

[30240604 面向计算机科学的离散数学-图论 2023]

面向计算机科学的离散数学

图论—绪论

苏航

suhangss@mail.tsinghua.edu.cn

清华大学 计算机系

本节课主要内容

◆ 课程介绍

- 《离散数学2》（图论）

◆ 初识图论

- 往届成果

- 该如何学习图论？

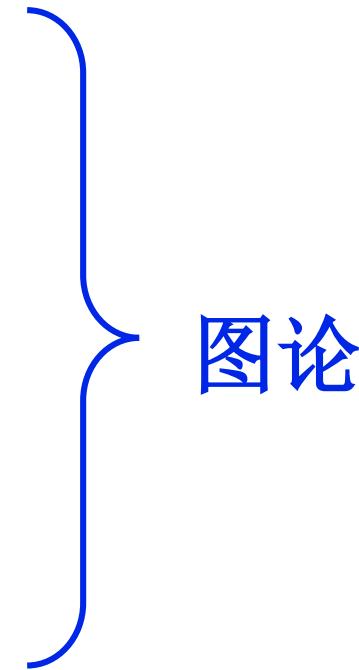
◆ 图的基本概念和定义

课程目标

- ◆ 调研能力：广博群书、挖掘与筛选资料，提取知识
- ◆ 表达能力：用严谨的术语准确讲解知识点
- ◆ 创新思维的培养与凝练：以创新思维思考知识点，
挖掘其中逻辑关系
- ◆ 核心思路
 - 为什么 (Why)
 - 是什么 (what)
 - 难在哪 (Challenges)
 - 怎么做 (How)

课程安排

- ◆ 第一章 基本概念
- ◆ 第二章 道路与回路
- ◆ 第三章 树
- ◆ 第四章 平面图与图的着色
- ◆ 第五章 匹配
- ◆ 第六章 网络流
- ◆ 第七章 图的连通性



图论

本节课主要内容

- ◆ 课程介绍
- ◆ 初识图论与往届成果
- ◆ 图的基本概念和定义

放心去飞

◆ 什么是图？

- 一笔画图
- 地图
- 交通流
- 计算机网络
- 集成电路设计
- 电厂仓库选址
- 生产物流系统
-

◆ 什么图论问题？

- 七桥问题
- 最短路径
- 最大流
- 最大匹配
- 四色猜想
-

放心去飞

- ◆什么是图？

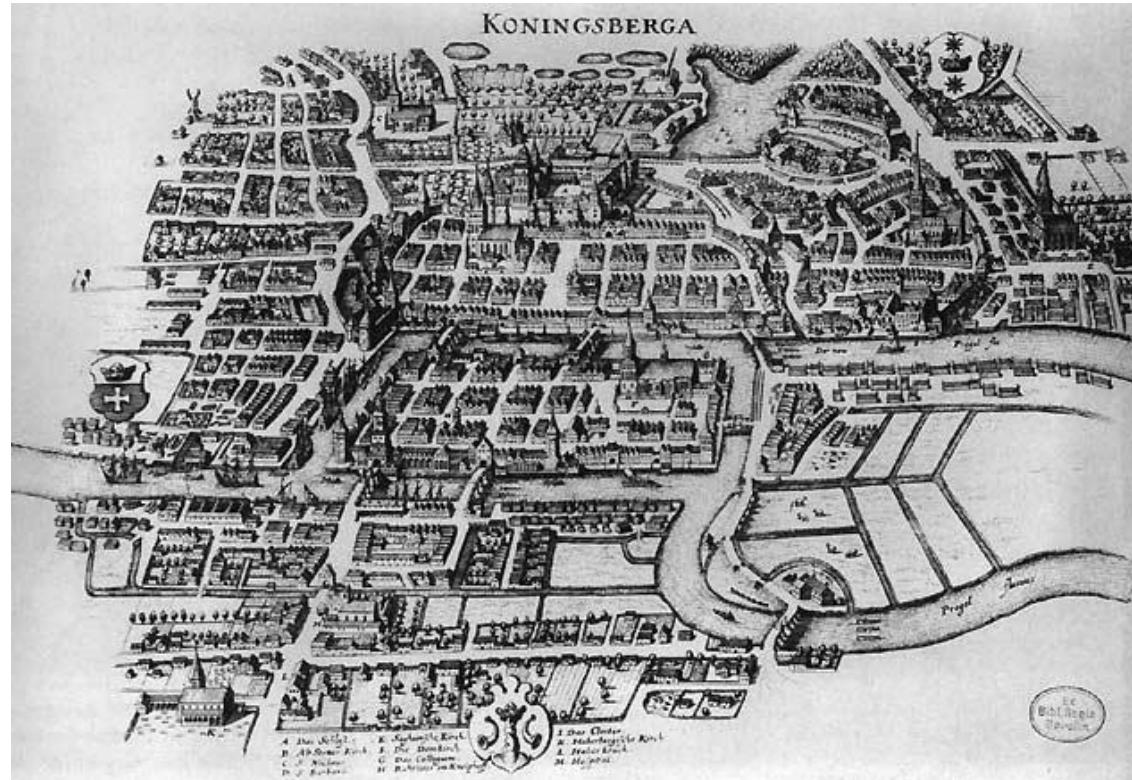
- 一笔画图
- 地图
- 交通流
- 计算机网络
- 集成电路设计
- 电厂仓库选址
- 生产物流系统
-

- 什么图论问题？

- 七桥问题
- 最短路径
- 最大流
- 最大匹配
- 四色猜想
-

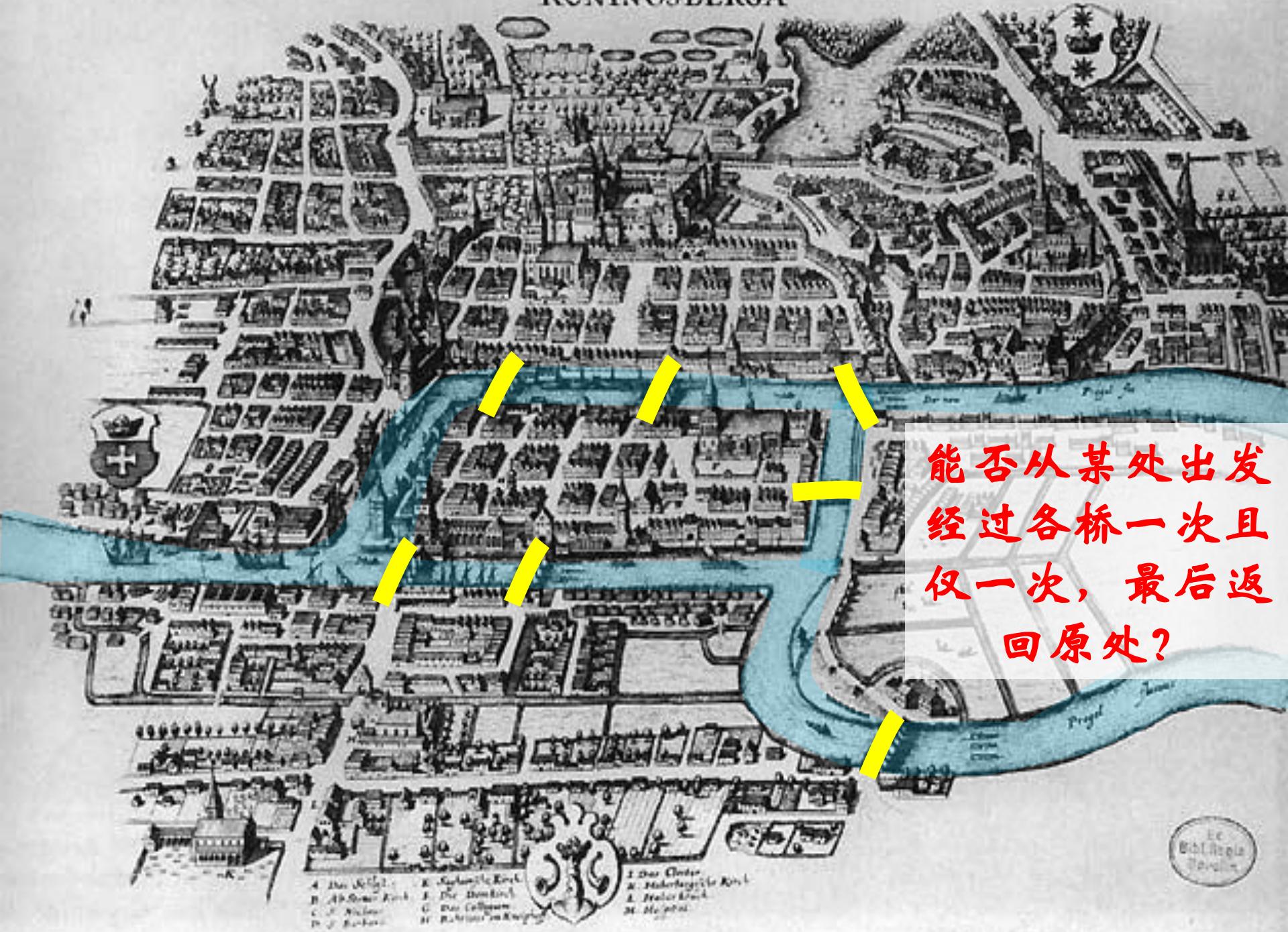
图论的起源

◆ 欧拉与哥尼斯堡城七桥问题（1736年）



注：哥尼斯堡城是今日的俄罗斯加里宁格勒

KONINGSBERGA

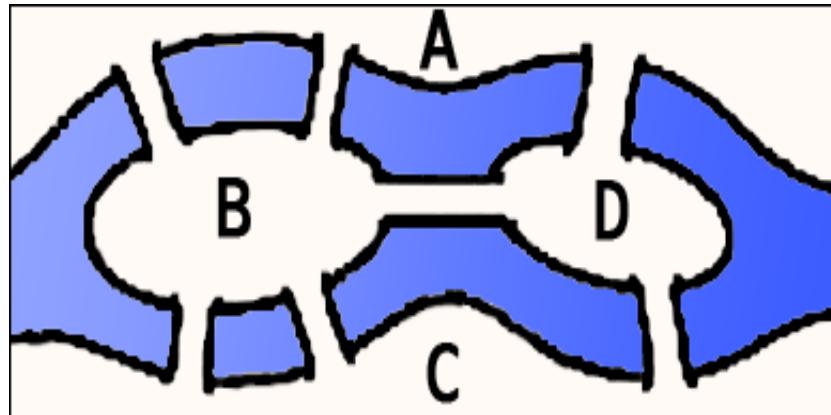


能否从某处出发
经过各桥一次且
仅一次，最后返
回原处？

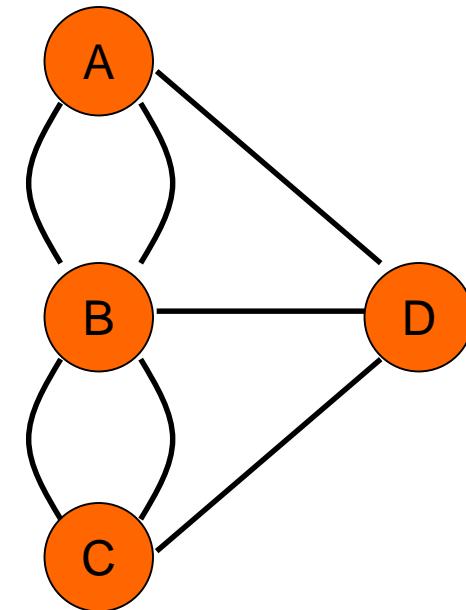
图论的起源

- 著名的哥尼斯堡7桥问题

- PregeI河流经哥尼斯堡，将该市分为两部分，河中还有两个小岛。18世纪初，该市有7座桥将它们相连，如左图所示。当时人们问：能否从某处出发，经过各桥一次且仅一次，最后返回原处？



图论建模



考虑桥的流量如何做?
部分区域是否存在桥不
确定?

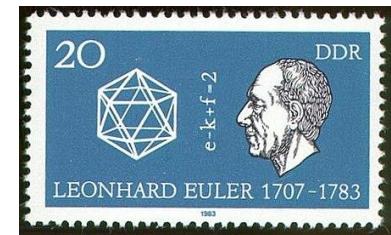
在生活中发现问题……

莱昂哈德·欧拉

- ◆ 引进了“函数”的概念，并且第一个将函数的写为 $f(x)$
- ◆ 以 e 表记自然对数的底（现在也称作欧拉数）
- ◆ 建立了流体力学里的欧拉方程
- ◆ 柯尼斯堡七桥问题是最早运用图论和拓扑学的典范
- ◆
- ◆ 人生最后7年，欧拉的双目完全失明，他还是以惊人的速度产出了生平一半的著作!!



前苏联于1957年发行的邮票，纪念欧拉诞辰250周年



前东德发行的欧拉逝世200周年纪念邮票

图论的历史与发展

◆ 欧拉与哥尼斯堡七桥问题

- 人们常称1736年是图论历史元年，因为在这一年伟大数学家欧拉（Euler）发表了图论的首篇论文——《哥尼斯堡七桥问题无解》——欧拉是图论的创始人

◆ 1936年《有限图与无限图理论》

- 匈牙利数学家寇尼格（Konig）出版图论的第一部专著
- 图论发展史上的重要里程碑，标志着图论进入突飞猛进发展阶段

◆ 近40年来，图论发展异军突起，活跃非凡

- 计算机科学的发展为图论的发展提供了计算工具
- 现代科学技术的发展需要借助图论来描述和解决课题中的各种关系

图论的重要性 (Why?)

◆ 图论在许多领域的应用广泛

- 描述事务之间关系的手段或称工具
- 广泛应用于电信网络、电力网络、运输能力、开关理论、编码理论、控制论、反馈理论、随机过程、可靠性理论、化学化合物的辨认、计算机的程序设计、故障诊断、人工智能、印制电路板的设计、图案识别、地图着色、情报检索.....
- 应用于诸如语言学、社会结构、经济学、运筹学、兵站学、遗传学等等方面。

图论的重要性 (Why?)

- ◆ 图论是计算机学科的重要基础课
 - 培养计算机科学的离散思维能力
 - 最短路径、树、染色、匹配、最大流……
 - 《程序设计》、《数据结构》、《计算机网络原理》……

社交网络

facebook

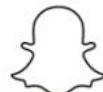
- 2 billion MAU
- 26.4 billion minutes/day

twitter

- 320 million MAU
- Peak: 143K tweets/s

Instagram

- 700 million MAU
- 95 million pics/day



snapchat

- 300 million MAU
- 30 minutes/user/day

Alibaba Group
阿里巴巴集团

- >777 million trans. (alipay)
- 200 billion on 11/11

新浪微博
weibo.com

- 462 million users
- influencing our daily life

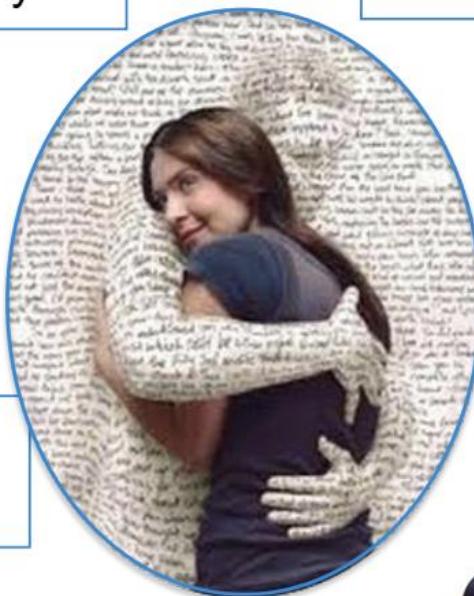
头条 今日头条

- ~1.5 billion MAU
- 70 minutes/user/day



WeChat

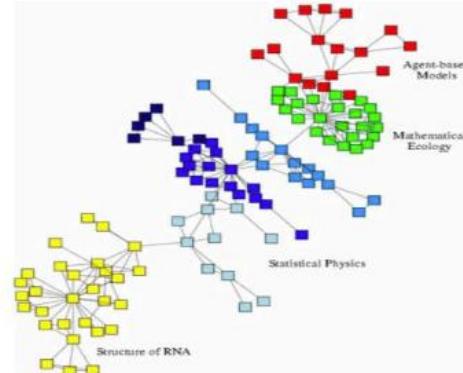
- QQ: 860 million MAU
- WeChat: 1.1 billion MAU



