



人工智能前沿之

决策智能




计算机学院

余皓然

2020/12/7

主要内容

决策智能部分研究方向介绍

-  棋牌博弈
-  数字经济
-  多智能体协作

阅读资料介绍及经验分享

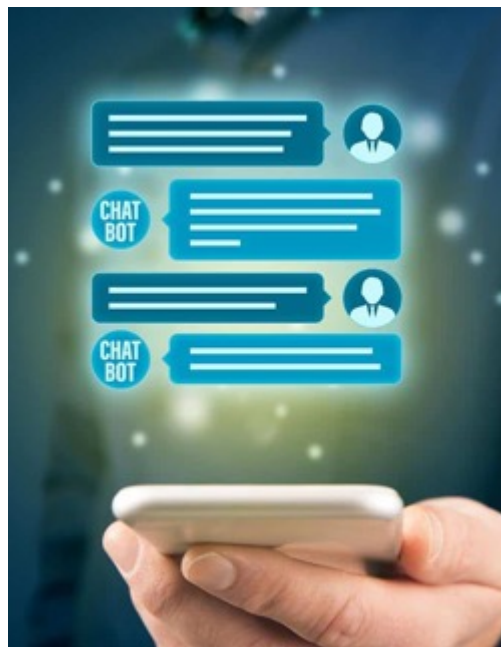


决策智能部分方向介绍

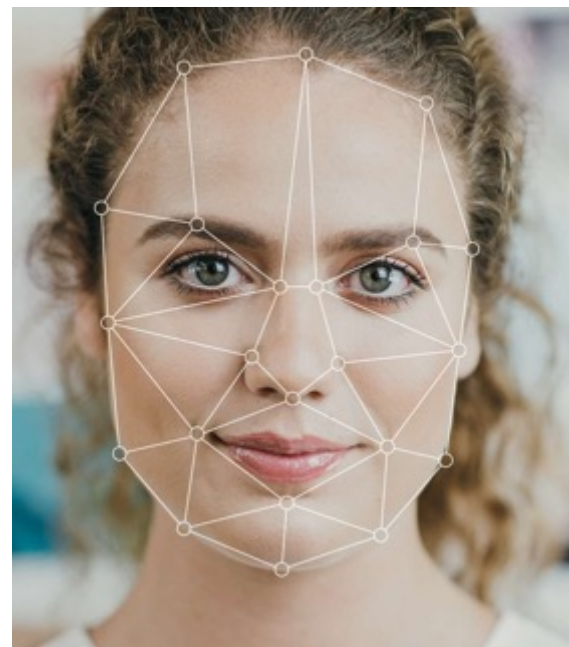
AI之感知智能



语音识别



语言处理

