select: 查詢資料

資料控制語言（DCL, data control language）

- Grant: 設定權限
- Revoke: 取消權限
- Commit: 確認交易完成
- Rollback: 交易失敗回復原始狀態

關連式資料庫的組成

實體（entity）-> Table

屬性（attribute）-> Field

關連性（relationship）-> 心中

通訊錄（I）

姓名 / 住址 / 電話

- 請用 Excel 建

姓名	住址	電話
王大明	台中市台灣大道一段 50 號	1111
李大媽	台中市台灣大道一段 50 號	1111
王小毛	台中市台灣大道一段 50 號	1111

資料庫中稱此
表格為**資料表**
(table)

通訊錄（II）

問題：資料重複 → 維護困難且浪費儲存空間

姓名	住址	電話
王大明	台中市公益路二段100 號	2222
李大媽	台中市台灣大道一段 50 號	1111
王小毛	台中市公益路二段100 號	2222

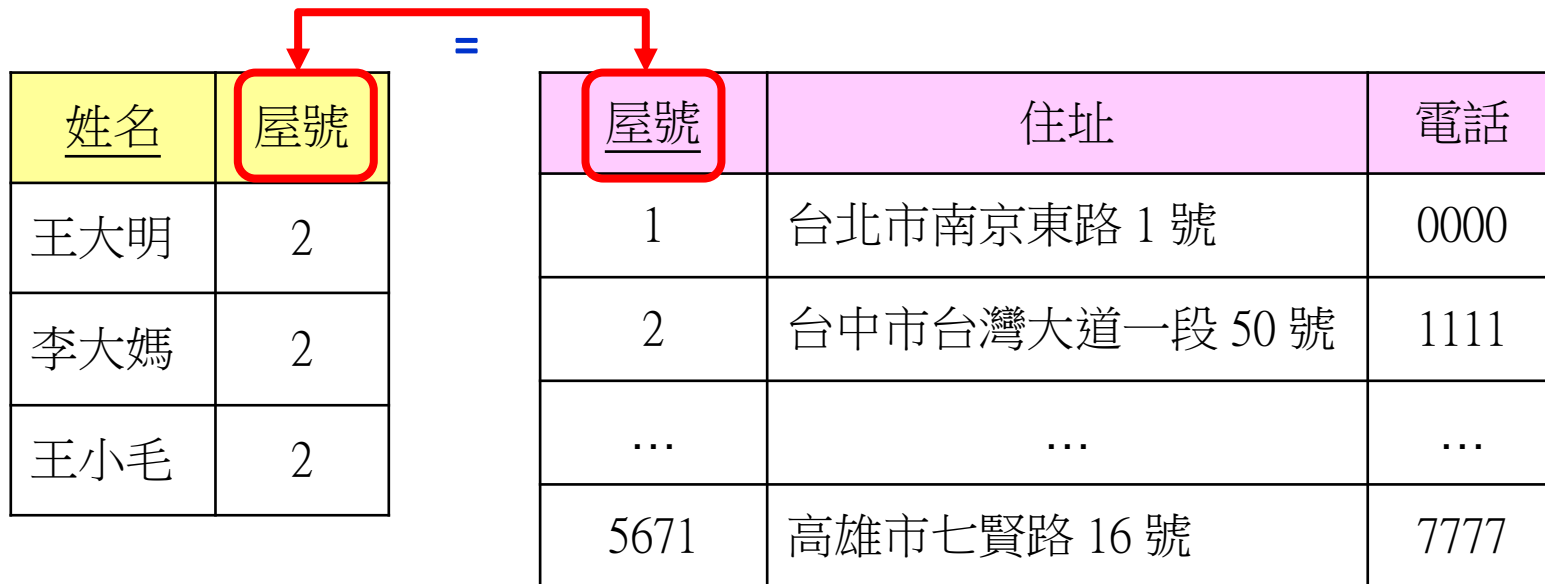
全形、半形
還是國字？

搬家後這筆資料
不小心沒改到。

數字前後要
不要空格？

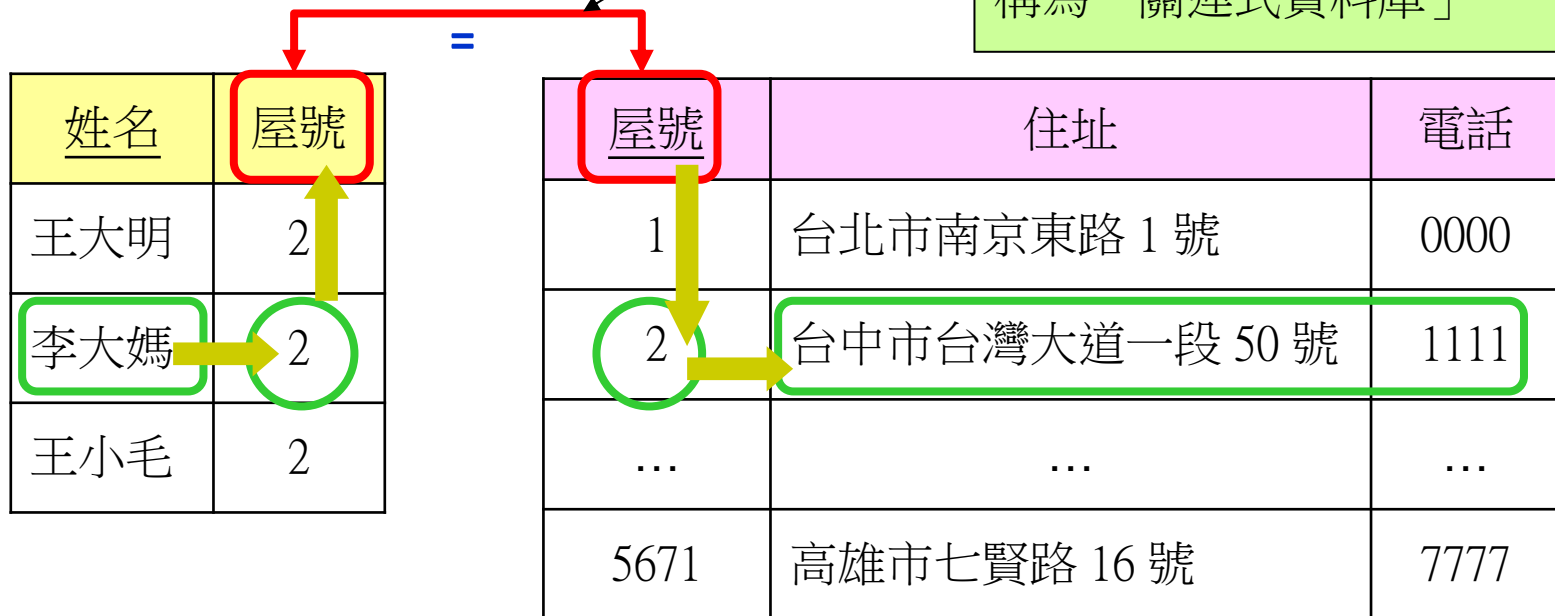
通訊錄（III）

解決方法：拆成兩個資料表



通訊錄 (IV)

查詢：李大媽住哪？



通訊錄 (V)

關連式資料庫的目的

- 解決資料重複的問題 → 維護容易
- 例如：王小毛要搬到高雄去唸書

姓名	屋號
王大明	2
李大媽	2
王小毛	5671

屋號	住址	電話
1	台北市南京東路 1 號	0000
2	台中市台灣大道一段 50 號	1111
...
5671	高雄市七賢路 16 號	7777

將 2 改為 5671 就
搬家完成了

索引 (Index)