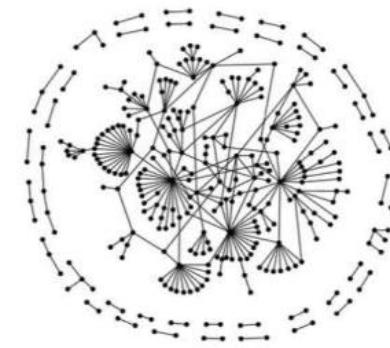
图是一种基本的数据组织形式



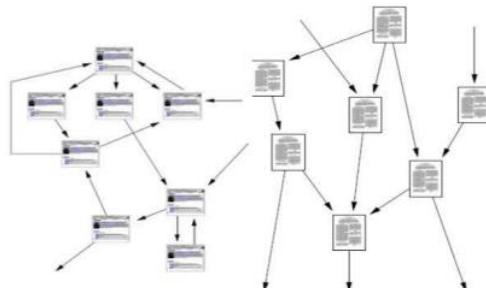
社交网络



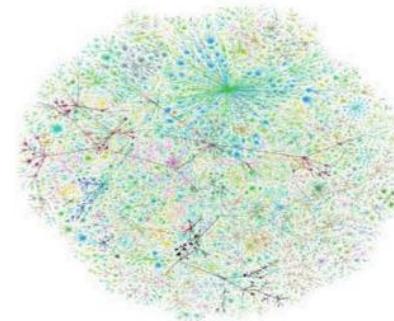
经济网络



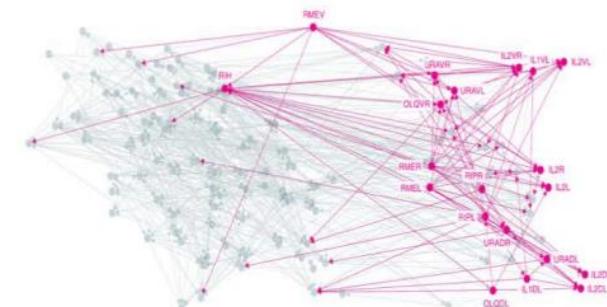
神经网络



学科网络



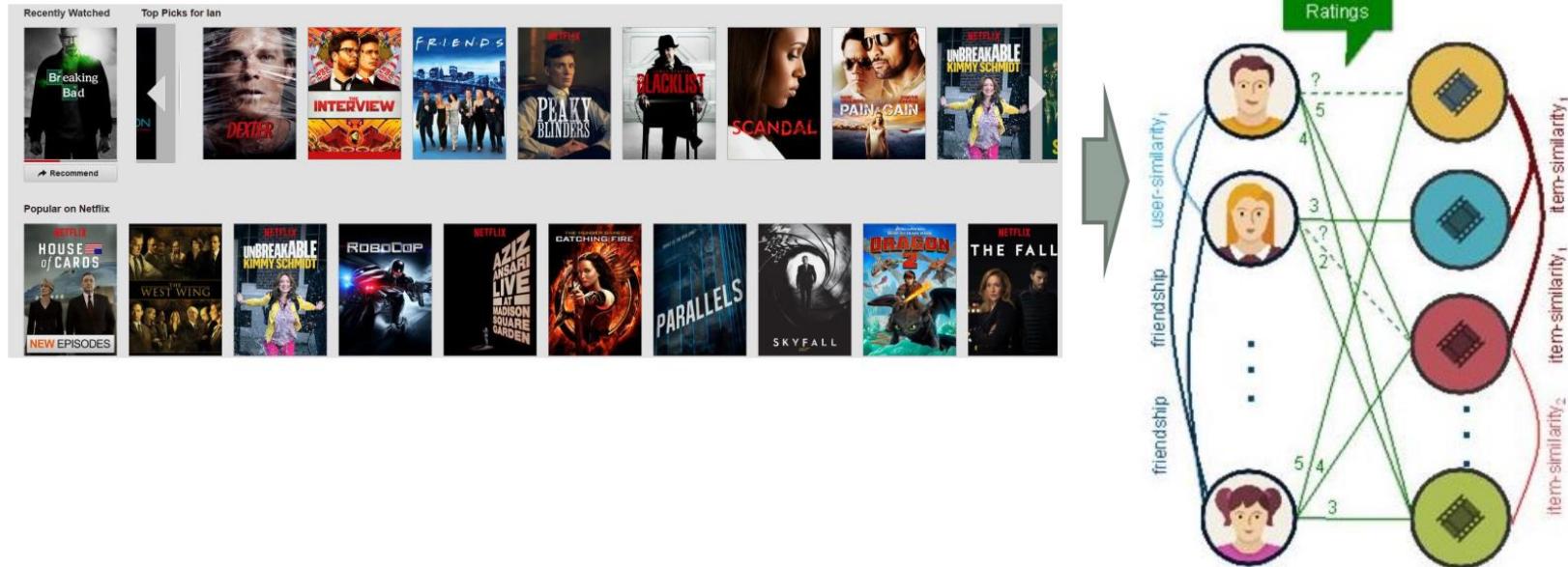
因特网



人工神经网路

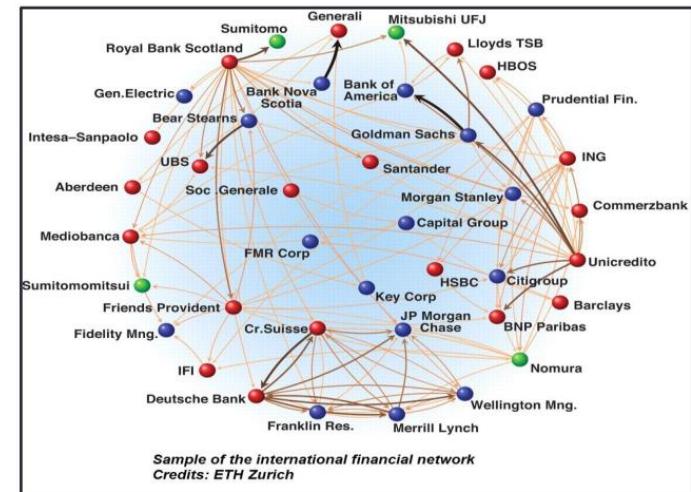
图论相关技术具有广泛应用

- ◆ 推荐系统可以转化为二分图上的链接预测问题



图论相关技术具有广泛应用

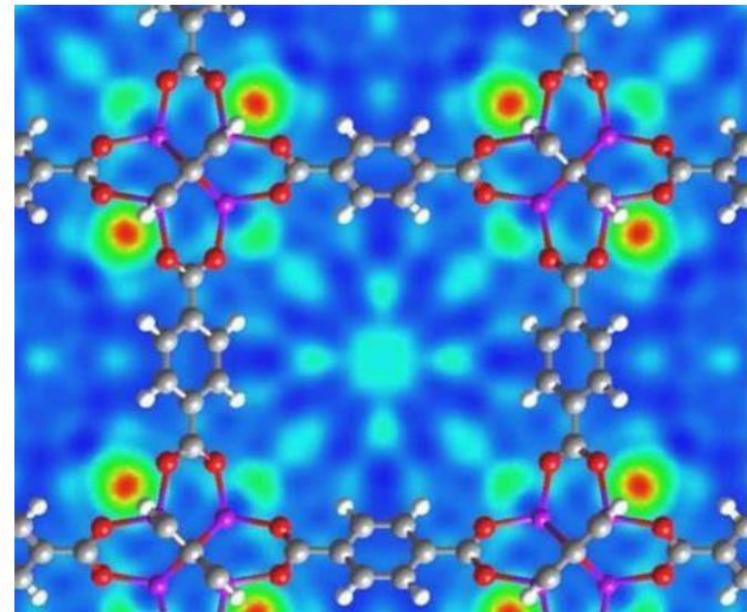
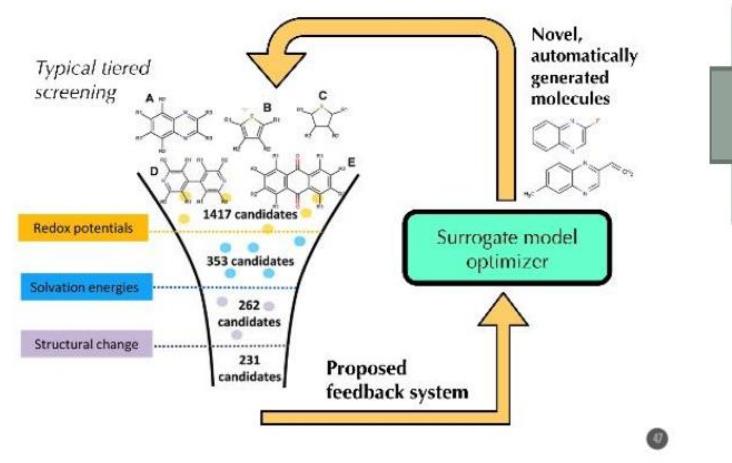
- ◆ 信用和风险评估可以转化为节点分类问题



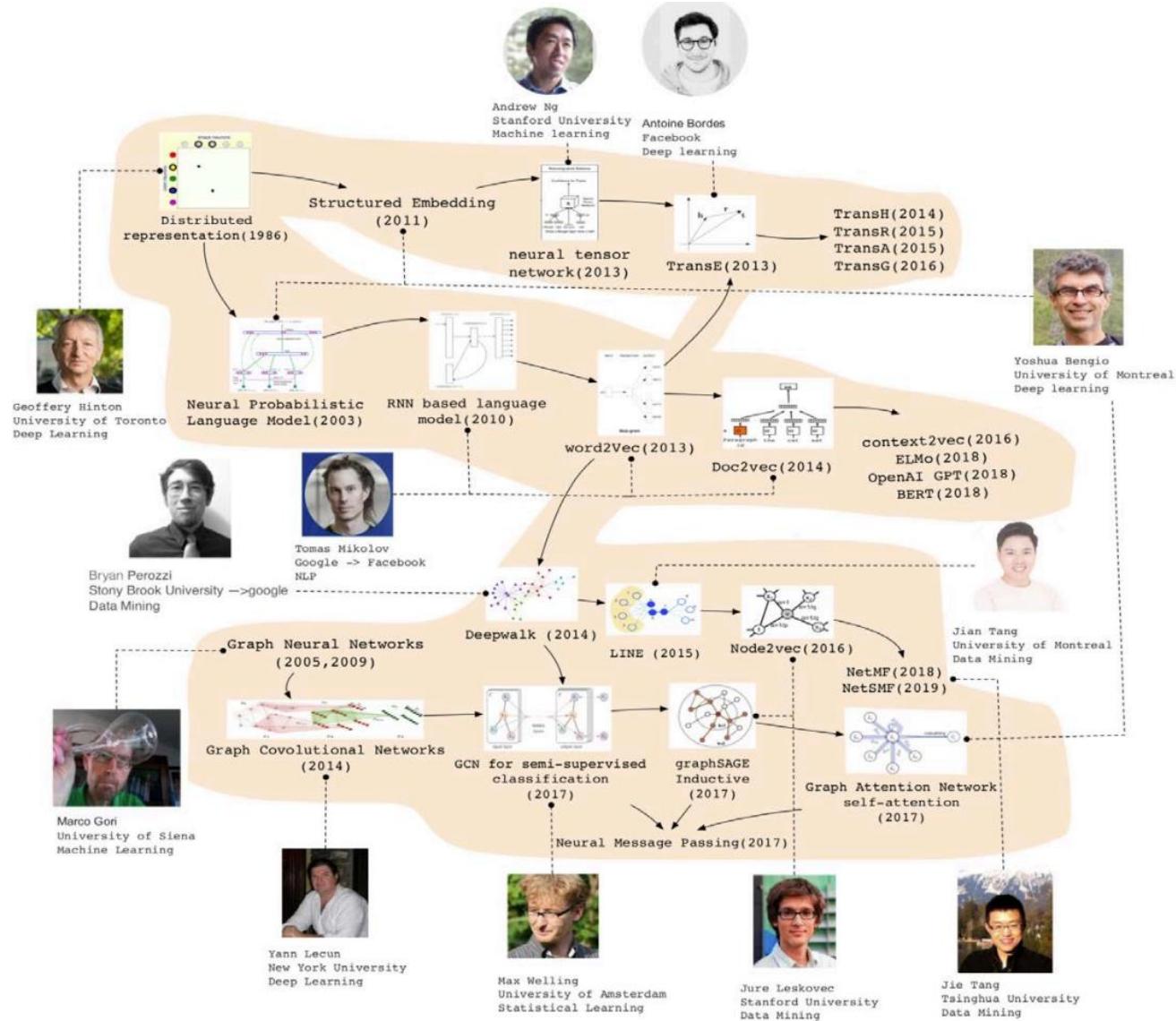
图论相关技术具有广泛应用

- ◆ 部分材料和药物发现问题可以转化为子图模式发现问题

Materials discovery engine concept



图上数据挖掘



如何学好图论及代数结构

◆ 离散数学

- 以离散量为研究对象，为计算机专业的基础数学课
- 主要内容：集合论、数理逻辑、**图论**

◆ 课程特点

- 与三角函数、微积分不同，**反证法、构造法**是解题主要方法
- 从基本概念入手充分掌握定义、定理，并重视算法
- 保质保量完成作业和报告

要求及学习方法

- ◆ 听课：基础内容，拓展知识面
- ◆ 参与：积极参与课堂讨论
- ◆ 作业：巩固基础知识
- ◆ 实验：编程&理解关键算法
- ◆ 思考：知其然知其所以然
- ◆ 实践：拓展应用本领（课外）
- ◆ 创新：如果我是欧拉
- ◆ 前沿：学术前沿研究成果

减轻工作量，不用预习课本；但课后要认真看书

子曰：

「学而不思则罔，
思而不学则殆。」

我们能学习到什么？

- ◆ 什么是最重要的？

- 定义概念？公式？定理定律？证明过程？

- ◆ 本课程有助于训练：

- 基本知识—图论所涉及的概念、定理……

- 逻辑思维—数学课

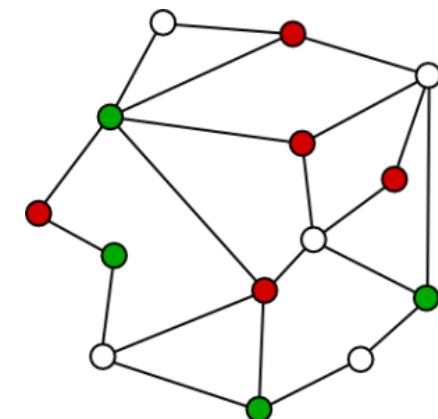
- 思维模式—从问题和为什么入手的分析方法

- 创新能力—做实验编程序、自己设计定义定理

- 语言表达—课堂演讲

图的基本概念

- ◆ 世界上各事物之间，自然界内诸现象之间经常存在着某些必然的联系，需要人们通过研究分析，去揭示这些关系
- ◆ 人们常把**事物、现象用顶点（或结点）表示**，用有向的或无向的边来表示它们之间的联系
- ◆ 这就构成了图论中所讨论的图



图论的基本问题

◆ 存在性问题

- 是否存在可行解

◆ 计数问题

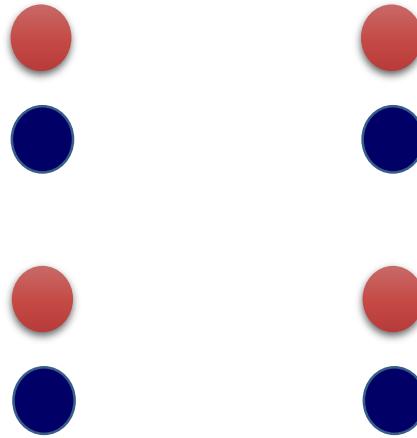
- 有多少个可行解

◆ 优化问题

- 寻找某种意义上的“最优”解

存在性问题

- ◆ 四对夫妻在两个球场上打混双比赛
- ◆ 时间两个对手，每半个小时交换搭档和对手
- ◆ 是否存在这样的安排
 - 每个男人和每个女人分别做一次搭档和对手
 - 每个男人和其它男人至少做一次对手



计数问题

- ◆ 六个人投资俱乐部规定每年轮换主席和财务主管
- ◆ 最多经过多少年，他们将不得不使同一个人重复出现在同样的位置上



优化问题

- ◆ 老板有三个雇员Bill、Quentin和Robin，他们的工资分别是\$6, \$7和\$8。目前有三项工作要进行分配，每个人的工作时间如下，请问应该如何分配才能使得支付的总费用最少？

| | Bill | Quentin | Robin |
|--------|------|---------|-------|
| Task 1 | 7.5 | 6 | 6.5 |
| Task 2 | 8 | 8.5 | 7 |
| Task 3 | 5 | 6.5 | 5.5 |