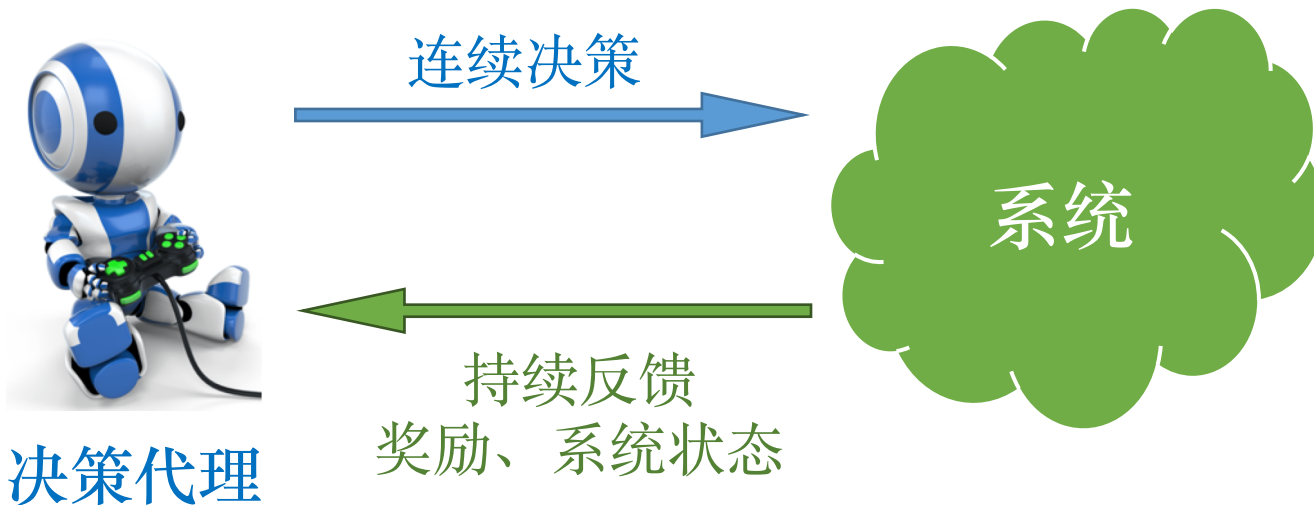


人脸识别

目的：令AI具有听觉、视觉、理解语言的能力

代表方法：深度神经网络

AI之决策智能



目的：令AI具有在动态不确定系统中的决策能力

代表方法：强化学习

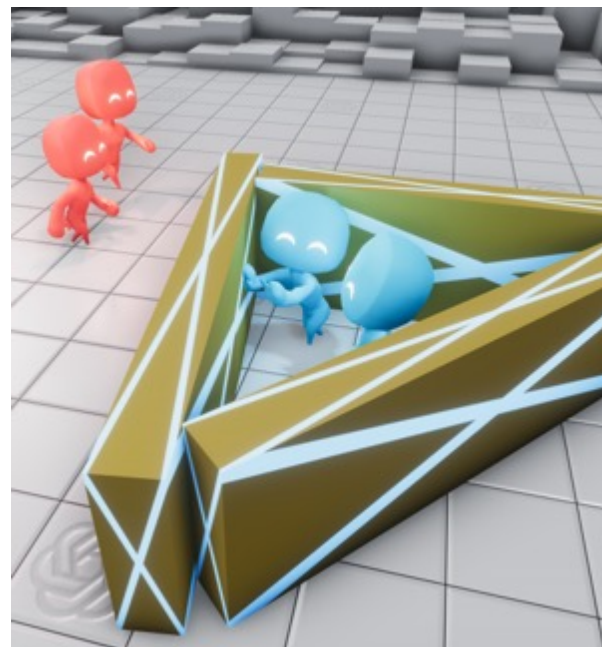
决策智能实例介绍



棋牌博弈

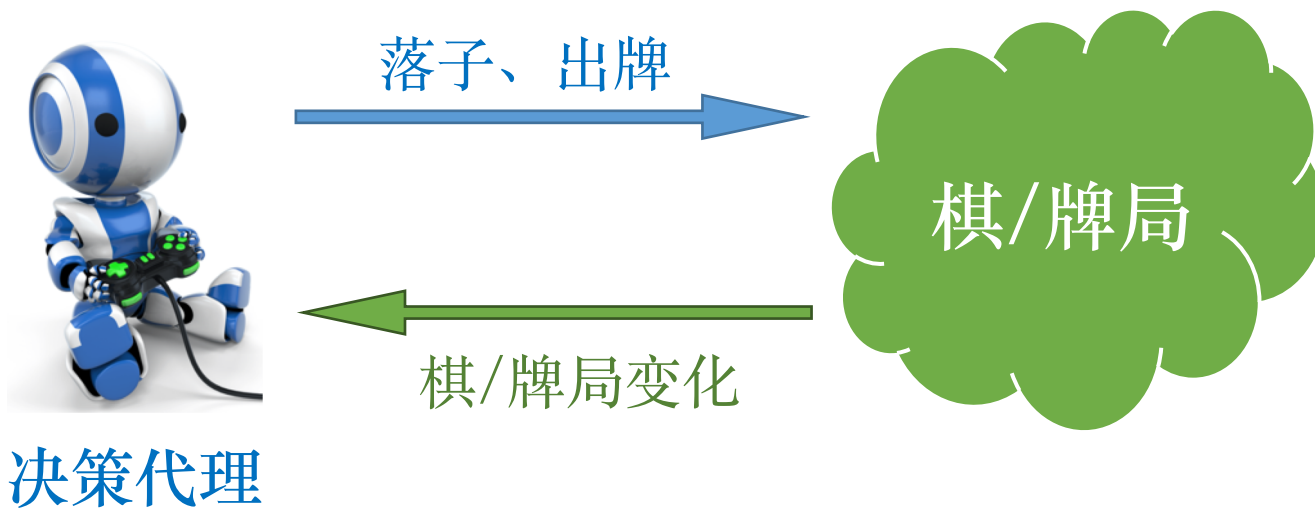


数字经济



多智能体协作

1. 棋牌博弈



