

# 数据库系统概论新技术篇

## 图数据管理概述

卢卫

中国人民大学信息学院

2017年6月

# 提纲

- ❖ 基本概念
- ❖ 图数据管理技术的产生和发展
- ❖ 图数据管理的应用及面临的挑战



# 提纲

## ❖ 基本概念

❖ 图数据管理技术的产生和发展

❖ 图数据管理的应用及面临的挑战



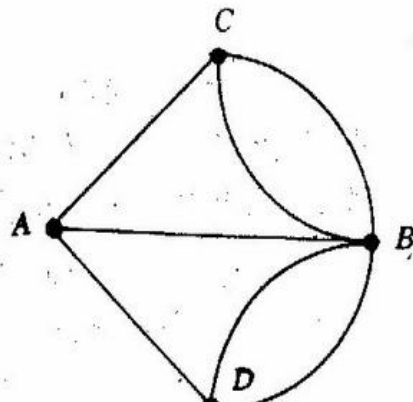
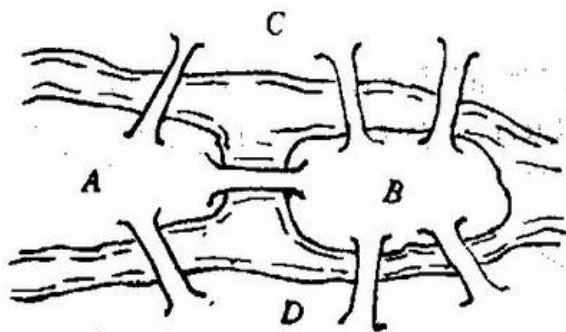
# 基本概念-图

❖ 一种灵活的数据结构，现实生活当中各类应用都可以用 **图** 来进行建模

■ 1736哥尼斯堡七桥问题



欧拉



开创了数学的一个新的分支

——图论与几何拓扑

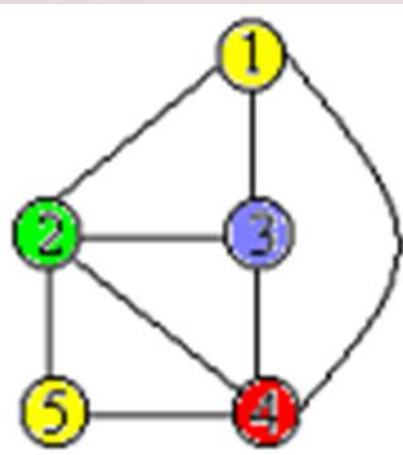
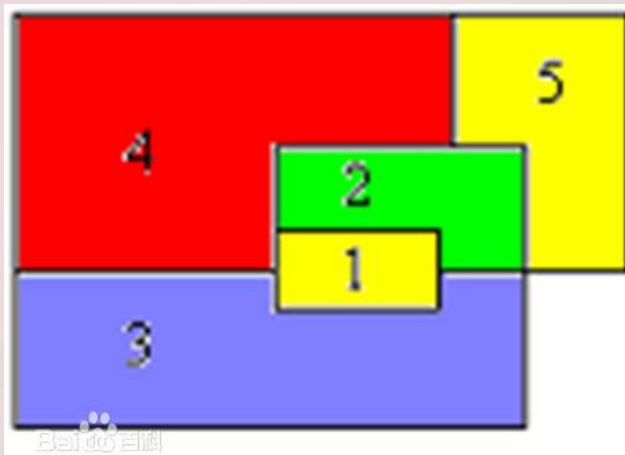
# 基本概念-图

- ❖ 一种灵活的数据结构，用来模型化数据对象之间的关系
  - ❖ 图可以对广泛的、各种各样的实际应用进行建模
    - 七桥问题
    - 旅行商问题
- 