图论典型问题

- ◆ 工程完成时间问题

- ◆ 匹配问题

- ◆ 背包问题

工程完成时间问题

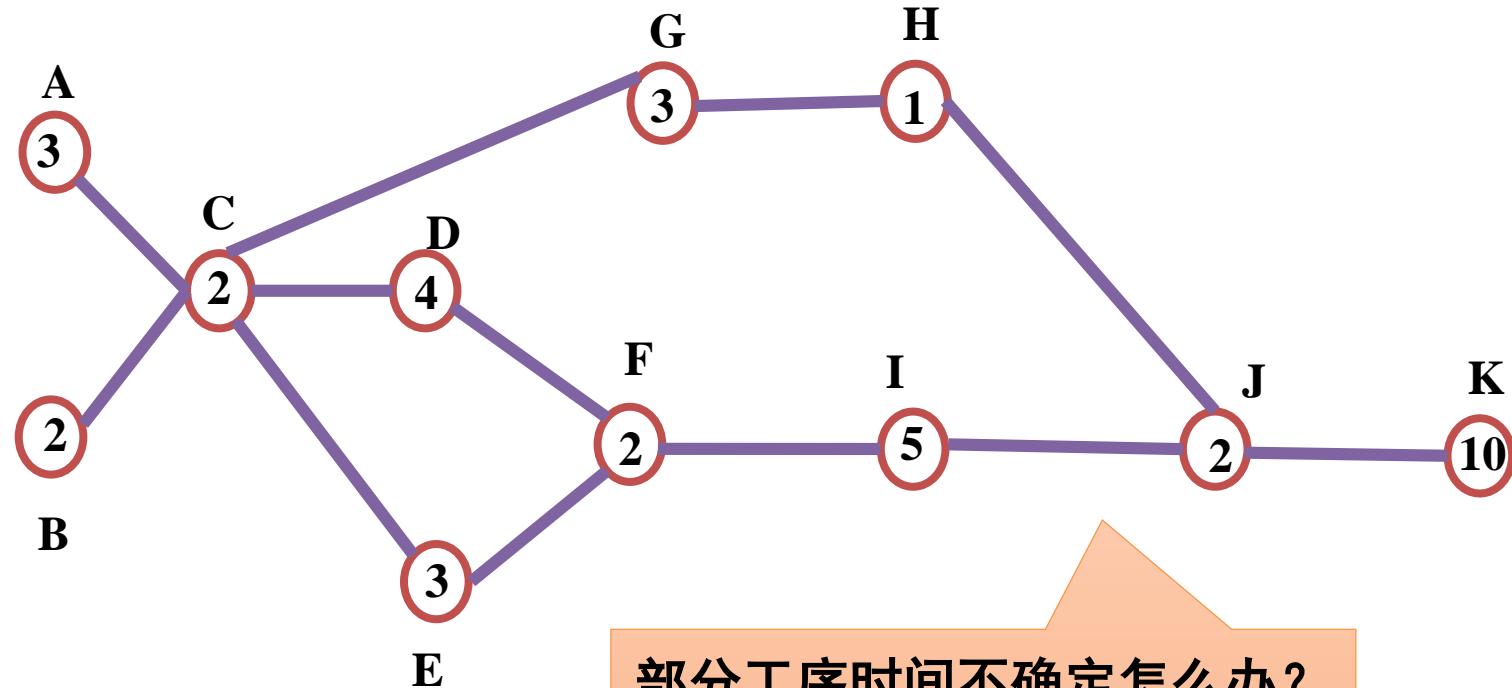
◆ 市场营销方案制定

| 工作 | 前继工作 | 天数 |
|-----------|------|----|
| A 经理选择商品 | 无 | 3 |
| B 采购员确定商品 | 无 | 2 |
| C 选择广告商品 | A, B | 2 |
| D 准备图片 | C | 4 |
| E 准备文字 | C | 3 |
| F 设计广告 | D, E | 2 |

| 工作 | 前继工作 | 天数 |
|--------|------|----|
| G 汇集清单 | C | 3 |
| H 打印标签 | G | 1 |
| I 打印广告 | F | 5 |
| J 贴标签 | H, I | 2 |
| K 发送广告 | J | 10 |
| | | |

工程完成时间问题

◆ 图表示和关键路径分析



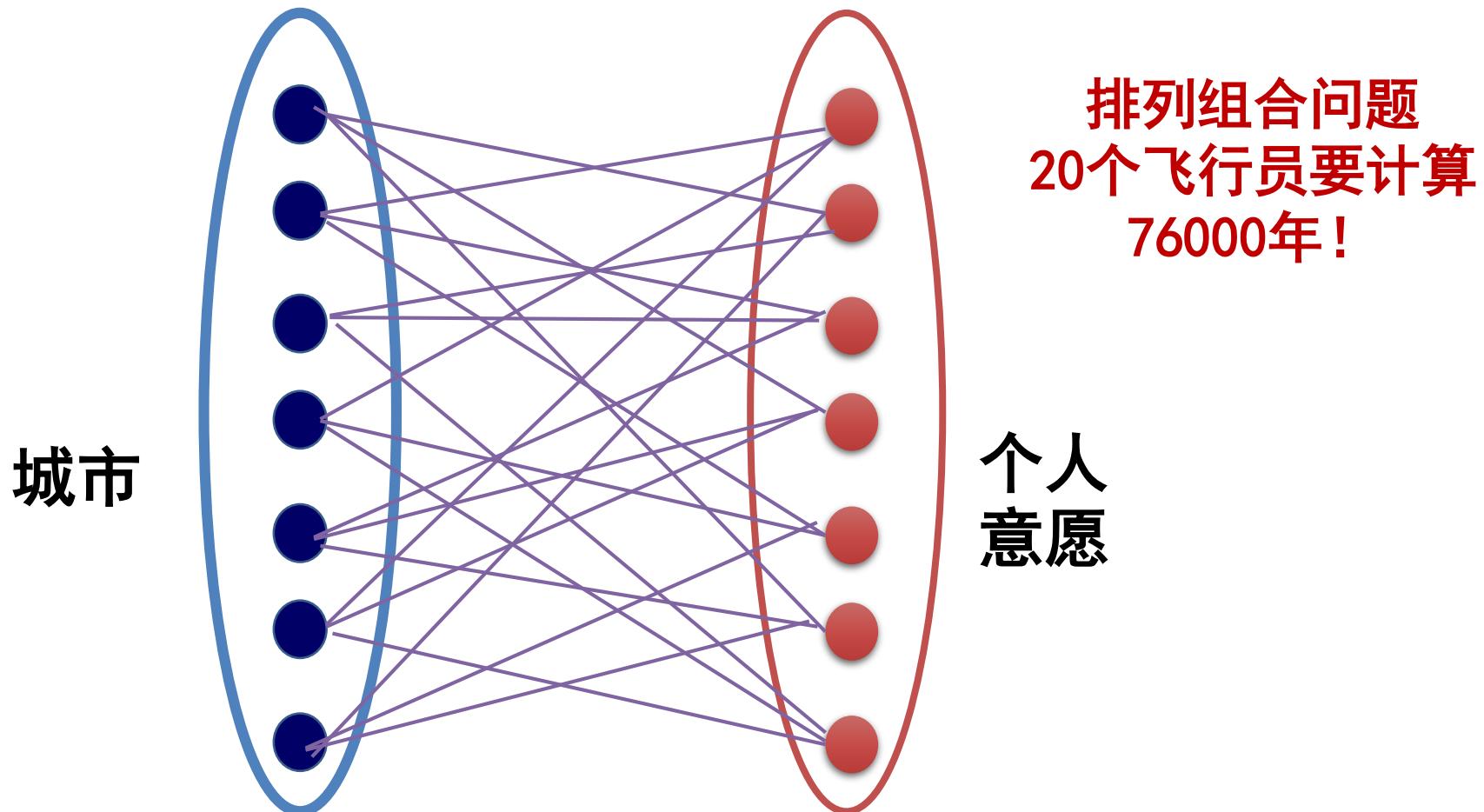
部分工序时间不确定怎么办?
是否考虑风险等其它维度?
大规模问题如何处理?

关键路径分析

- ◆ 计划评估技术（PERT, Program Evaluation and Review Technique）
- ◆ 诞生于美国海军北极星潜艇和导弹项目中
- ◆ 涉及数百个步骤的大型项目编排进程：复杂系统！
- ◆ 减少项目所需要的时间，需要压缩关键路径！

匹配问题

- ◆ 航空公司飞行员匹配问题，尽可能让所有的飞行员满意



背包问题

- ◆ 有一个徒步旅行者，其可携带物品重量的限度为 a 公斤，设有 n 种物品可供他选择装入包中。已知每种物品的重量及使用价值（作用），问此人应如何选择携带的物品（各几件），使所起作用（使用价值）最大？

| 物品 | 1 | 2 | ... | j | ... | n |
|----------|-------|-------|-----|-------|-----|-------|
| 重量（公斤/件） | a_1 | a_2 | ... | a_j | ... | a_n |
| 每件使用价值 | c_1 | c_2 | ... | c_j | ... | c_n |

以往课程学习成果展示

◆ 实际应用问题

- 公交网络搭建，送餐路线制订
- 清华交通网络优化

仅用于
开阔思路

◆ 专业课题

- 河流水利设施规划，DNA判定
- 智能組句拼音輸入法，互联网组播传输模型

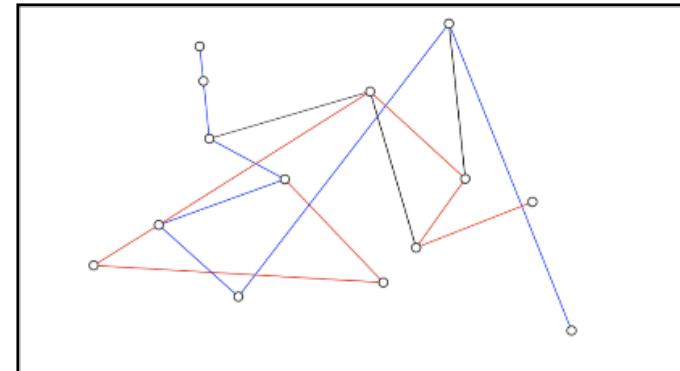
◆ 你的最爱⑧

- 连连看，五子棋，数独，游园路线，各类游戏

公交网络搭建

◆ 问题描述

- 给定车站和可能的路线，搭建公交网络，要求：
 - 覆盖所有车站
 - 减少乘客的等待时间
 - 减少需要的车辆数量



◆ 问题求解

- 公交路线选取
 - 最小生成树
- 公交线路划分
 - 道路和回路

• 无限遐想

- 乘客数量不等
- 出行时间各异
- 等候时间还是车上时间
- 换乘次数最少
- 如何设定车站地址
- 最少的车辆数量
- 乘客呈现概率分布
-

送餐路线制订

◆ 问题描述：等待时间最短？

- 有k个送餐员，n个订单
- 最小化总路径长度
- 分配订单并为每个送餐员制定送餐路线

◆ 问题求解

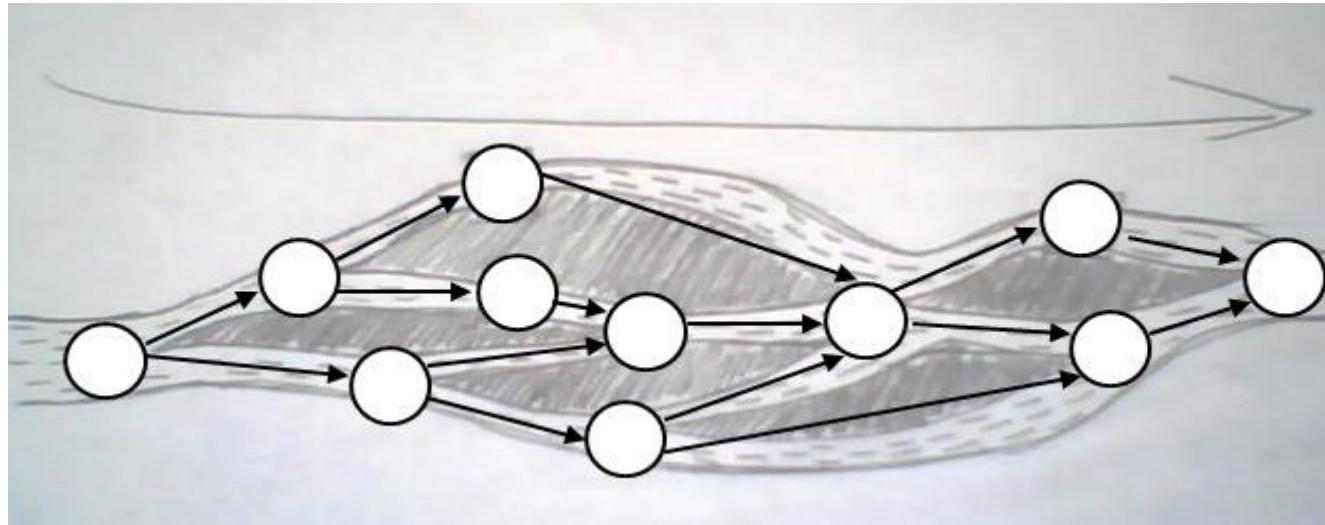
- 寻找不同订单位置之间的路径
 - 两点间的最短路径
- 将n个订单分成k组
 - 最小生成树
- 为每组订单求制订路线
 - 旅行商问题



- 无限遐想
 - 半小时内必达
 - 离线问题还是在线问题？
 - 最少多少个送餐员？
 - 最多能接多少订单？
 -

河流水利设施规划

- 问题描述
 - 在河道沿岸选取合适的位置建设水利设施进行扩容
 - 保证河道能够承受流量
 - 尽量减少建设开销
- 问题求解
 - 河道流量分析
 - 最大流
 - 选取河道进行扩容
 - 图的割切



DNA判定

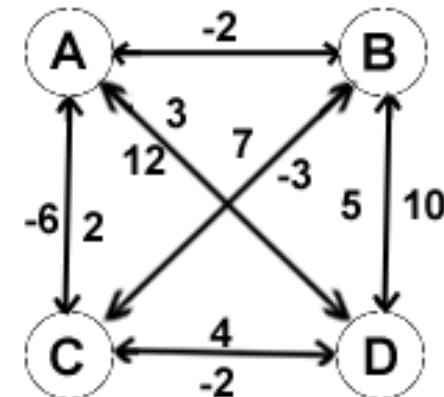
- 问题描述

- 判定一段生物体DNA序列是否包含一些特定性的序列
- 特征序列本身多变，而且可能会有自身的重复

CCTCCGCTCTCCCTCTTCCCCCTTCCCCGTGGTGTGG
CCGCCGCCTCATCCTCCATCCCTCTTCCTCCGCCGCGTCG
TTGCCATTATCTAACTATTCACCCCTTCCATCCACGGCACG
TTGCCATTATCTAACTATTCACCCCTTCCATCCACGGCACG
CCTCTTCTCCCTTTCTTAGCTTCCTCCTACCGATG
CCTTCTCCTTCTTCTTCCTCATTATTGCGTGTG
TCTTCTTCTTCTTCTTCTACCTCTCGATCGATCTCCAG
CCTCCGCCGCCGCCGTTGCCCTCCATGCTGCCCTC
TCTCTTCTCCCTCCCTACCATGGCTGCTTTCTATCATG
TCTCCAGAAAATTAACCTTCCCCCCCCCTTCTTCATG