1. 棋牌博弈

以击败人类顶级玩家为标志

1997

国际象棋



IBM-深蓝

2016

围棋



谷歌-AlphaGo

2017

德州扑克



CMU-libratus

2019

麻将



微软-Suphx

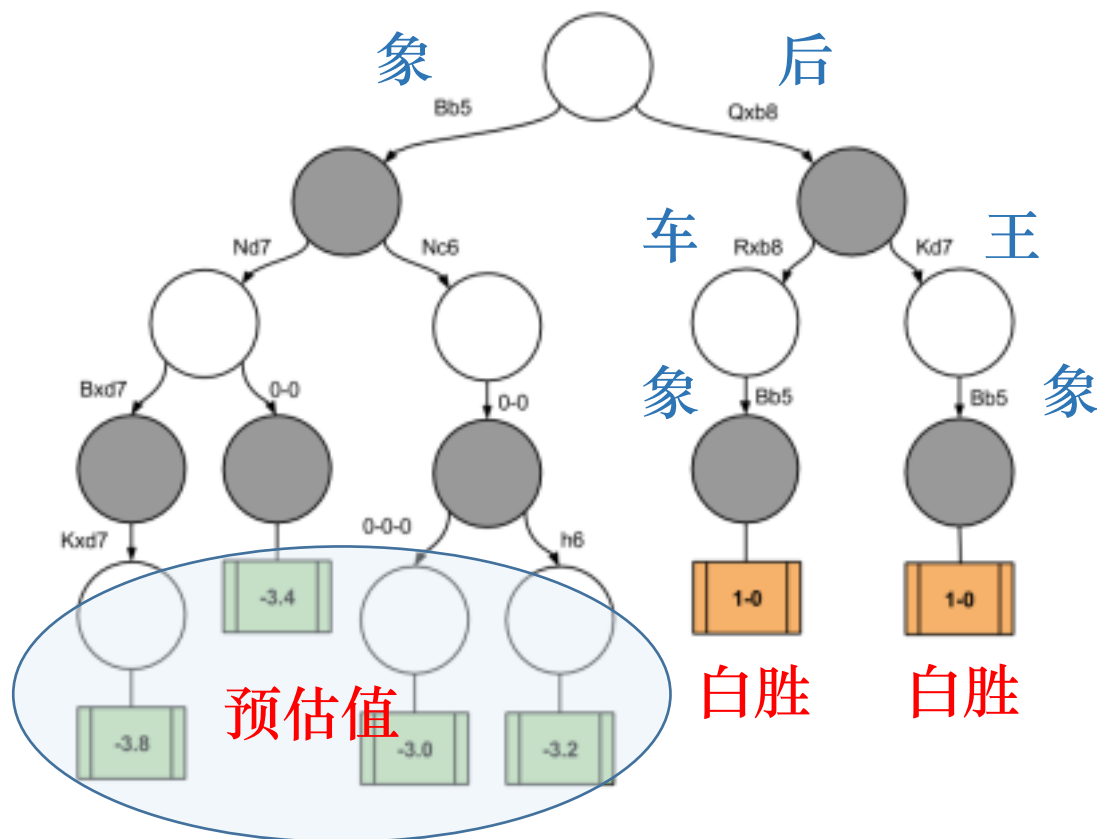
研究趋势： 问题复杂度增高、由完全信息博弈到非完全信息博弈

本节部分内容源自刘铁岩博士报告“Suphx: Superhuman Mahjong AI”