唯一目的：加快資料搜尋速度

- 以空間換取時間
- 想想原文書最後的索引就會知道運作原理了

特性：

- 一個資料表可設定的索引數目不限
- 索引可由一個或一個以上的欄位所組成
 - 例如：可設「姓名」或是「姓名 + 屋號」

建議設索引的欄位：

- 關連線兩端的欄位
- 查詢「條件」的欄位（見SQL Command教學）

主索引、主鍵、Primary Key、PK

目的：

- 設成主索引的欄位其值：
 - 不可重複
 - 不可空白不填

特性

- 主索引通常預設是索引
- 可由一個或一個以上的欄位構成

參考索引、外來鍵、Foreign Key、FK

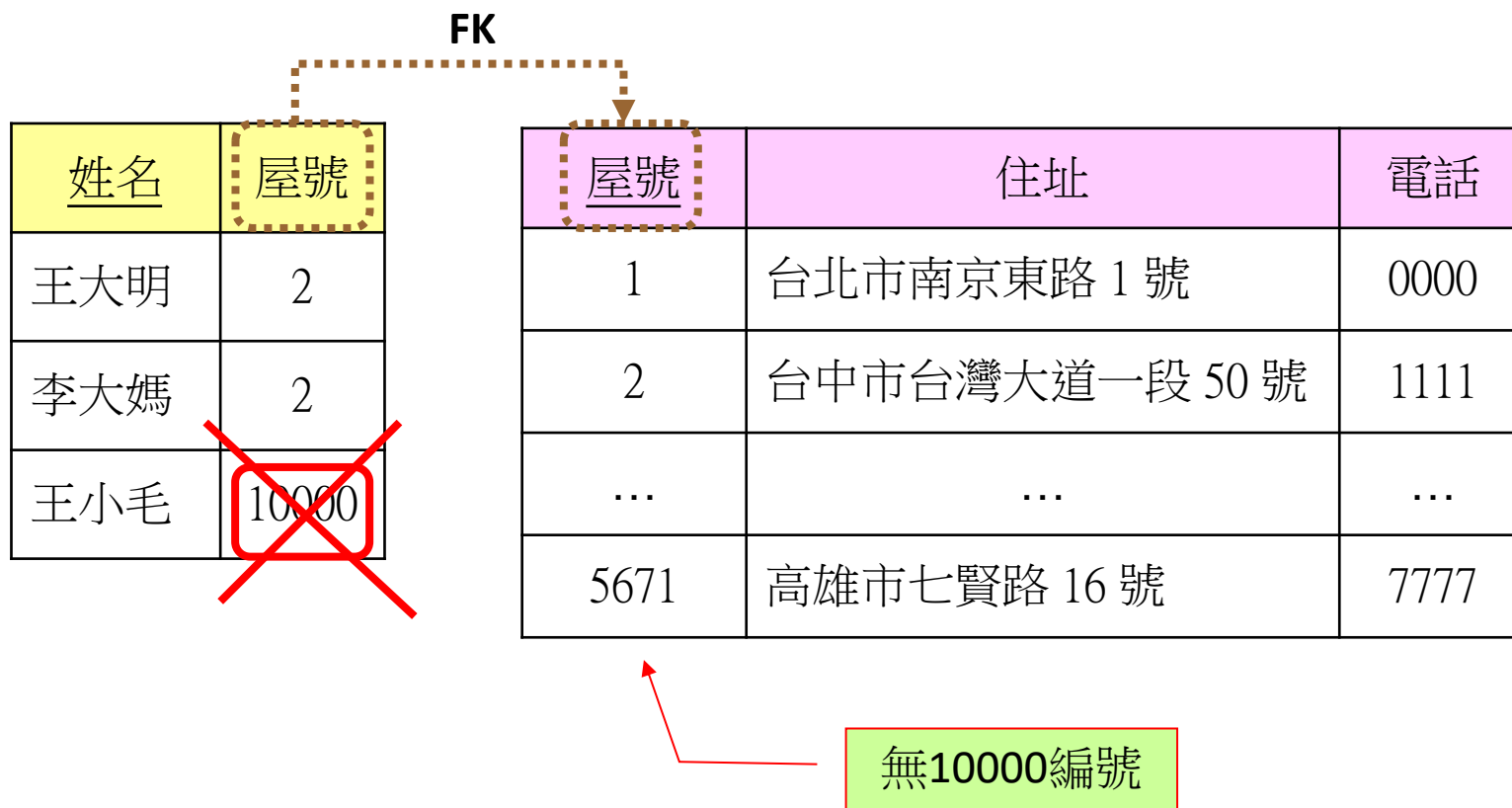
目的：

- 維持資料正確性

特性：

- 參考欄位的值只能在被參考欄位值的範圍內
- 被參考欄位必須為主索引或不可重複
- 如未設定"串接刪除" (cascade delete)，則無法直接刪除被參考資料