- 问题求解

- 图描述特征序列中不同片段连接的倾向性
 - 图论建模
- 利用特征图对DNA序列进行判定
 - 算法设计，迭代



连连看

◆ 连连看游戏AI

- 获取当前状态
 - 截图
 - 图像识别
- 计算可消去的方块
 - 寻找相同方块
 - 搜索消去道路
- 执行消去操作
 - 模拟鼠标事件



五子棋人机对战游戏

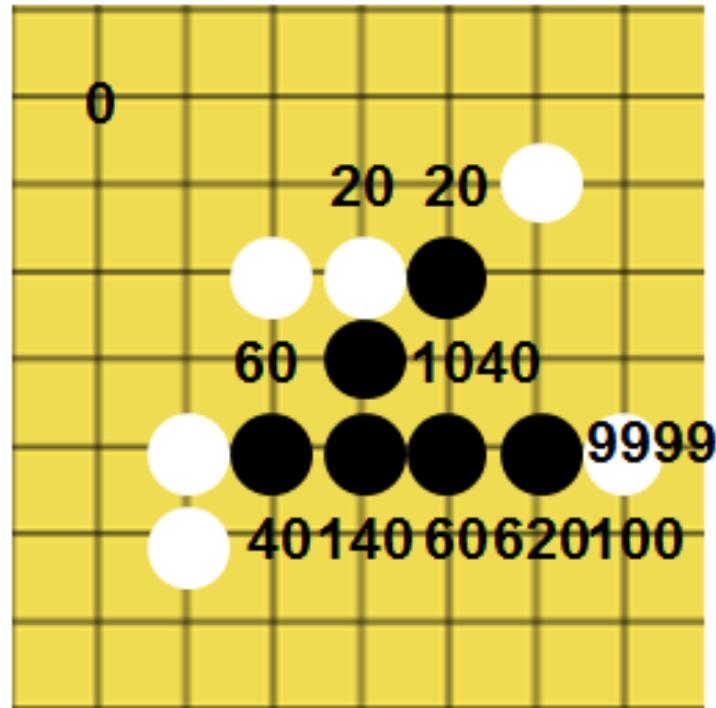
◆ 五子棋AI

□ 权值分析法

- 对每一个位置进行评分，得出分数最高的点即为下一步的位置

□ 多步分析

- 假设每个位置是否落子，然后向下继续假设多步，建立搜索树
- 分析每种情况输赢进而得出最初一步的最佳位置



数独游戏

◆ 数独AI

□ 行列排除法

- 根据行与列中数字的分布，选择具有唯一可能的方格

□ 穷举法

□ 假设法

- 以可能解较少的方格为出发点进行穷举搜索
- 搜索中需要进行回溯

| | | | | | | | | |
|---------------|-------------|-----------------------|-----------------|---------------|---------------|-----------------|---------------|---------------|
| 9 | 2,4,5, 6 | 2,3,4, 5,6,7, 8 | 4,5,8 | 3,4,5 | 4,8 | 2,3,6, 7,8 | 1 | 2,3,6, 7,8 |
| 1,3,8 | 2,4 | 2,3,4, 7,8 | 4,8 | 6 | 1,4,8 | 9 | 3,7 | 5 |
| 1,3,5, 6,8 | 5,6 | 3,5,6, 8 | 2 | 1,3,5, 9 | 7 | 3,6,8 | 4 | 3,6,8 |
| 7 | 3 | 5,6,8, 9 | 1 | 5,9 | 6,8,9 | 5,6 | 2 | 4 |
| 5,6,8 | 4,5,6, 9 | 4,5,6, 9,8 | 4,5,6, 7,8,9 | 2,4,5, 7,9 | 2,4,6, 8,9 | 1,3,5, 6,7 | 3,5,6, 7 | 3,6,7 |
| 2 | 1 | 4,5,6 | 4,5,6, 7 | 4,5,7 | 3 | 5,6,7 | 8 | 9 |
| 6 | 8 | 2,6,9 | 3 | 1,2,4, 7,9 | 5 | 2,4,6, 7 | 6,7,9 | 2,6,7 |
| 4 | 2,5,6, 9 | 1 | 6,7,9 | 8 | 2,6,9 | 2,3,5, 6,7 | 3,5,6, 7,9 | 2,3,6, 7 |
| 3,6,5 | 7 | 2,3,5, 6,9 | 4,6,9 | 2,4,9 | 2,4,6, 9 | 2,3,4, 5,6,8 | 3,5,6, 9 | 1 |

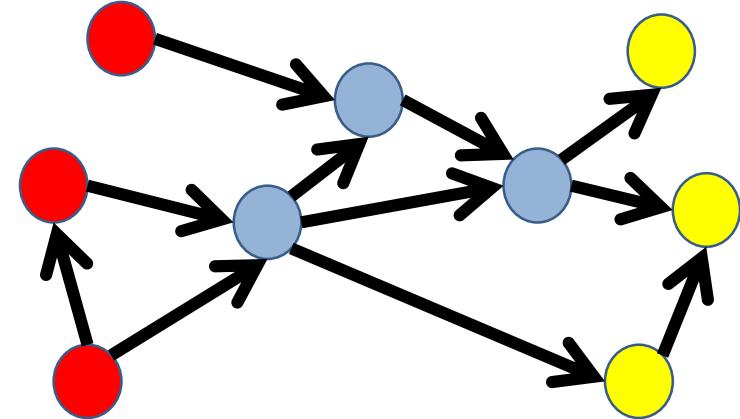
交通网络优化



交通网络优化

◆ 问题建模

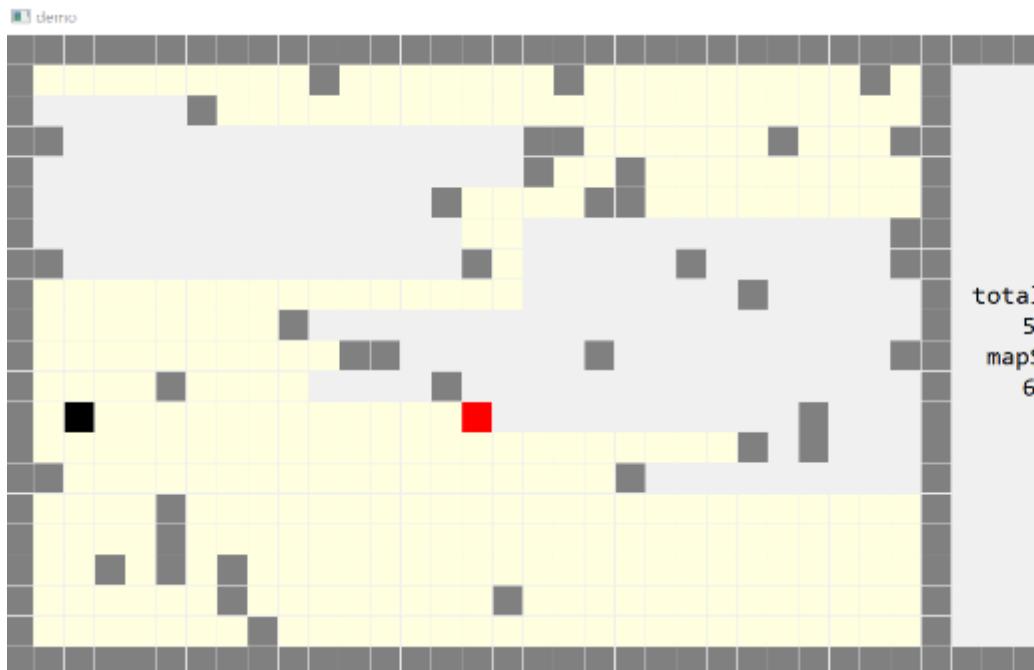
- 将所有教室、食堂、十字路口看做结点
- 将连接两地的道路看做两点间的边
- 同学从一些结点出发---源
- 到达一些结点---汇
- (通过一条道路所需的时间与这条路上的人数有关)
- 目标：
 - 使人们以最短的时间到达目的地
 - 最大限度地利用道路容量



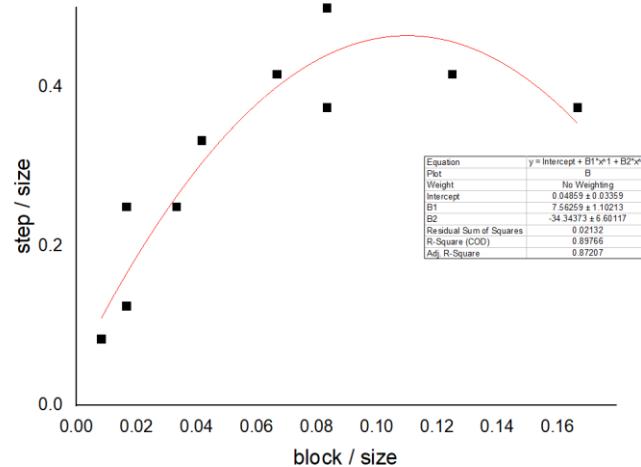
转化网络流图问题

扫地机器人的全覆盖路径规划

- ◆ 针对整体不变，局部有扰动的半在线算法
- ◆ 图论建模：旅行商问题+贪心
- ◆ 强化学习

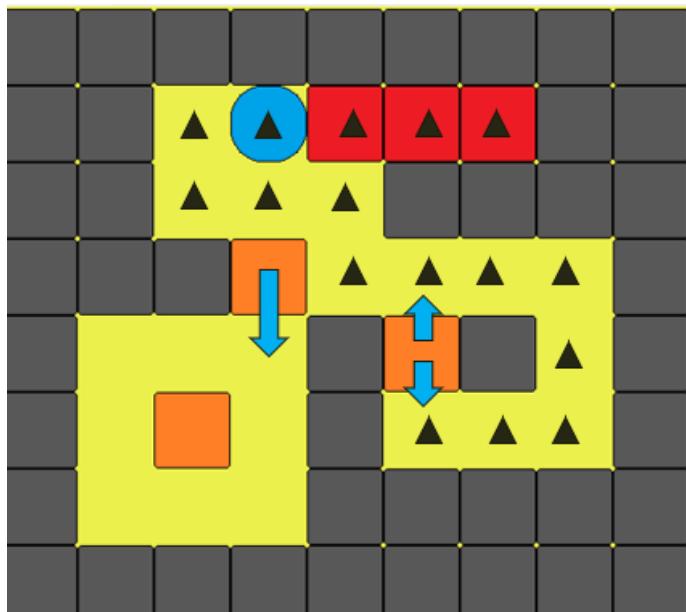


totalStep:
505
mapSize:
600

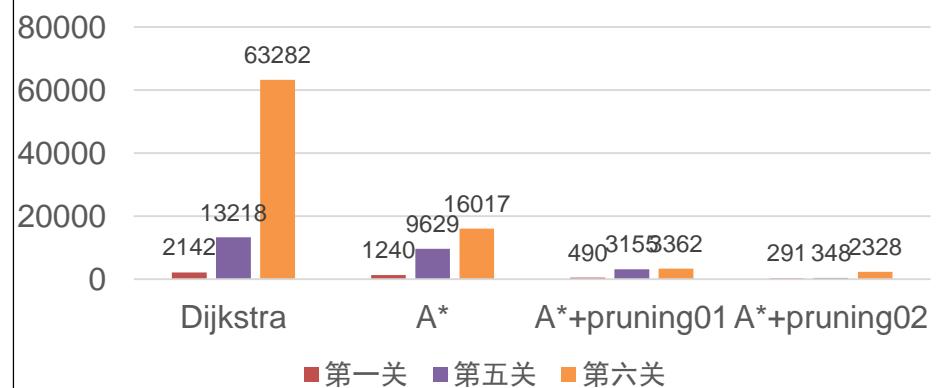


自动推箱子求解器

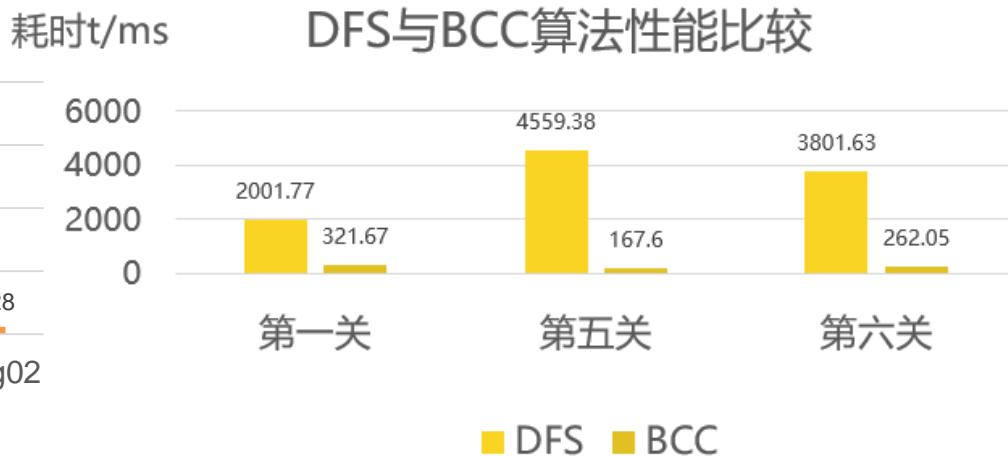
- 基于状态节点的图搜索
- DFS、BFS、Dijkstra、A*、剪枝
- 双连通分量(BCC)优化算法



几种算法展开节点数比较

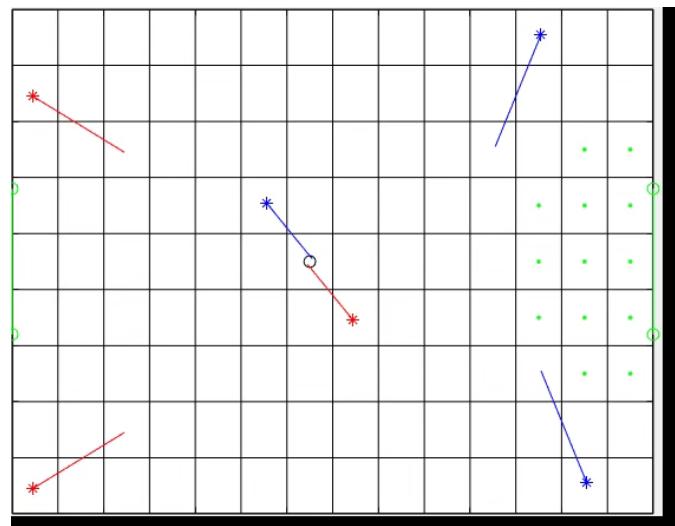


DFS与BCC算法性能比较



足球机器人路径规划算法

- ◆ 针对人形足球机器人比赛RoboCup
 - 将足球场简化为“棋盘”
 - 将运动员和球简化为“棋子”
 - 将比赛过程类比于下棋
- ◆ 用A*算法为进攻球员规划路径
 - 以进球为目标
 - 以球到对方球门最短距离为评估
 - 以Dijkstra计算该最短距离
 - 多方博弈问题？
- ◆ (终于有不那么原始的足球机器人比赛算法了



智能拼音输入法

◆ 问题提出

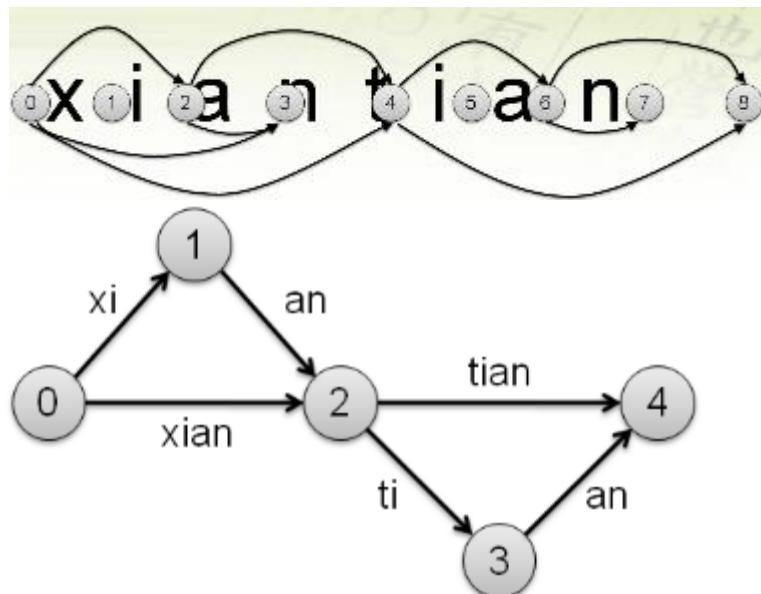
- 智能拼音输入应对拼音输入法重码过多
 - 音节解析问题
 - fangan --- “反感”？“方案”？
 - fanan --- “发难”？“翻案”？
 - xingai ---难道仅仅是“信概”？
 - 出现的概率
 - 统计语言模型：哪个词出现概率大？连接其他词呢？
 - 设计fangan？