1. 棋牌博弈—国际象棋



IBM-深蓝

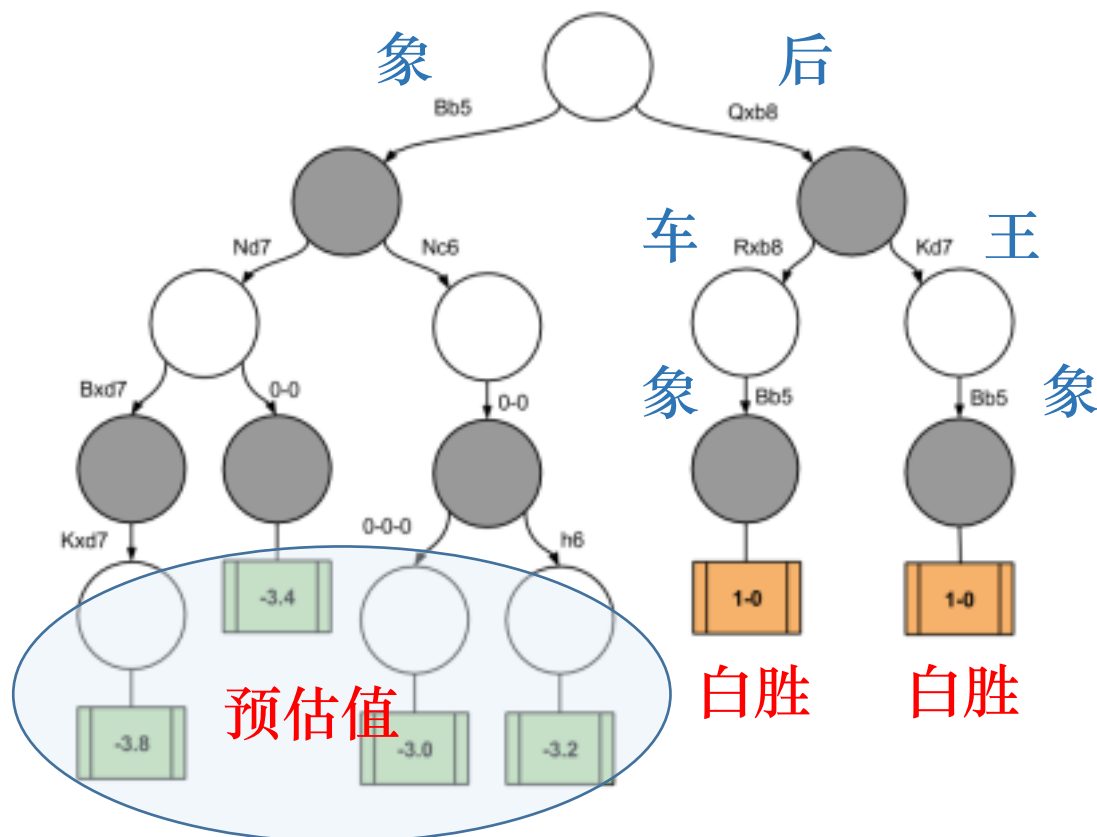


研究方法：用博弈树对棋局所有可能发展进行推演、根据反向推导决定当前落子决策

1. 棋牌博弈—国际象棋



IBM-深蓝

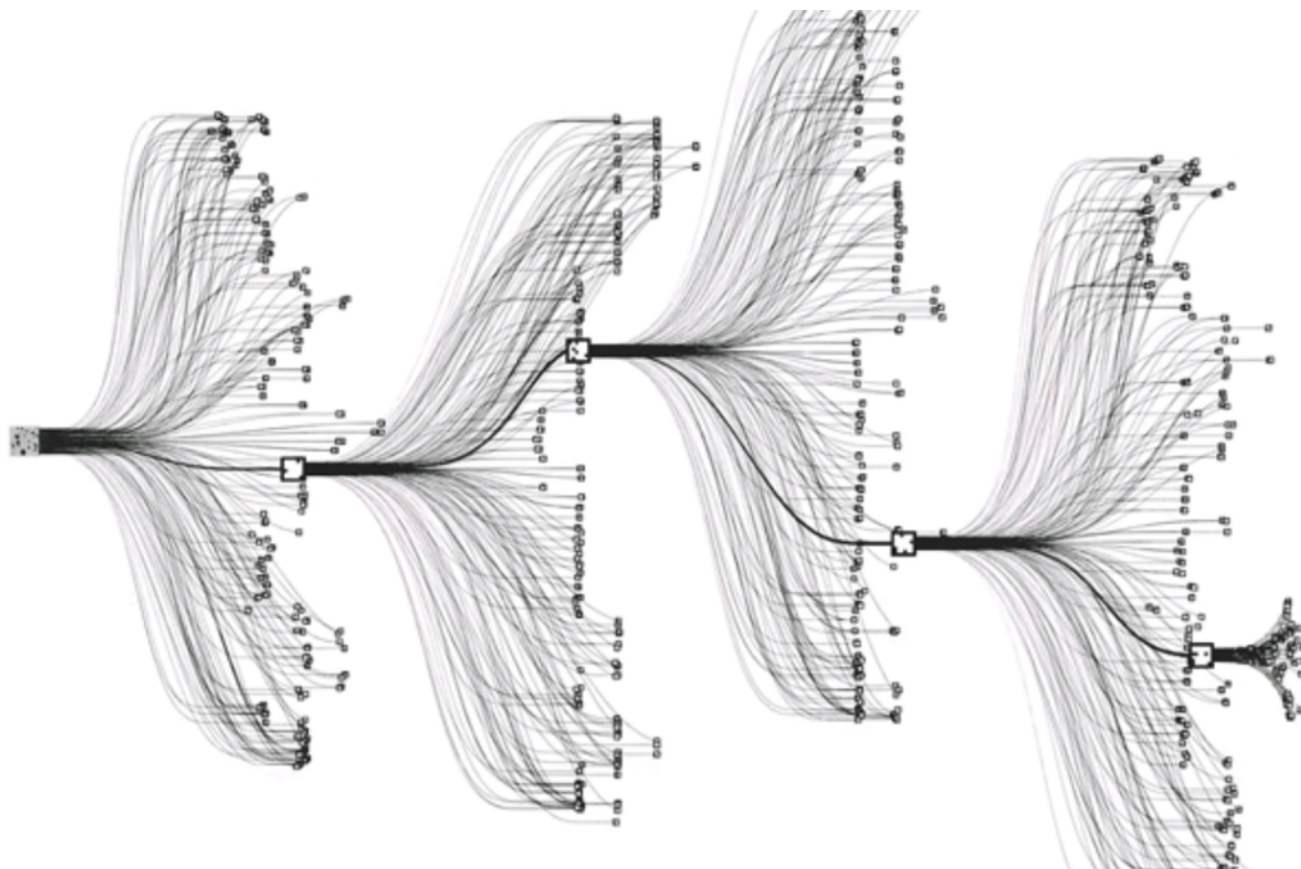


由于算力限制，AI不能对完全博弈树进行分析（ 10^{120} 变化），只能分析未来有限步（如6步）变化。对叶子结点的预估值计算影响算法性能

1. 棋牌博弈—围棋



谷歌-AlphaGo



棋局变化极多 (10^{360} 变化)，用博弈树方法难以分析较优决策