# 基本概念-图

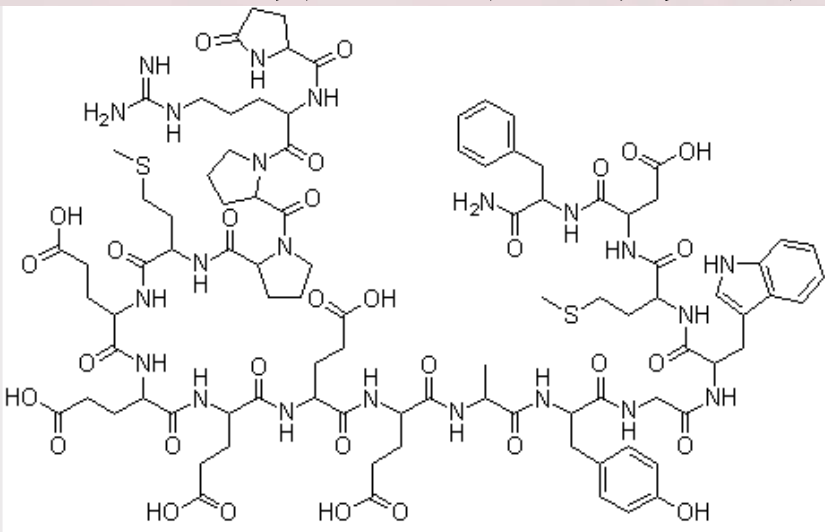
- ❖ 一种灵活的数据结构，用来模型化数据对象之间的关系
- ❖ 图可以对广泛的、各种各样的实际应用进行建模
  - 七桥问题
  - 旅行商问题
  - 图着色问题