与图有什么关系？

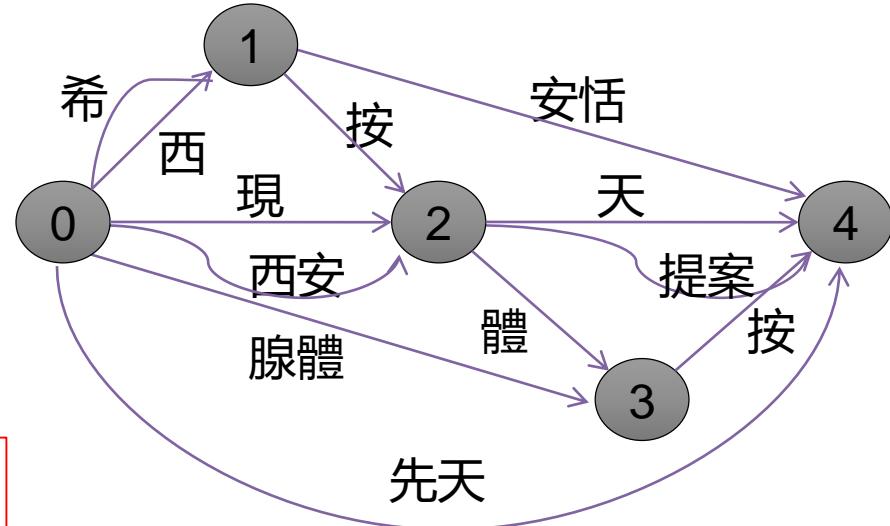
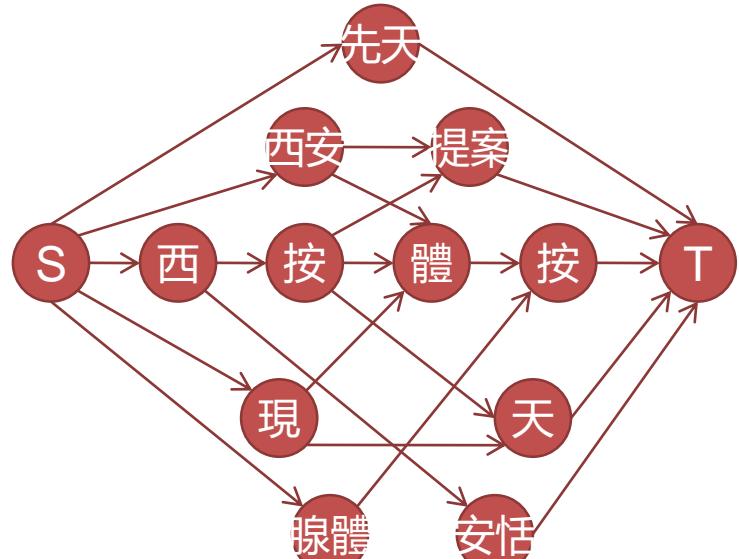
智能拼音输入法

建立音节图

- 将可能的字符连起来
- 转化成图论模型



xiantian, xiantianqi



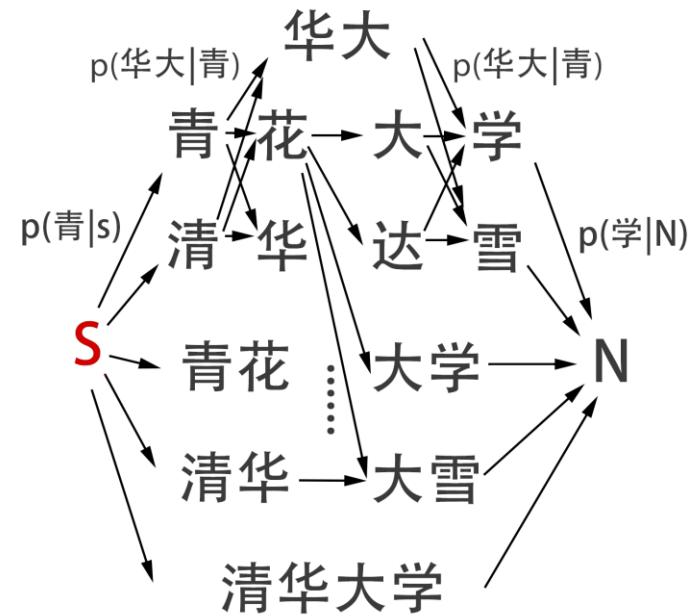
拼音输入法——原理

◆ 图论建模——词图

- 结点：字/词/句首尾标记
- 边：前n-1个字/词到下一个的概率
 - n元关系：实验确定n
 - 获得：统计语料的词频
- 目标：源到汇的最长道路
 - 此时这句话出现概率最大

◆ 问题求解

- 动态规划（Viterbi算法）
 - 从前向后给每个结点选择父结点
 - 使其到源道路最长



游园路线规划

◆ 从实际生活中找问题

- 游乐园（Disneyland）游玩时，如何选择合适路径及游玩方案，使得游玩方案相对更优？

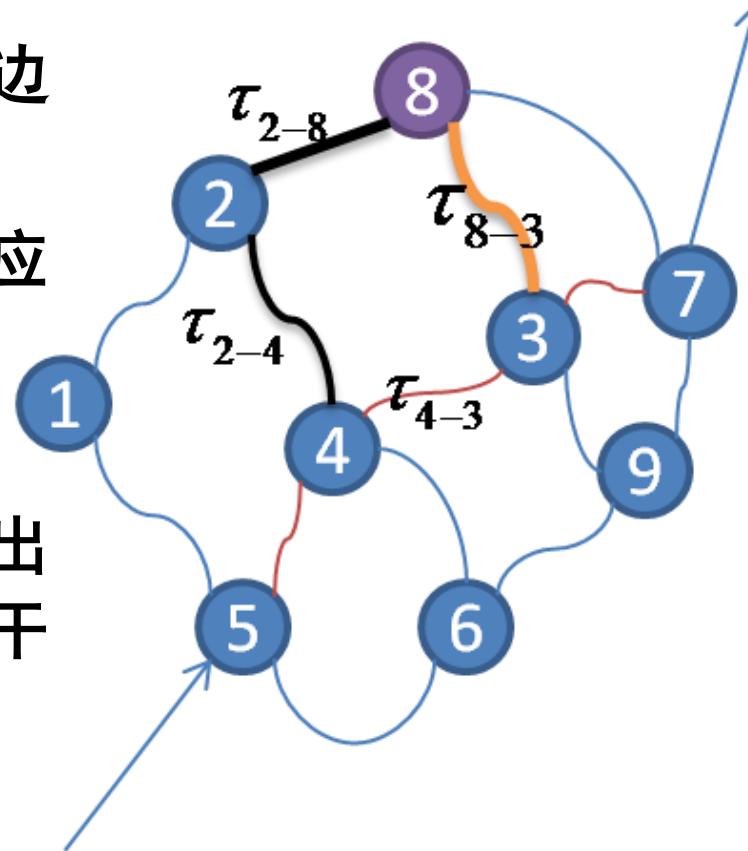
◆ 抽取关键因素

- 游玩地点及路径
- 时间分配
- 满足游客需求（想不想去）
-



游园路线规划建模

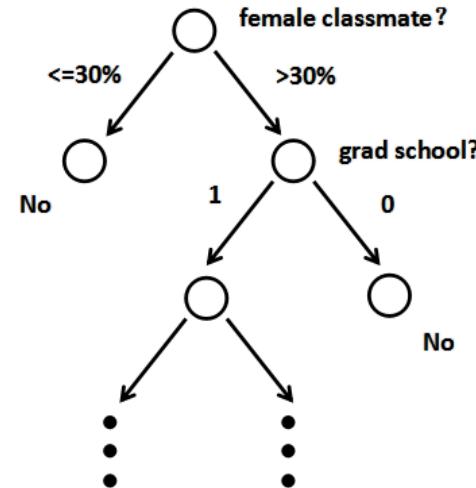
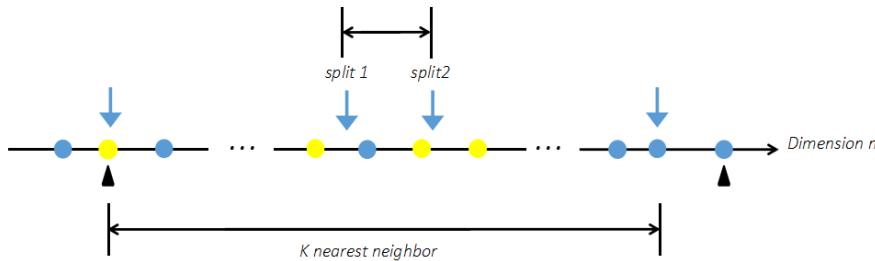
- 将游乐设施等作结点，道路为边
- 给每个结点一个权值，对应“重要程度”，此外结点还对应有花费的时间，相当于长度
- 每条边有一个时间值
- 最终目的是给出一条从入口到出口的道路及道路上要去玩的若干结点



最终目标：在总耗费时间不超过给定时间的情况下，使得总权值最大。

多分支决策树算法设计

什么是决策树?
二分支vs多分支?
分类连续特征的数据?
可不可以提高效率?



算法设计：

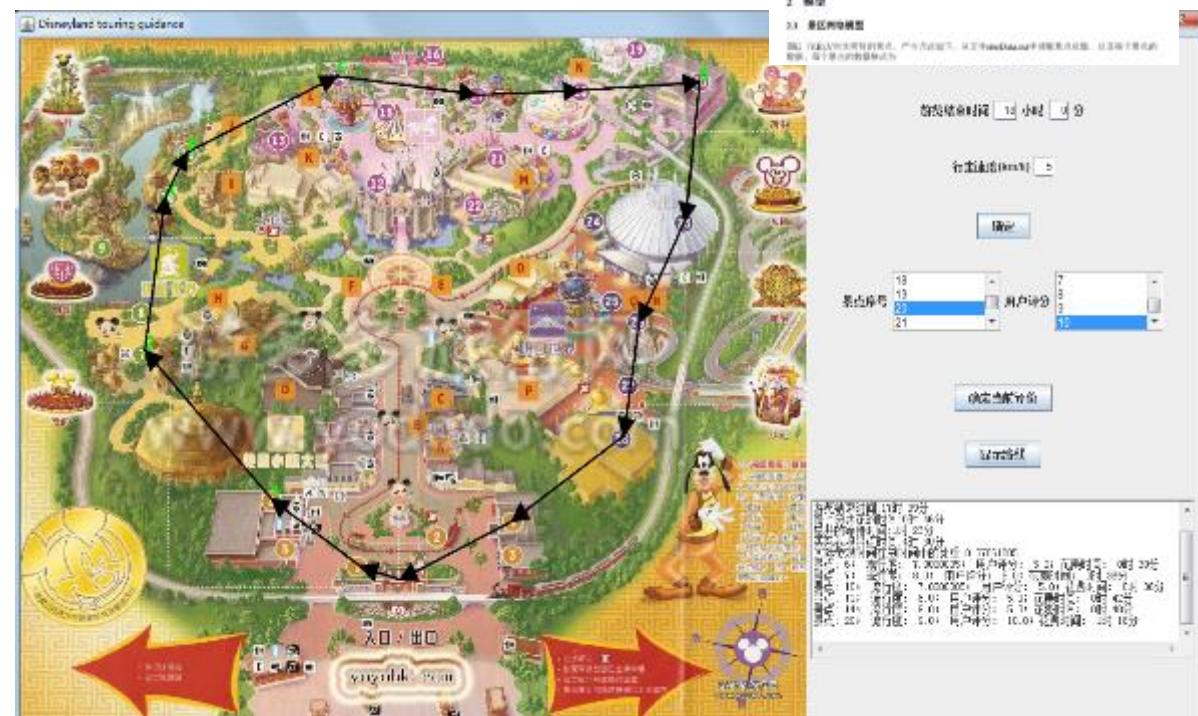
通过计算一个维度上离一个点最近的K个点的邻域的数据组成，决定决策树构建方式，通过迭代更新邻域扫描整个轴

Disneyland 旅游解决方案

Graph Theory
design for touring routing

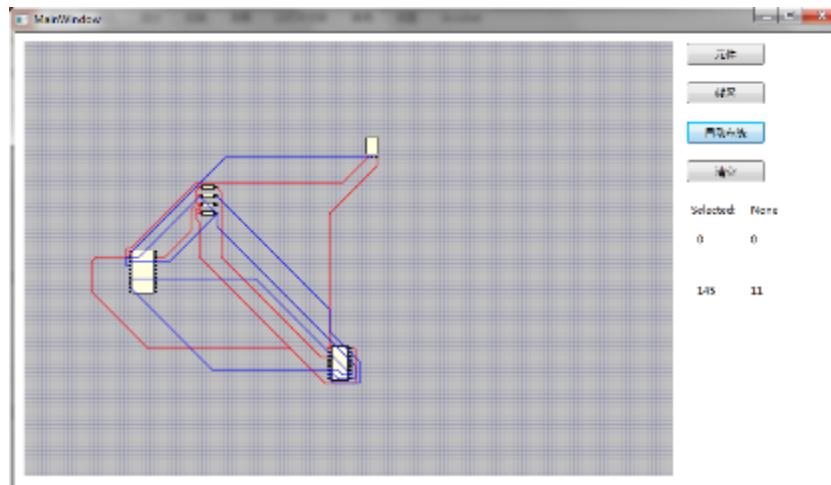
Department of engineering physics
Shantou University, China
www.sustech.edu.cn

- ◆ NP问题简化：景点、路线，分离→旅行商问题
- ◆ 模拟退火算法，降低复杂度
- ◆ A*算法，加速搜索
- ◆ 工作量
 - 自学Java
 - 近1500行代码
 - 说明文档6页
- ◆ 个人信息
 - 王寿文
 - 工程物理系