1. 棋牌博弈—围棋



谷歌-AlphaGo

AlphaGo三个主要技术——（1）有监督学习

用卷积神经网络从3000万专业棋手落子学习针对不同棋局的落子决策



一次落子

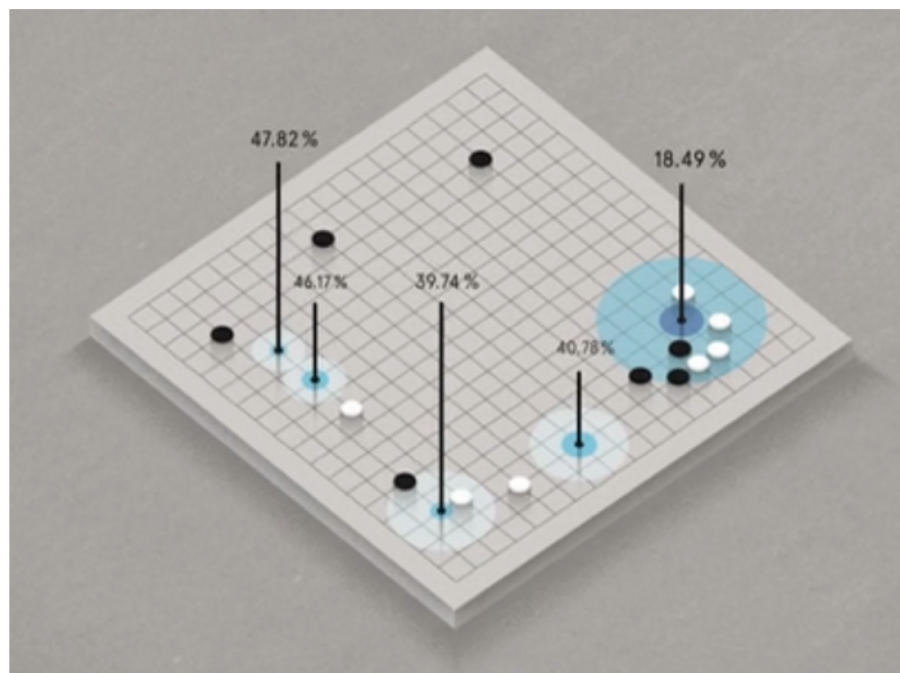
1. 棋牌博弈—围棋



谷歌-AlphaGo

AlphaGo三个主要技术——（2）强化学习

用已得到的决策AI进行3000万局自我博弈，根据博弈胜负倒推计算每一步棋的价值，得到价值网络



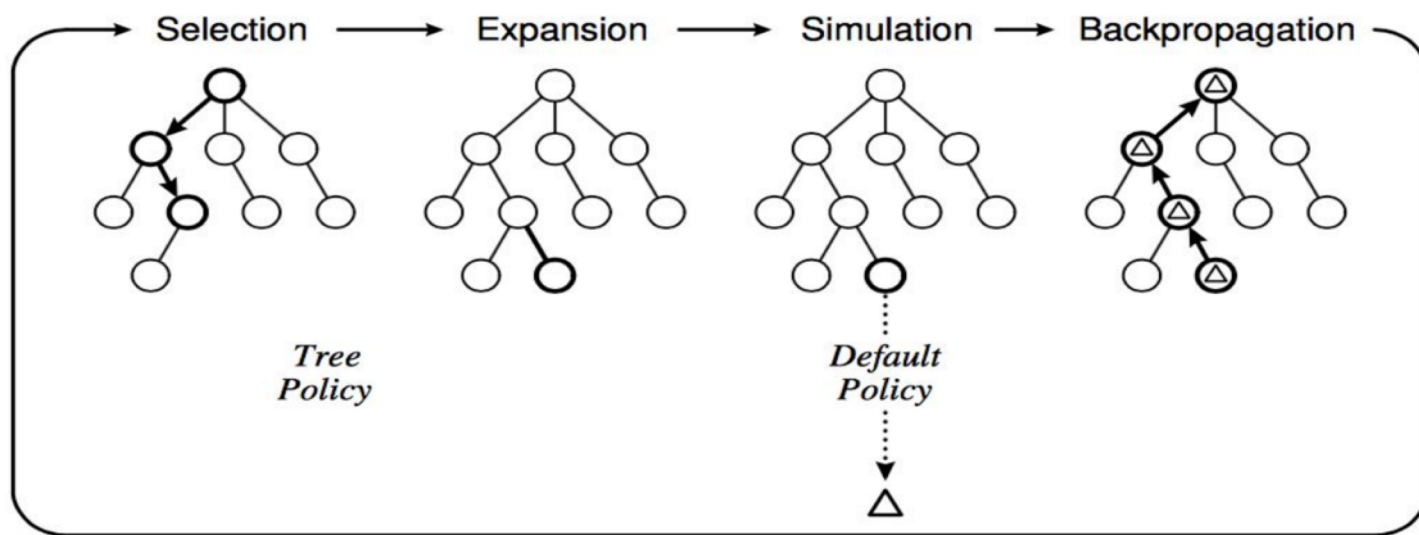
1. 棋牌博弈—围棋