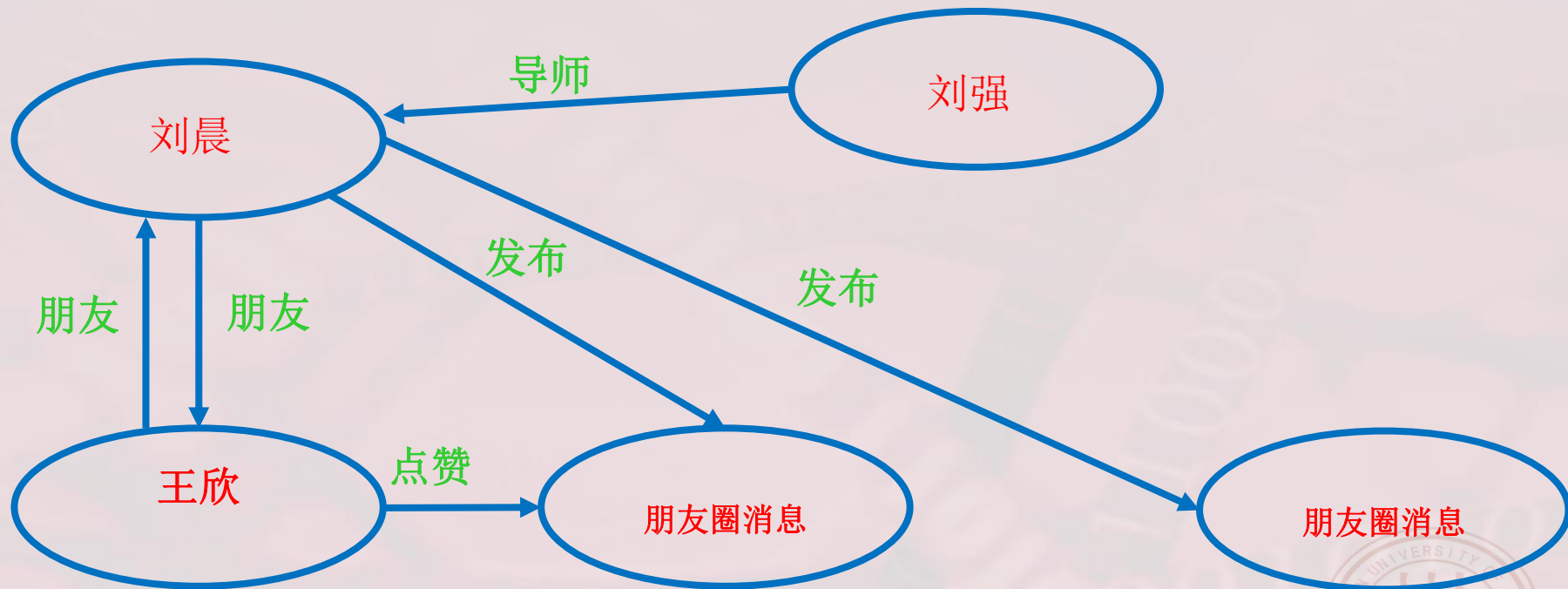


# 基本概念-图（续）

❖ 图的表示方法。四元组： $(V, E, L, F)$ ； $V$ 表示顶点的集合， $E$ 表示边的集合， $L$ 表示标签的集合， $F$ 是标签生成函数，用来为图中顶点和边生成标签



# 基本概念-图（续）





# 基本概念-属性图

- ❖ 属性图由顶点、边、标签、和属性组成
- ❖ 顶点可以包含一组属性，每个属性由键值对组成，即属性名和属性值
- ❖ 顶点可以被赋予一个或多个标签，每个标签代表该顶点的类别
- ❖ 边是有方向的，边 $v_i \rightarrow v_j$ 表示源点 $v_i$ 到终点 $v_j$ 的联系
- ❖ 类似于顶点，边也可以包含属性和标签