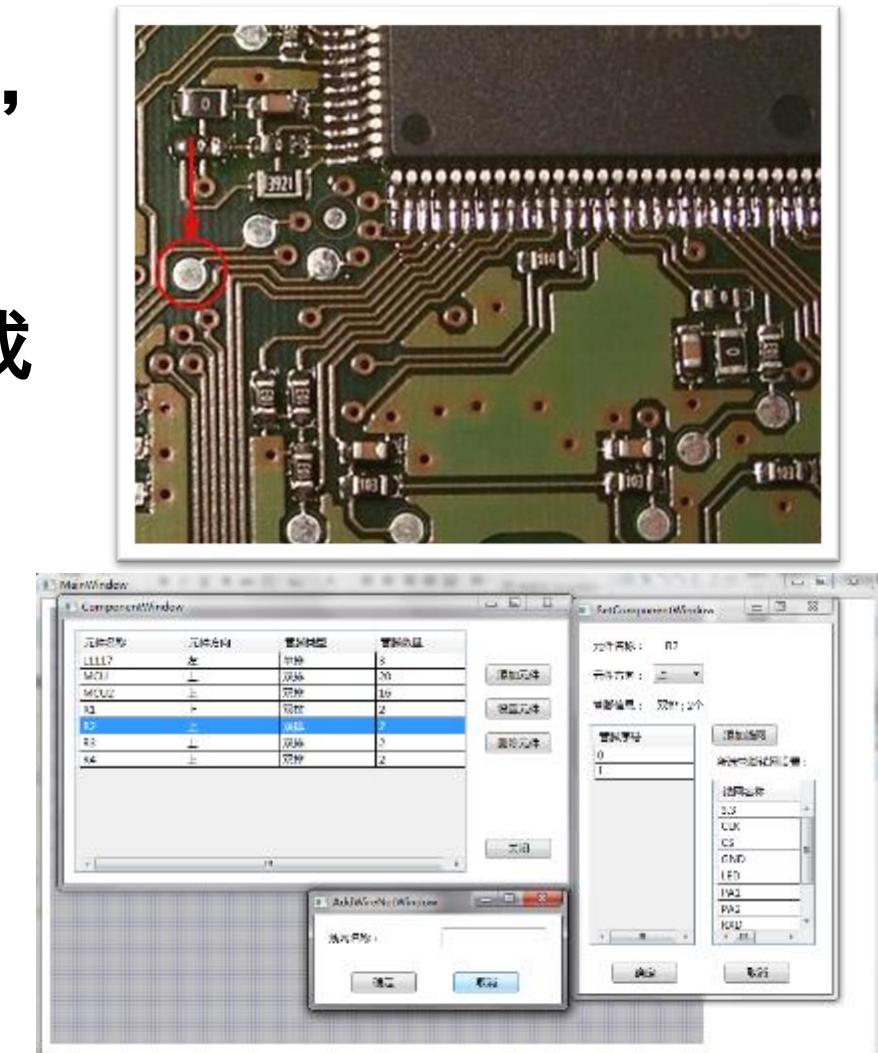


PCB自动布线程序

- 通过点击、拖拽设置元件，直观的操作体验
- 数据库存储、检索数据
- Kruscal 算法实现线路生成
- A*算法探路
- 精02 匡冶 兰天



自动布线效果图



元件管理界面

电力系统建模



- **发电中心**
- **中转中心**
- **负荷中心**
- **输电线路** → 边

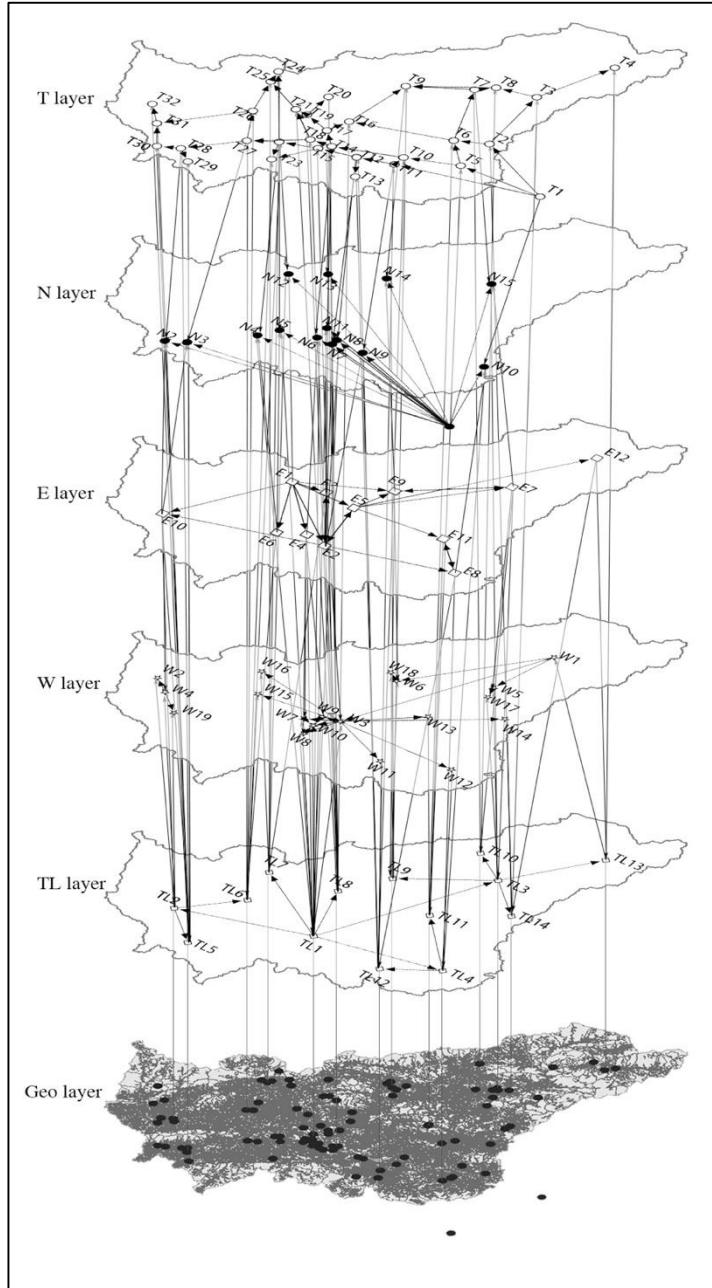
城市韧性问题研究

◆ 问题描述

- 计算灾害破坏后城市基础设施韧性水平（城市及基础设施在冲击前吸收抵抗、适应恢复的能力）
- 考虑节点失效概率影响

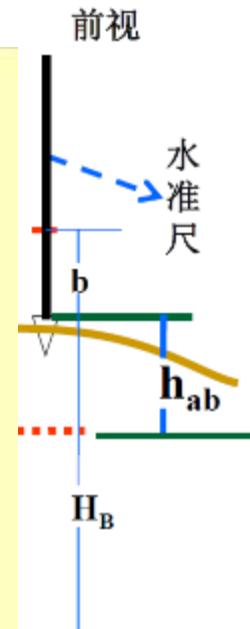
◆ 问题求解

- 图描述基础设施关联关系
- 图论建模
- 利用最短关键路径对基础设施韧性进行评估
- Floyd算法设计，迭代



水准网闭合差计算

水准网



出租车调度系统

- ◆ 十字路口和道路通行时间
- ◆ 汽车
 - 位置、容量
- ◆ 乘客
 - 出发点、目的地、时间窗口、拼车意愿

| | 乘客1 | 乘客2 |
|------|--------|--------|
| 用时 1 | 77 min | 50 min |
| 用时 2 | 87 min | 57 min |
| 费用 1 | 78 元 | 40 元 |
| 费用 2 | 54 元 | 32 元 |



中文分词系统

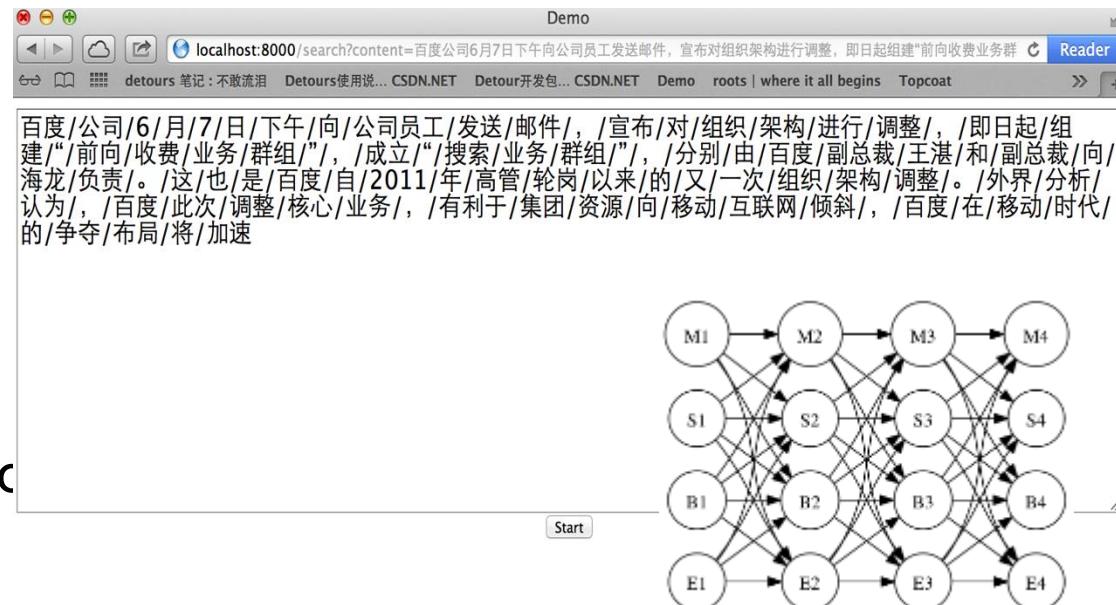
◆ 算法设计

- 基于词典的统计模型
- 基于BMES标注的隐式马尔科夫
- 四状态标注的Viterbi算法
- ⑩ 未登录词模型
- 算法识别率测试
 - 采用公开测试集
- 系统界面
- 开放的Web形式
- ◆ 开发方式：
 - Python语言，Django

测试结果

(SIGHAN2005 BAKEOFF公开语料)

- 准确率 = 0.88
- 召回率 = 0.86



基于KM算法的排课系统

- ◆ 建立图模型和规划模型
- ◆ 转化为求四次二分图的最佳匹配
- ◆ 四次KM算法
- ◆ 三次修改代价矩阵

$$\min \sum Z_{ij} C_{ij} \quad (1)$$

s.t.

$$\forall i, \sum_j Z_{ij} = 2; \forall j, \sum_i Z_{ij} = 4 \quad (2)$$

$$\forall j, \sum_{Z_{ij}=1} (i \% 4) = 6, \sum_{Z_{ij}=1} (i \% \% 4) = 10, \sum_{Z_{ij}=1} (i \% 2) = 2 \quad (3)$$

$$\forall j, \text{if } Z_{ikj} = 1, k = 1, 2, 3, 4 \text{ then } i_{k+1} - i_k > 31, k = 1, 2, 3, \quad (4)$$

$$\text{no more than one } k \text{ s.t. } 65 \leq i_k \leq 128 \quad (5)$$

$$\forall j, \sum_i Z_{ij} C_{ij} \leq 4 \quad (6)$$

$$\text{for some } m \text{ and } n, \exists \text{ three } k, \text{ s.t. } Z_{ikm} = Z_{ikn} \quad (7)$$

$$1 \leq i \leq 320 \quad (8)$$

$$1 \leq j \leq 160 \quad (9)$$

The screenshot shows a software application window titled "排课操作" (Scheduling Operation). The window contains several sections: "显示设置" (Display Settings) with fields for "学号" (Student ID) set to 1, "实验序号" (Experiment Number) set to 0, and buttons for "调课方案" (Lesson Adjustment Plan) and "个人课表" (Personal Course Table); "排课信息" (Scheduling Information) with fields for "总代价" (Total Cost) set to 8, "平均代价" (Average Cost) set to 0.05, and "最高代价" (Highest Cost) set to 4; and "模式选择及参数设置" (Mode Selection and Parameter Settings) with dropdowns for "读入学生代价矩阵" (Read Student Cost Matrix), "代价限值" (Cost Limit) set to 5, "代价调整" (Cost Adjustment) set to 5, and buttons for "期中考试" (Midterm Exam) set to 33 and "64".

| 魏成涛 | 实验时间 | 换课人 | 实验时间2 | 主动代价变化 | 被动代价变化 | 总代价变化 |
|--------|---------|------|----------|--------|--------|-------|
| 1,实验1 | 实验时间 78 | 魏成涛 | 实验时间 155 | 0 | 0 | 0 |
| 2,实验1 | 实验时间 78 | 郑艳兰 | 实验时间 154 | 0 | 0 | 0 |
| 3,实验1 | 实验时间 78 | 樊萍 | 实验时间 158 | 1 | 0 | 0 |
| 4,实验1 | 实验时间 78 | 曾意隆 | 实验时间 147 | 1 | 0 | 0 |
| 5,实验1 | 实验时间 78 | 张若彤 | 实验时间 134 | 1 | 0 | 0 |
| 6,实验1 | 实验时间 78 | 何浩铭 | 实验时间 133 | 0 | 0 | 0 |
| 7,实验1 | 实验时间 78 | 钟鹏飞 | 实验时间 139 | 0 | 0 | 0 |
| 8,实验1 | 实验时间 78 | 李健 | 实验时间 91 | 0 | 0 | 0 |
| 9,实验1 | 实验时间 78 | 吴雨卫 | 实验时间 127 | 1 | 0 | 0 |
| 10,实验1 | 实验时间 78 | 李玲 | 实验时间 123 | 0 | 0 | 0 |
| 11,实验2 | 实验时间 59 | 王玲川 | 实验时间 150 | 1 | 0 | 0 |
| 12,实验2 | 实验时间 59 | 唐河锦翊 | 实验时间 148 | 0 | 0 | 0 |

| 整体课表 | 工位 1 | 工位 2 | 工位 3 | 工位 4 | 工位 5 | 工位 6 | 工位 7 | 工位 8 |
|---------|------|------|------|------|------|------|------|------|
| 实验时间 1 | 周思纹 | | | 潭琴微 | 张伟 | 宋楚文 | 柏祥 | 李 |
| 实验时间 2 | 李晨 | 胡溢玲 | 刘子瑜 | 匡茂源 | 乐旋 | 李常 | 李璟弘 | 雷 |
| 实验时间 3 | 欧俊鹏 | 杨江龙 | 李浩云 | 欧阳泽 | 郑艳兰 | 蒋璐阳 | 黄焱灵 | 冯 |
| 实验时间 4 | 杨松霖 | 周哲先 | 唐龙 | 谭佳宇 | 李湘东 | 刘韬 | 李琪琛 | 唐 |
| 实验时间 5 | 邓恒香 | 何鹏 | 谢紫榕 | 肖泽阳 | 钟鹏飞 | 黄兴文 | 欧文奇 | 刘 |
| 实验时间 6 | 杨东旺 | 蒋秦政 | 黄将 | 曾意隆 | 魏成涛 | 李玲 | 李健 | 蔡 |
| 实验时间 7 | 张伟浩 | 周旭 | 李俊豪 | 唐佳坤 | 唐能智 | 蒋浩健 | 袁文鹏 | 唐 |
| 实验时间 8 | 朱怡心 | 王永强 | 樊萍 | 蒋华 | 周世东 | 欧阳昊 | 何浩铭 | 雷 |
| 实验时间 9 | 黄婷婷 | | | 欧阳能静 | 颜浩宇 | | | 雷 |
| 实验时间 10 | 刘枳琰 | 张颖 | 伍宁超 | 贺源镔 | 肖俊名 | 雷明 | 陈志群 | 雷 |

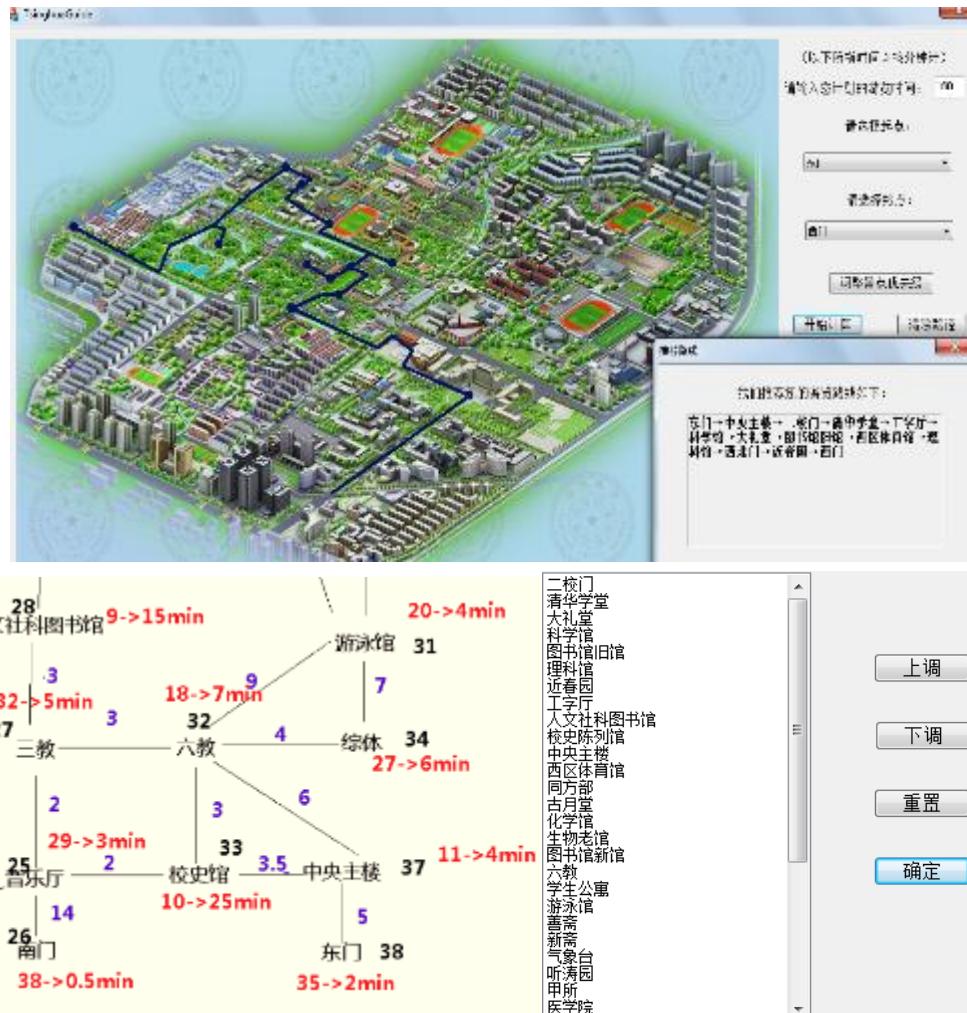
清华导游程序

◆ 问题描述

- 根据游客给定的时间、起点和终点，给出在清华校内浏览的最佳路线
- 游客自定义景点优先级

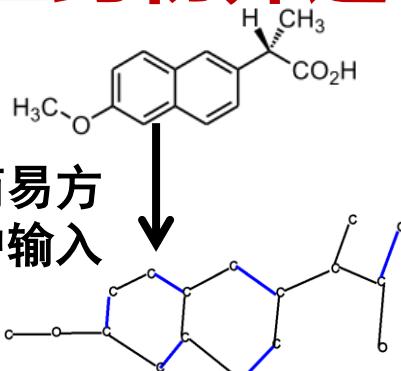
◆ 问题求解

- 图论建模，实地测量
- 优先级制定
- 改进的遗传算法解决旅行商问题



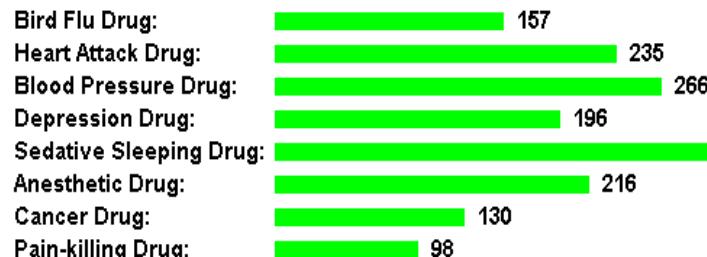
潜在药物筛选

待测目标
化学结构



matrix
17
0 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0
0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 4 3
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

权矩阵转化
(图论建模)



相似度匹配：分析得到未知药物的可能药效

This molecule is a potential: Sedative Sleeping Drug

深度优先
(递归算法)