谷歌-AlphaGo

AlphaGo三个主要技术—— (3) 蒙特卡洛树搜索

根据 (1) 的决策网络和 (2) 的价值网络，对现有棋局进行有限步数的推演（50~60步），综合决定当前决策



1. 棋牌博弈—德州扑克



CMU-libratus



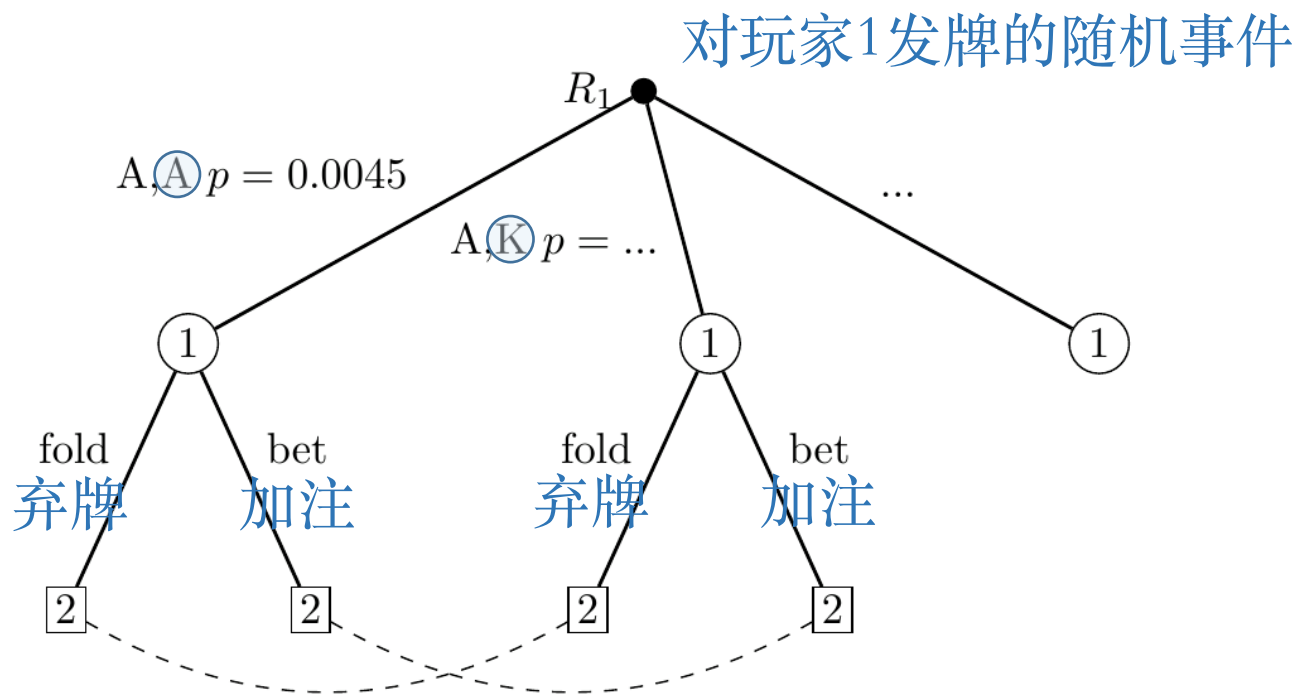
Science, Aug 2019

1. 棋牌博弈—德州扑克



CMU-libratus

非完全信息博弈 例子：两个玩家各有一张牌，玩家2拿到A

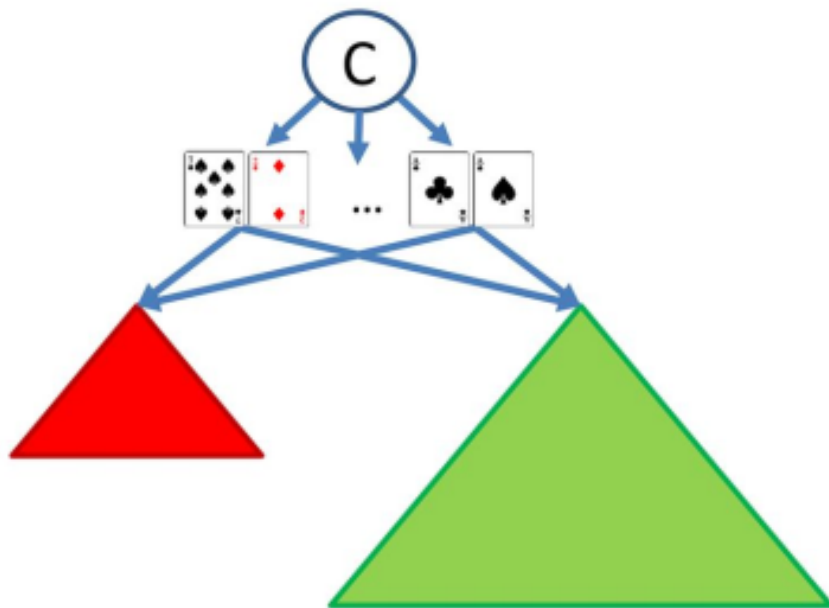