參考索引|範例



正規化

資料庫設計時的遵循準則

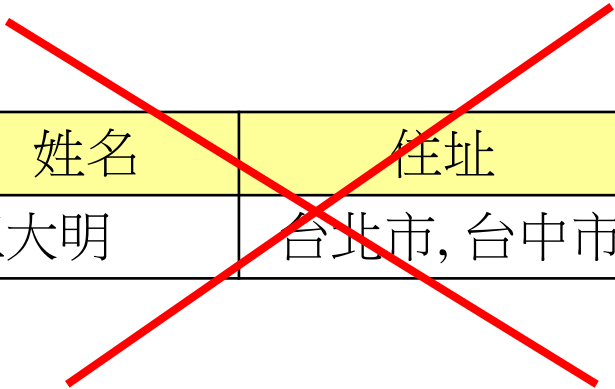
- 有 1 ~ 5 正規化以及 BC 與 DK 共 7 個正規化
- 實務上通常只用 1 ~ 3 正規化而已

越後面的正規化越可以考慮不使用，只有第一正規化一定要遵守

第一正規化 (1NF)

每一個欄位內只能放一個值

- 例如：王大明住在台北市與台中市




姓名	住址
王大明	台北市, 台中市

姓名	住址
王大明	台北市
王大明	台中市

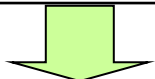
第二正規化 (2NF)

要求資料表中的所有欄位都要與PK完全依賴

違反第二正規化



<u>UID</u>	姓名	<u>住址</u>	電話
A01	王大明	台中市台灣大道一段 50 號	1111
A01	王大明	台北市南京東路 1 號	0000
C05	吳大華	高雄市和平一路 2 號	3333



<u>UID</u>	姓名	<u>屋號</u>
A01	王大明	2
A01	王大明	1
C05	吳大華	18

<u>屋號</u>	住址	電話
1	台北市南京東路 1 號	0000
2	台中市台灣大道一段 50 號	1111
18	高雄市和平一路 2 號	3333

第二正規化

又違反第二正規化



<u>UID</u>	姓名	<u>屋號</u>
A01	王大明	2
A01	王大明	1
C05	吳大華	18
A02	李大媽	2

<u>UID</u>	姓名
A01	王大明
C05	吳大華
A02	李大媽

<u>UID</u>	<u>屋號</u>
A01	2
A01	1
C05	18
A02	2

<u>屋號</u>	住址
1	台北市南京東路 1 號
2	台中市台灣大道一段 50 號
18	高雄市和平一路 2 號

第三正規化 (3NF)

非鍵欄位之間彼此不能依賴

違反第三正規化

<u>電話</u>	屋號	電信業編號	業者名
2222	2	C1	中華
3333	3	C2	遠傳

<u>電話</u>	屋號	電信業編號
2222	2	C1
3333	3	C2

<u>電信業編號</u>	業者名
C1	中華
C2	遠傳

思考

電話如果跟屋子資料表放在一起，有違反哪一個正規化嗎？如果沒有，那在資料儲存上是否有什麼限制？如果有，應該怎麼改？

關連圖

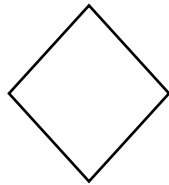
Entity-Relationship (ER) Diagrams

目的：

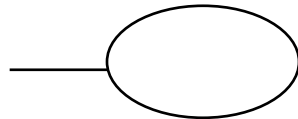
- 此為資料庫的「架構圖」，通常在系統分析時製作，確定無誤後才開始實作資料庫



Entity (資料表)