# 基本概念-属性图（续）

:Person

:Master



# 基本概念-图数据模型

## ❖ 什么是数据模型

- 数据结构、数据操作、数据的完整性约束条件

## ❖ 从关系数据模型→图数据模型

关系数据模型

关系

关系代数

实体完整性、参照完整性、用户自定义完整性

# 基本概念-图数据模型

## ❖ 什么是数据模型

- 数据结构、数据操作、数据的完整性约束条件

## ❖ 从关系数据模型→图数据模型

关系

关系数据模型 并、交、差、笛卡尔积、选择、投影、连接、除法

实体完整性、参照完整性、用户自定义完整性

# 基本概念-图数据模型（续）



属性图

图代数？

完整性约束



# 基本概念-图数据模型（续）



## 属性图

可达性查询；最短路径查询；图匹配；节点排序；聚类；频繁模式挖掘

## 完整性约束？





# 基本概念-图数据模型（续）



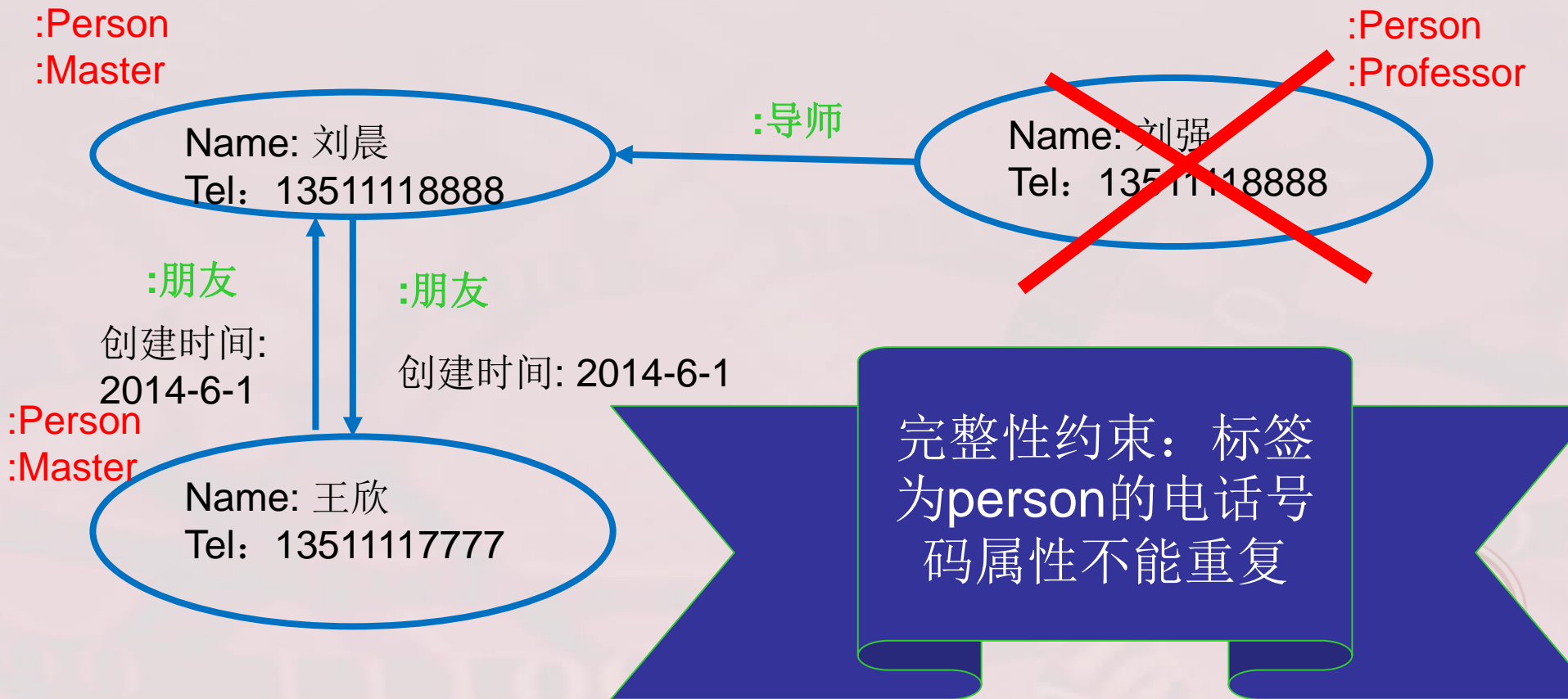
## 属性图

可达性查询；最短路径查询；图匹配；节点排序；聚类；频繁模式挖掘

## 完整性约束？



# 基本概念-属性图（续）



# 提纲

- ❖ 基本概念
- ❖ 图数据管理技术的产生和发展
- ❖ 图数据管理的应用及面临的挑战