1. 棋牌博弈—德州扑克



CMU-libratus

- Libratus: (1) 近似求解非完全信息博弈的纳什均衡
(2) 根据即时更新的牌局信息，对子博弈进行精细化求解

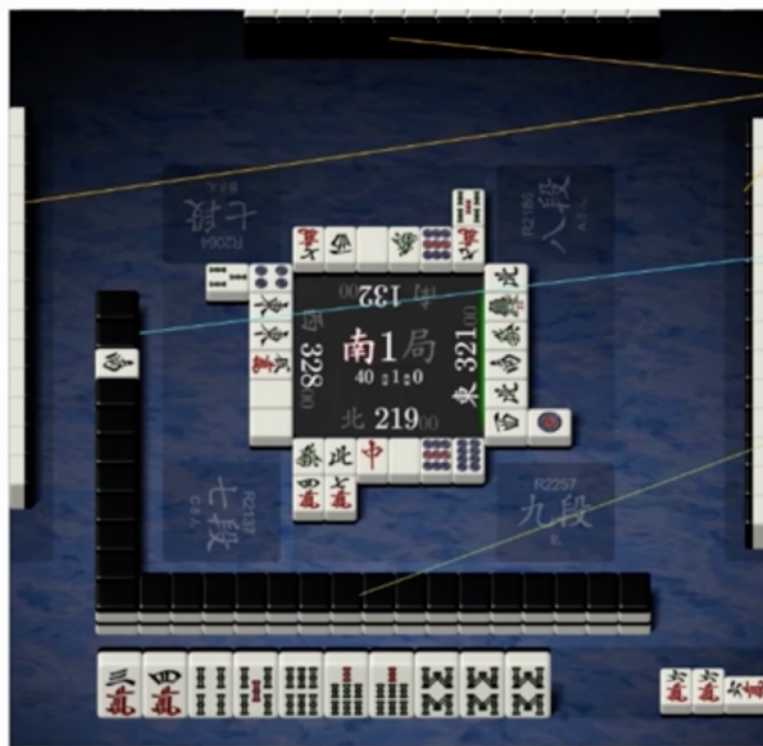


1. 棋牌博弈—麻将



微软-Suphx

未知信息远多于德州 (10⁴⁸ vs 10³)、博弈树不规则、计分规则复杂



39张手牌未知

84张墙牌未知

主要思路技术可见刘铁岩博士报告“Suphx: Superhuman Mahjong AI”

1. 棋牌博弈



相较真实世界的交互博弈，更为简单、规则明确