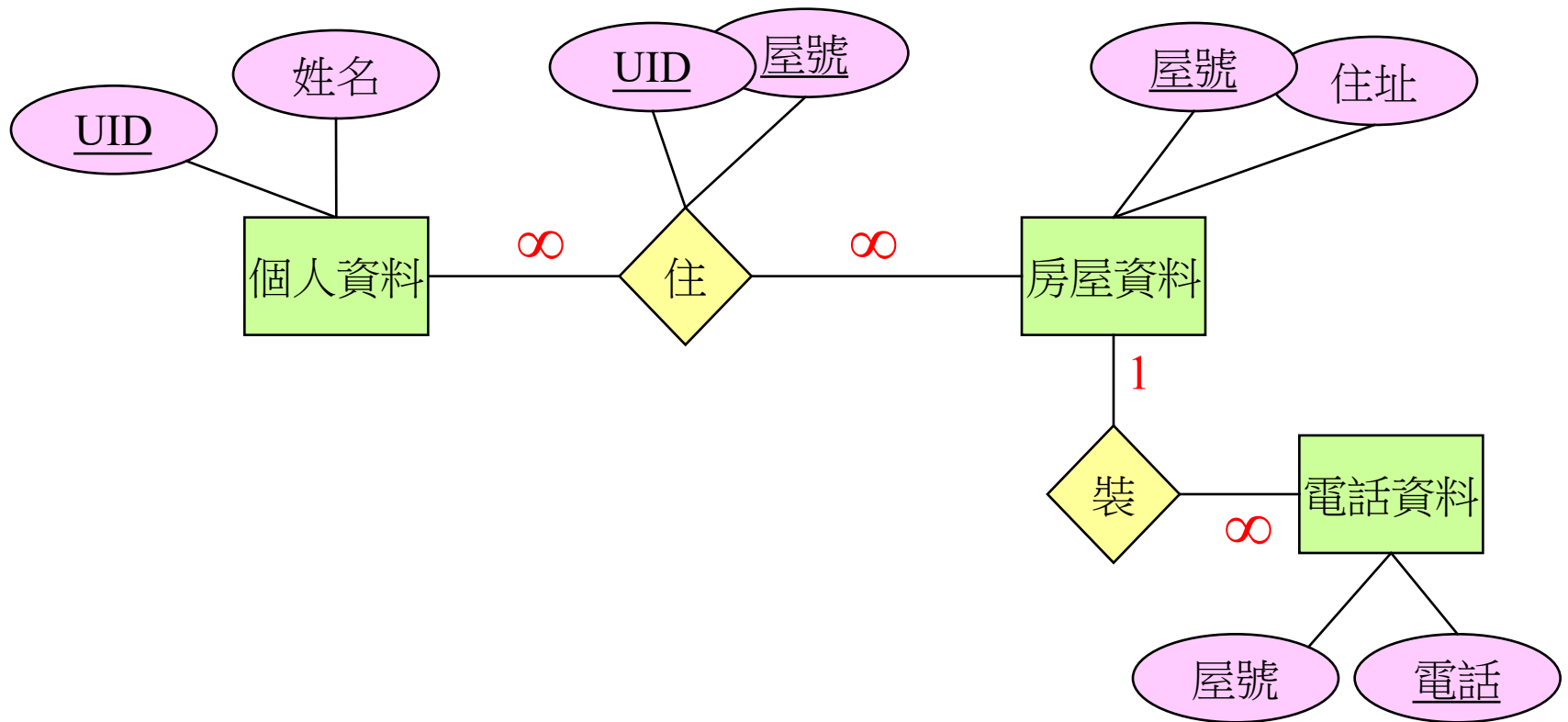


Relationship (關連)



Attribute (欄位)

ER範例



資料字典

個人資料

序號	說明	欄位名稱	資料型態	NULL?
*1	身份證字號	UID	CHAR(10)	N
2	中文姓名	CNAME	NVARCHAR(50)	Y