与标准库比较

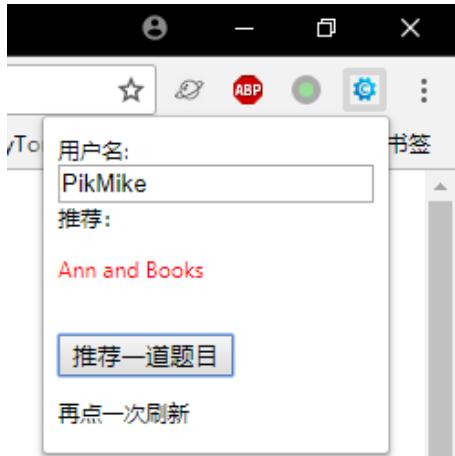
CodeForces做题推荐系统

- ◆ 如何在CodeForces的题海中找到适合自己的题目？

| 遇到的问题 | 解决方案 |
|--------------|------------------------------|
| 如何量化题目难度？ | 通过选手提交记录推测用时 使用时间的对数为题目难度 |
| 选手水平持续变化怎么办？ | 主要保留近期数据，丢弃老数据 |
| 数据噪音很大怎么办？ | 去除一定量可疑选手 |

- ◆ 主要方法：基于协同过滤算法预测出每个题目对于每位选手需要做多少时间，以此进行合适难度的题目推荐。
- ◆ 最终效果：与做题时间估测平均误差为13.3分钟
- ◆ 源码和插件下载→ 

实际应用截图→



多机调度问题

- ◆ 如何让大作业分工均等？让程序运行时间再-1s？
- ◆ 合理分工调度，就能做得更好。
- ◆ 难点：啃证明、数据生成、调参、算法选取与实现
- ◆ 经典NP问题，不断改进算法，1500行代码，十余种算

| | 经典贪心算法 | 随机算法 | 爬山算法 |
|--------|---------------|----------------|-----------|
| 时间复杂度 | $O(n \log n)$ | $O(Tn \log n)$ | $O(Tn^2)$ |
| 空间复杂度 | $O(n)$ | $O(n)$ | $O(n)$ |
| 解的相对比例 | 0% | 差50% | 改进30% |

| | 搜索 | 折半搜索 | 动态规划1 | 动态规划2 |
|--------|---------------------------|-------------------------------|-------------|-----------------------|
| 时间复杂度 | $O(\text{Stirling}(n,m))$ | $O(\text{Stirling}(n/2,m)^2)$ | $O(m^3 n)$ | $O((m+\log n) * 2^n)$ |
| 空间复杂度 | $O(n^2)$ | $O(\text{Stirling}(n/2,m))$ | $O(2^n)$ | $O(2^n)$ |
| 极限数据规模 | $n=20(66s)$ | $n=22(256s)$ | $n=20(50s)$ | $n=26(74s)$ |

GZ Search 思修慕课搜索引擎

GZ SEARCH

中国经济

GZ一下

答案 新闻

GZ为您找到相关结果约74个

宋氏三姐妹及家族成员旗袍在沪集中展出...

上海6月13日电(王子涛)宋氏三姐妹及亲属的近30件旗袍近日在上海大规模集中亮相，这些旗袍由宋家的亲属们长期保存，近十年间陆续从美国和加拿大带回中国。 宋庆龄和她的姐姐宋霭龄、宋美龄，她们生活上...

<http://www.chinanews.com/cul/2018/06-13/8536693.shtml>

从好莱坞到百姓舞台豫剧名家三步曲传承曲艺文化...

郑州6月13日电：从好莱坞到百姓舞台豫剧名家三步曲传承曲艺文化 记者李贵刚
从中原地区到走出国门，从东亚、欧洲到美国百老汇，从好莱坞杜比大剧院再到市场化尝试.....多年来，著名豫剧表演艺术家李树...

<http://www.chinanews.com/cul/2018/06-13/8537115.shtml>

林燕妮离世：莫问佳人何在知交多半零落...

6月4日晚，香港传奇女性林燕妮因肺癌逝世，享年75岁，这位前半生“生如夏花之绚烂”的女子晚景可谓“烟花寂寞”。要强了一辈子，林燕妮不愿让人看到病榻前被疾病摧毁的容颜，所以，去世的消息是通过其...

<http://www.chinanews.com/cul/2018/06-13/8536645.shtml>

亚洲巨型企业自航绞吸船“天鲲号”完成首次海试...

6月12日，经过近4天的海上航行试验后，亚洲巨型自航绞吸船“天鲲号”返航至江苏启东船厂码头停泊。海试期间，“天鲲号”的动力、推进系统等设备成功经受了海洋环境考验。“天鲲号”可谓疏浚“重器”，去年11月...

<http://www.chinanews.com/tp/hd2011/2018/06-12/821806.shtml>

不可思议！日本八幡平现“龙眼”奇景...

当地时间2018年6月5日，日本，在横跨岩手县与秋田县的八幡平山顶附近的“镜沼”，伴随着积雪融化，日前出现了奇特的“龙眼”景象。据报道，八幡平“龙眼”奇景只在每年5月下旬至6月初的雪融期间短暂出现。当...

<http://www.chinanews.com/tp/hd2011/2018/06-13/821881.shtml>

九旬教师守候留守儿童语文英语书法全都教...

近日，安徽马鞍山，90岁的退休教师叶连平，义务为学生辅导英语18年，分文不收。叶连平原本是初中语文老师，2000年，他在家中开设课堂（后更名为“留守儿童之家”），给小学和初中的留守儿童们平日义务辅导英...

<http://www.chinanews.com/tp/hd2011/2018/06-13/821945.shtml>

探秘海南高考评卷场考生答题卡武警24小时值守...

6月12日，在海南师范大学高考评卷场，工作人员仔细拆封答卷并进行扫描入电脑，整个过程井然有序。在评卷现场的入口，除了有武警把守之外，进出评卷现场的评卷老师和工作人员每次进入都要在门口的身份证识别系统中...

<http://www.chinanews.com/tp/hd2011/2018/06-13/822108.shtml>

世界杯32强·D组阿根廷·图片频道-中国新闻网...

中新网微信公众号中新网客户端本网站所刊载信息，不代表中新社和中新网观点。刊用本网站稿件，务经书面授权。未经授权禁止转载、摘编、复制及建立镜像，违者将依法追究法律责任。QQ群公安网安：110102...

<http://www.chinanews.com/tp/hd2011/chart/2018/06-08/698.shtml>

搜索页

from Spider. (202)
Network thread received a data string from Spider. (203)
Network thread received a data string from Spider. (204)
ok !
Network thread received a data string from Spider. (205)
Network thread received a data string from Spider. (206)
Network thread received a data string from Spider. (207)
Network thread received a data string from Spider. (208)
Network thread received a data string from Spider. (209)
“鲲”？
Network thread received a data string from Spider. (210)
Network thread received a data string from Spider. (211)
Network thread received a data string from Spider. (212)
Network thread received a data string from Spider. (213)
Network thread received a data string from Spider. (214)
Network thread received a data string from Spider. (215)
Network thread received a data string from Spider. (216)
Network thread received a data string from Spider. (217)
Network thread received a data string from Spider. (218)
Network thread received a data string from Spider. (219)
Network thread received a data string from Spider. (220)
Network thread received a data string from Spider. (221)
Network thread received a data string from Spider. (222)
Network thread received a data string from Spider. (223)
Network thread received a data string from Spider. (224)
Network thread received a data string from Spider. (225)
Network thread received a data string from Spider. (226)
Network thread received a data string from Spider. (227)
Network thread received a data string from Spider. (228)
Network thread received a data string from Spider. (229)
Network thread received a data string from Spider. (230)
Network thread received a data string from Spider. (231)
Network thread received a data string from Spider. (232)

应该会有，敬请期待。”不过记者追问具体上线时间时，对方未正面回答。北京安博(成都)律师事务所律师陈军告诉记者，“鲲游戏”广告大多属于互联网广告，这一类广告的管控没有专门立法。但是不管作为广告发布者或者广告主而言，都不应发布虚假广告。如果出现广告内容不实的情况下

处虚假广告，“通过虚假的广告让消费者耗费时间和流量下游戏，也是侵犯了别人的合法权益。” 游戏厂商为何偏爱“鲲”？业内人士：获取用户成本小
事实上，网络奇葩广告并不罕见。在《贪玩蓝月》意外走红之后，为什么“鲲”能够获得游戏厂商的青睐？国内一游戏平台的负责人李红(化名)告诉记者，公司也曾制作并发行过“鲲”元素的广告。这类广告的效果也是出奇的好。

“鲲”广告可以带来一大波流量，而这也是所有游戏发行者最为看重的。“网友看到这类广告会很好奇，因为作为上古神话里面的比较知名的动物，所有的鲲的形象都是想出来的。网友们看到这个广告会很好奇，想知道鲲是不是真的长这样，便会点击进去。”好奇心

来的是非常可贵，加之目前“鲲”的素材没有明确版权，所以广告的成本也就极低。李红说，传统的仙侠游戏广告获取一个有效用户，平均需要60元成本，而“鲲”广告获取一个用户最低只需要几块钱。“在这样的成本优势下，很多游戏厂商宁愿担负虚假广告的风险，也会来抢着分一杯羹。”另一位游戏行业的内部人士也认为，“鲲”作为引流的“利器”，只会吸引越来越多的游戏发行商。“像鲲这类虚假广告，也绝不会最后一次出现。” 李红透露，公司也确实有想法在游戏增加“鲲”的元素。“比如人物的坐骑和翅膀里，会慢慢带有鲲元素。”她说，平台之前也上线了一两款带有“鲲”元素的游戏，对比发现，用户的留存率比没有“鲲”的游戏高30%以上。

不过目前，“鲲”广告引流的游戏大多数仍然与“鲲”无关，长期如此，“鲲”这个引流利器被玩坏了怎么办？“顺其自然，鲲这个元素被淘汰了，还会有什么梗？”李红说。

华西都市报-封面新闻记者 杨尚智@GZ_RELATED@GZ_END@

后台页

A:都属于基本制度的范畴

B:不可存在于同一社会制度下

C:都是经济发展的手段

D:是社会主义和资本主义相区别的标志

答案：C【解析】计划和市场都是经济手段，计划多一点还是市场多一点，不是社会主义和资本主义的本质区别。
@GZ_URL=http://www.9120.cn/m/view.php?id=44981&page=10
GZ_CONTEXT@7 . 下列对计划和市场的关系，认识正确的是（ ）

A:都属于基本制度的范畴

B:不可存在于同一社会制度下

C:都是经济发展的手段

D:是社会主义和资本主义相区别的标志

答案：C【解析】计划和市场都是经济手段，计划多一点还是市场多一点，不是社会主义和资本主义的本质区别。

@GZ_RELATED@GZ_END@

Sending:

@GZ_TYPE@0@GZ_TITLE@0 . “字字写来都是血，十年辛苦不寻常”和“文不甚深，言不甚俗”分别讲的是中国古典文学中的（ ）

A:《水浒》和《聊斋志异》

B:《西游记》和《聊斋志异》

C:《儒林外史》和《三国演义》

D:《红楼梦》和《三国演义》

答案：D
@GZ_URL=http://www.9120.cn/m/view.php?id=44981&page=10
GZ_CONTEXT@8 . “字字写来都是血，十年辛苦不寻常”和“文不甚深，言不甚俗”分别讲的是中国古典文学中的（ ）

A:《水浒》和《聊斋志异》

B:《西游记》和《聊斋志异》

C:《儒林外史》和《三国演义》

D:《红楼梦》和《三国演义》

答案：D
@GZ_RELATED@GZ_END@

[14/Jun/2018 16:03:29] "GET /searchInput?searchText=E4%8B%AD%E5%9B%BD%E7%BB%8F%E6%85%8E&seriousSearch=G%6E%90%9C%E7%84%A2 HTTP/1.1" 200 194085

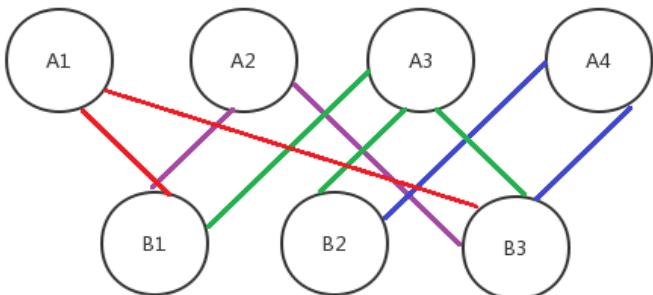
"Lyrics-MacBook-Pro.lo" 00:03 15-Jun-

教学楼集中化排课系统

- 集中排课 \Rightarrow 减少通勤 \Rightarrow 学堂路治堵
- 已知选课情况，设计排课方案，最小化两节课间的总通勤成本

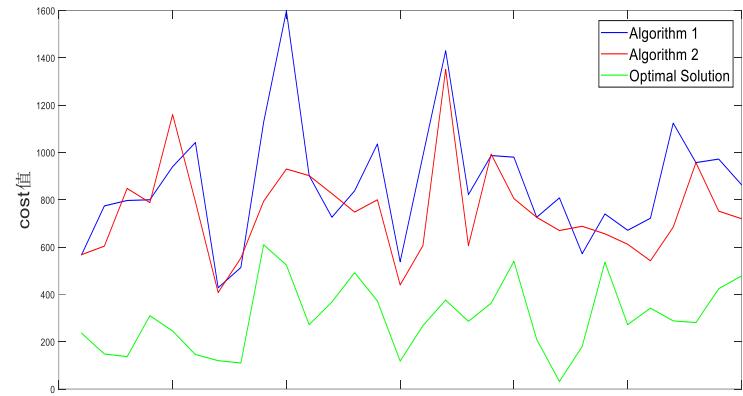
◆ $\arg \min_T G(T) = \sum_{i=0}^n \sum_{j=0}^{e_i} w_{ij} \times C_{T_i T_j}$

◆ *subject to* $\forall k \in \{1, 2, \dots, b\}, \forall l \in \{1, 2, \dots, r\},$
 $\sum_{i=0}^n \text{sgn}(T_{iR_l} = k) \leq \text{Capa}(B_{kR_l})$



选课二分图建模

结点：课程，边权：公共选课人数
 颜色：课程所在教学楼



可行算法、近似优化算法与最优算法的比较

排课结果展示

| 课程名 | 教学楼 | 教室号 |
|-------------|-----|-------|
| 微积分A(2) | 六教 | 6C300 |
| 模拟电子技术基础 | 六教 | 6C101 |
| 概率论与数理统计 | 六教 | 6C102 |
| 概率论与数理统计 | 六教 | 6C201 |
| 微积分A(2) | 六教 | 6C202 |
| 概率论与随机过程(1) | 六教 | 6A017 |
| 概率论与随机过程(1) | 六教 | 6A016 |
| 分子生物学(英) | 六教 | 6A018 |
| 数字逻辑电路 | 三教 | 3200 |
| 微积分B(2) | 三教 | 3300 |
| 微积分B(2) | 三教 | 2101 |
| 微积分B(2) | 三教 | 2102 |
| 高等微积分(2) | 三教 | 2301 |
| 建筑数学 | 三教 | 2302 |
| 建筑设计原理 | 新水 | 407 |
| 计算机程序设计基础 | 六教 | 6A117 |
| 中国文明 | 六教 | 6A215 |
| 俄语 | 六教 | 6A116 |
| 传统民居与乡土建筑 | 六教 | 6A118 |
| 微积分A(2) | 六教 | 6A214 |
| 微积分B(2) | 六教 | 6A216 |
| 离散数学方法 | 六教 | 6A314 |
| 微积分B(2) | 六教 | 6A316 |
| 大学物理A(1) | 六教 | 6A414 |
| 有机化学B | 六教 | 6A416 |
| 计量经济学(1) | 六教 | 6A315 |
| 计算机仿真 | 六教 | 6A415 |
| 微积分B(2) | 四教 | 4201 |
| 概率论(1) | 四教 | 4202 |
| 概率论(1) | 四教 | 4203 |
| 大学物理B(1) | 四教 | 4204 |
| 大学物理B(1) | 四教 | 4205 |
| 大学物理B(1) | 四教 | 4206 |
| 大学物理A(1) | 四教 | 4101 |
| 大学物理A(1) | 四教 | 4102 |