# 图数据管理技术的产生和发展



1970s

生物信息学  
社交网络

2010s

层次模型、网状模型、关系模型

网状模型表达能力最强；网状数据库之父、图灵奖获得者—**Charles Bachman**  
但结构复杂、查询语言不易掌握和使用、数据操作也较为复杂

面向事务型的图数据库

面向分析型的分布式图处理系统



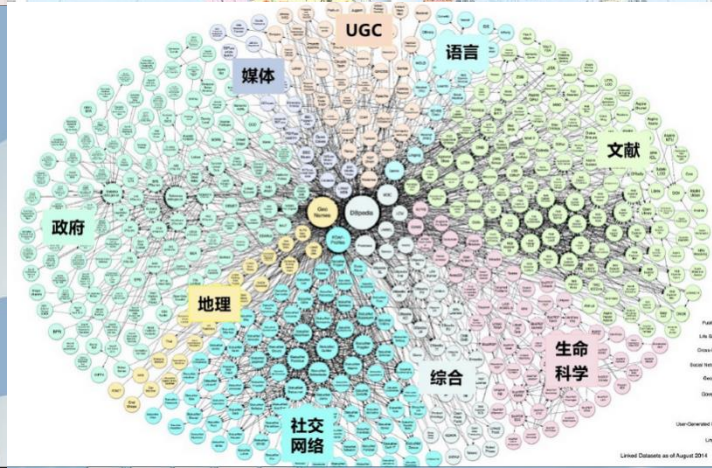
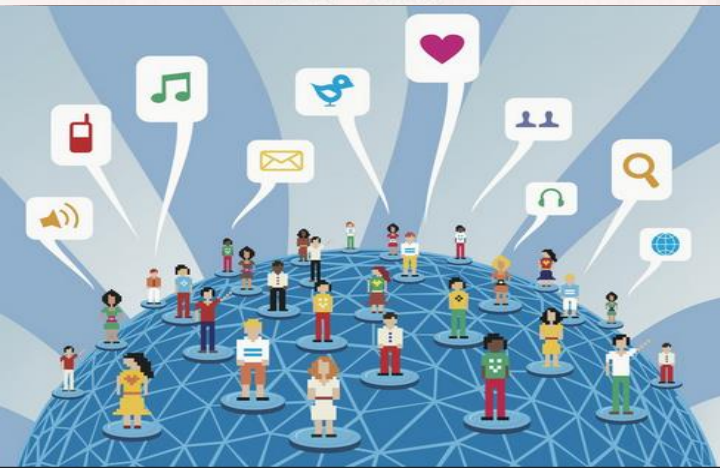
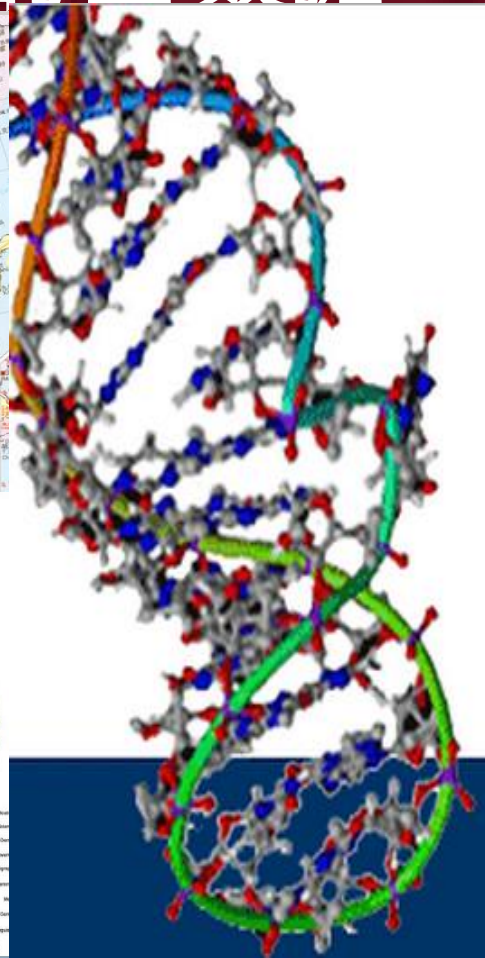
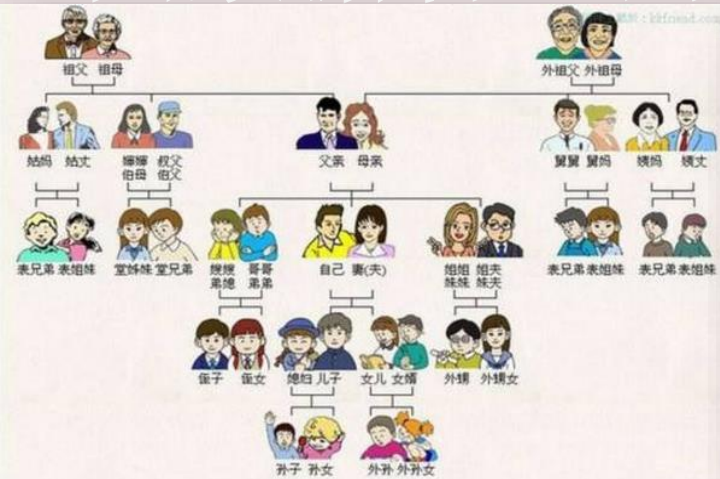
# 提纲

- ❖ 基本概念
- ❖ 图数据管理技术的产生和发展
- ❖ 图数据管理的应用及面临的挑战





# 图数据管理的应用-无处不在的图数据





# 图数据管理面临的挑战



❖ 应用需求的不同，图数据库管理系统分类：

■ 面向某一特定应用的专用图数据库管理系统

- Oracle RDF、HP Jena、RDF 3X

■ 通用图数据库管理系统

- 面向事务型（对应于关系型 RDBMS）
- 面向分析型（对应于数据仓库）