实时Low-poly图像生成器

◆ Low poly图像风格设计

- 将原有图片用三角形分割并填充纯色

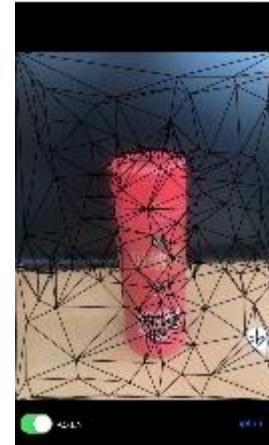
◆ 难点：快速边缘检测与三角剖分，

- 应用了基于图论的边缘检测算法及与三角剖分算法

- 全部内容由swift实现并在手机上成功

- 用GPUImage2以加速渲染，实时捕捉并渲染摄像头画面

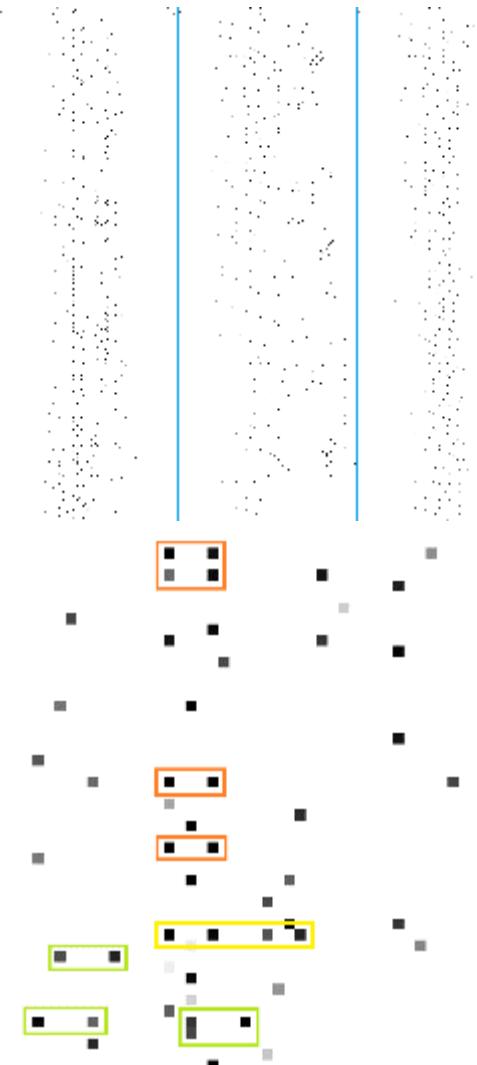
- 效果：640*480像素，平均30帧/秒



简要流程说明：
边缘检测并取点
三角剖分
着色

自动作曲

- ◆ 单次生成16小节多声部旋律
- ◆ 使用DCGAN模型
 - 训练集含498首音乐，随机取出大小为 88×256 的样本
 - 生成器和判别器各有10个卷积层
- ◆ 计算机辅助作曲、音乐制作自动化



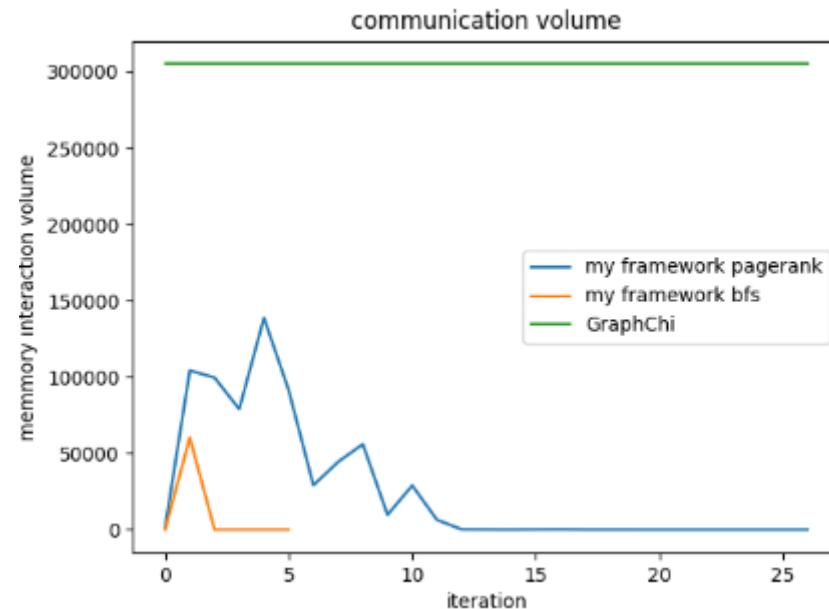
基于消息传递的图计算框架(AutoProgrammer)

◆ 问题描述

- 图算法实现千差万别，如何设计一个模板代码框架去高效实现不同的图算法，使程序员只需要实现给定的几个接口函数就可以实现不同的图算法

• 问题求解

- 提出了基于消息传递的算法抽象方式，并使其并行化
- 使用 Go 语言编写模板代码
- 通过模板代码实现 BFS、PageRank 和最短路径算法
- 消息传递机制相比现有算法框架显著降低了内存访问的次数



基于神经网络的输入法

谭竣文

- ◆ 输入法 → 翻译
- ◆ 对时序较为关注
 - 空格填补序列使拼音汉字等长
 - 神经网络作为分类器，每个位置计算概率分布
 - 使用 CNN 获取局部编码，代替常规的 Embedding
- ◆ 准确率约 97.7%
 - 平均编辑距离约 0.07
 - 平均交叉熵约 0.08
- ◆ 参数文件较小，约 90MB
- ◆ 可实时更新
- ◆ 代码大约30k



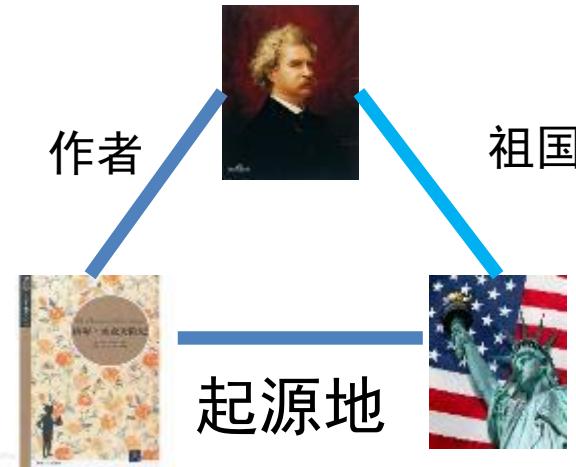
网络结构

Neural Input Method Engine a
[Default> jiyushenjingwangluodeshurufa
基于神经网络的输入法
Default>]

最终效果

长文本关系抽取

- 问题描述
 - 给一篇长文本，要求：
 - 找出所有实体两两之间的关系
- 问题难点
 - 需要记忆能力
 - 需要推理推理
- 问题求解
 - 设计记忆模块
 - 设计推理模块



Github网址：

<https://github.com/FuxeyHuang/MK-RNN>

计81 黄励新
计83 杨耀良

基于图论的旅行规划器

李家昊 计73

图论建模

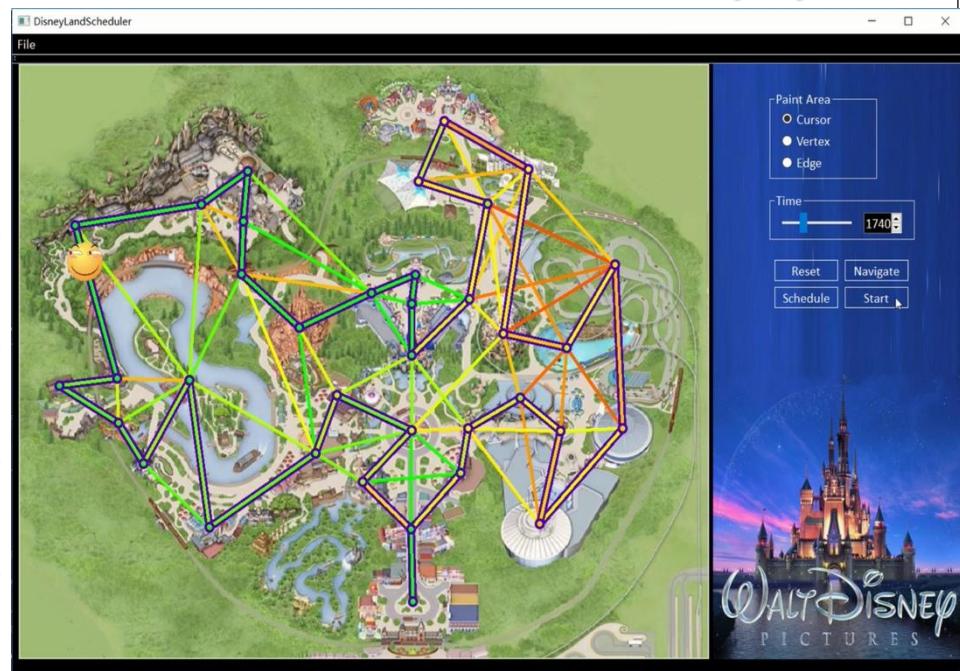
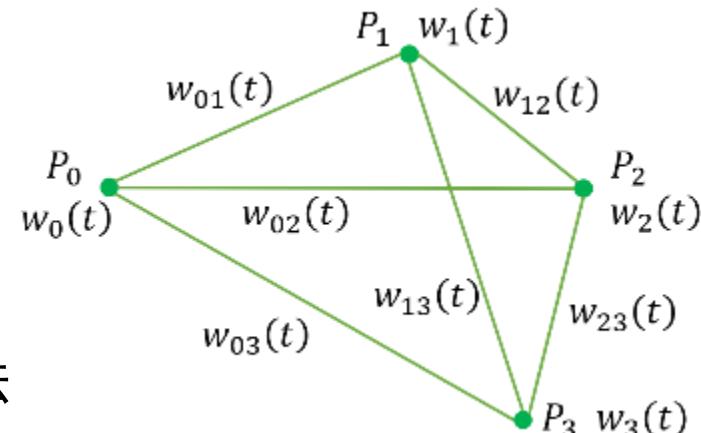
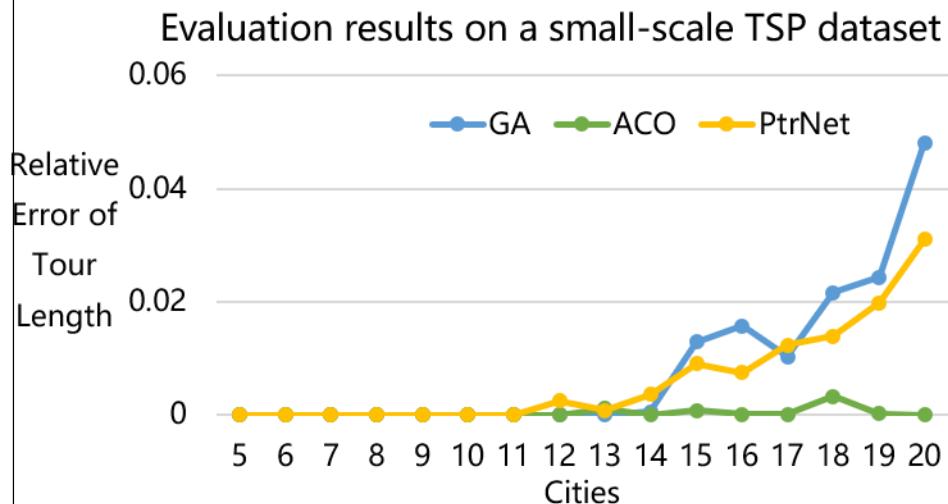
- 已知图上所有权重随时间变化的函数
- 规划一个旅行商问题

解决方案

- Pointer Network / 蚁群算法 + 遗传算法

工作量

- C++ 3042 行, Python 418 行.



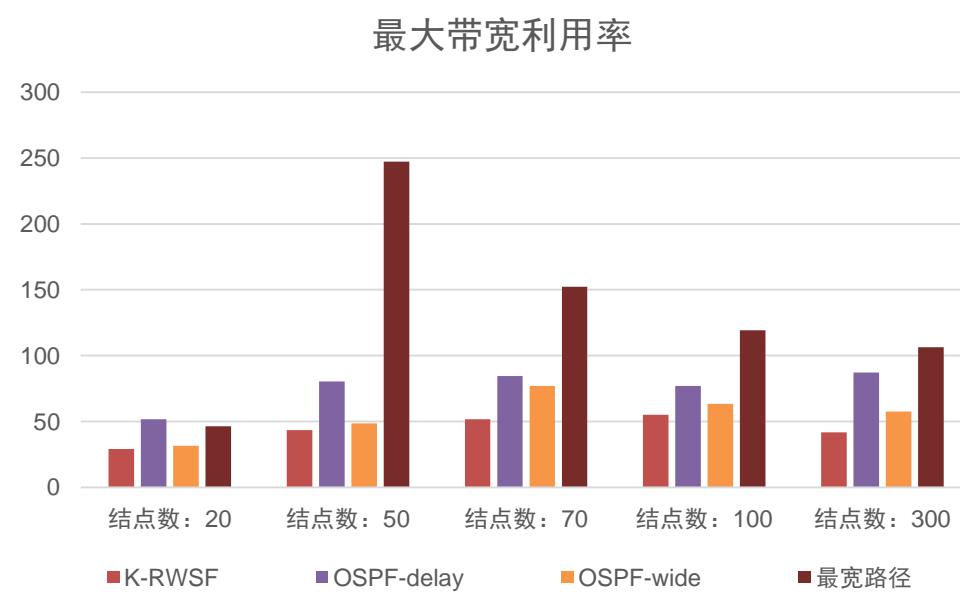
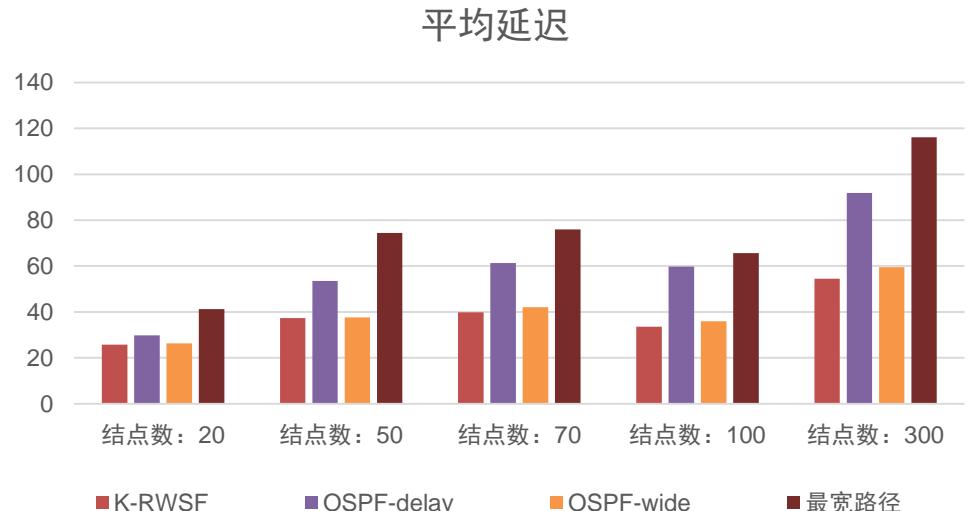
基于SD-WAN的双目标优化

- ◆ 延迟约束&负载均衡
- ◆ 应用背景：VIP kid等
- ◆ 优化目标：

$$\min \left\{ \max_{i,j=1 \sim n} \{ V_{i,j} \} \right\}$$

- ◀ $\min \left\{ \frac{\sum_{i=1}^n \sum_{j=1}^s t_{i,j}}{n} \right\}$
- $K=1$, 动态最宽路
- $K \rightarrow \infty$, 动态延迟最短路

- ◆ 导出子图最宽路



基于随机森林的股票量化分析

选题与方法

股票时时波动大，但涨跌有一定先兆。经典方法多是用lstm，但随机森林速度快，可以用上更相近的时间。

模型对比与可视化

对比

稳定持仓
• 对控股股票操作
• 交易时间短
• 避免过夜风险

投机方式
• 操作间隔：5分钟
• 做多：涨超过阈值
• 做空：跌超过阈值
• 持平预测不操作

谨慎操作
• 现实持平多，持平与可盈利点（涨、跌）比例3：1
• 预测持平更多

决策树

上面为可赢利点，下面为预测结果——决策比例低
随机森林为涨，绿色为跌

多层感知机

支持向量机

泛化能力

当日有效假设

svm是强分类器
线性svm，训练较快，但效果不行
高斯核svm，训练较慢，效果一般
只是一个超级函数，视为输入到输出的映射

小结

◆ 利用图论解决问题的一般过程

- 分析问题

- 确定输入信息、输出目标和约束条件等

- 使用图论进行建模

核心参数选取与约束条件简化

- 用图描述题设条件

- 将求解目标映射到图论体系中

- 使用现有的图论方法/设计新的算法进行求解

- 组合使用多种现有的图论方法

- 启发式算法

- 搜索，剪枝优化

◆ 期待更多的经典成果——将来自你们！

Questions?