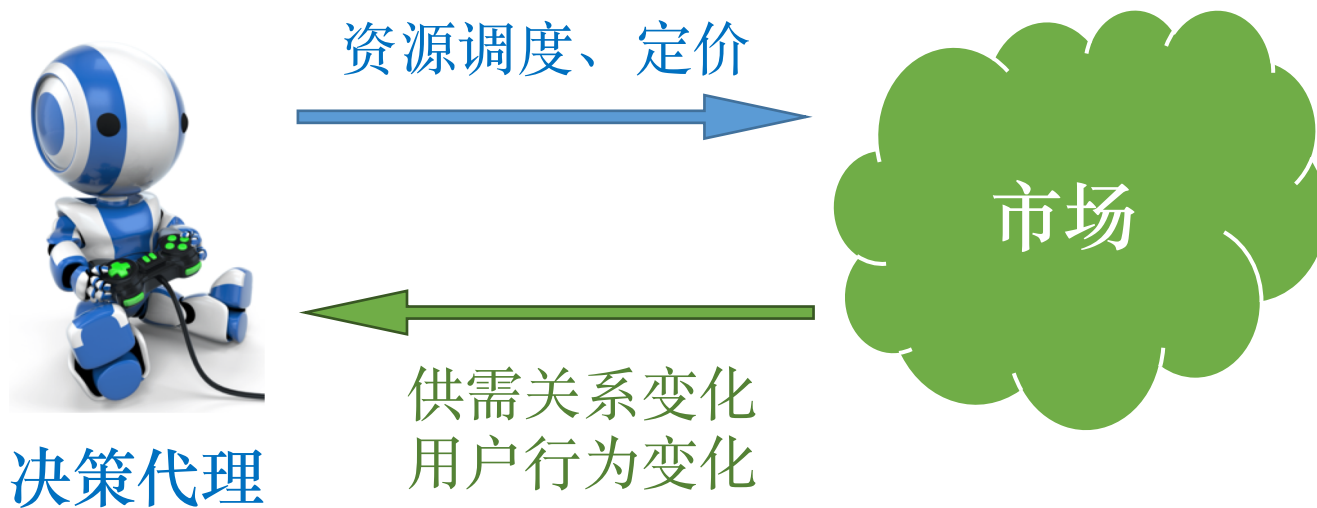


桌游AI设计，如**阿瓦隆**

J. Serrino et al, “Finding Friend and Foe in Multi-Agent Games”, NeurIPS 2019

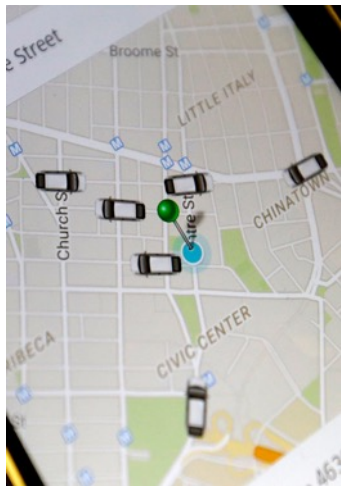
2. 数字经济

滴滴、百度等互联网公司关于资源调度、定价等的决策智能



2. 数字经济

共享出行



车辆调度
动态定价

广告位拍卖



动态拍卖底价

云服务器



动态定价

电力服务



动态定价

电子商务



商户排序

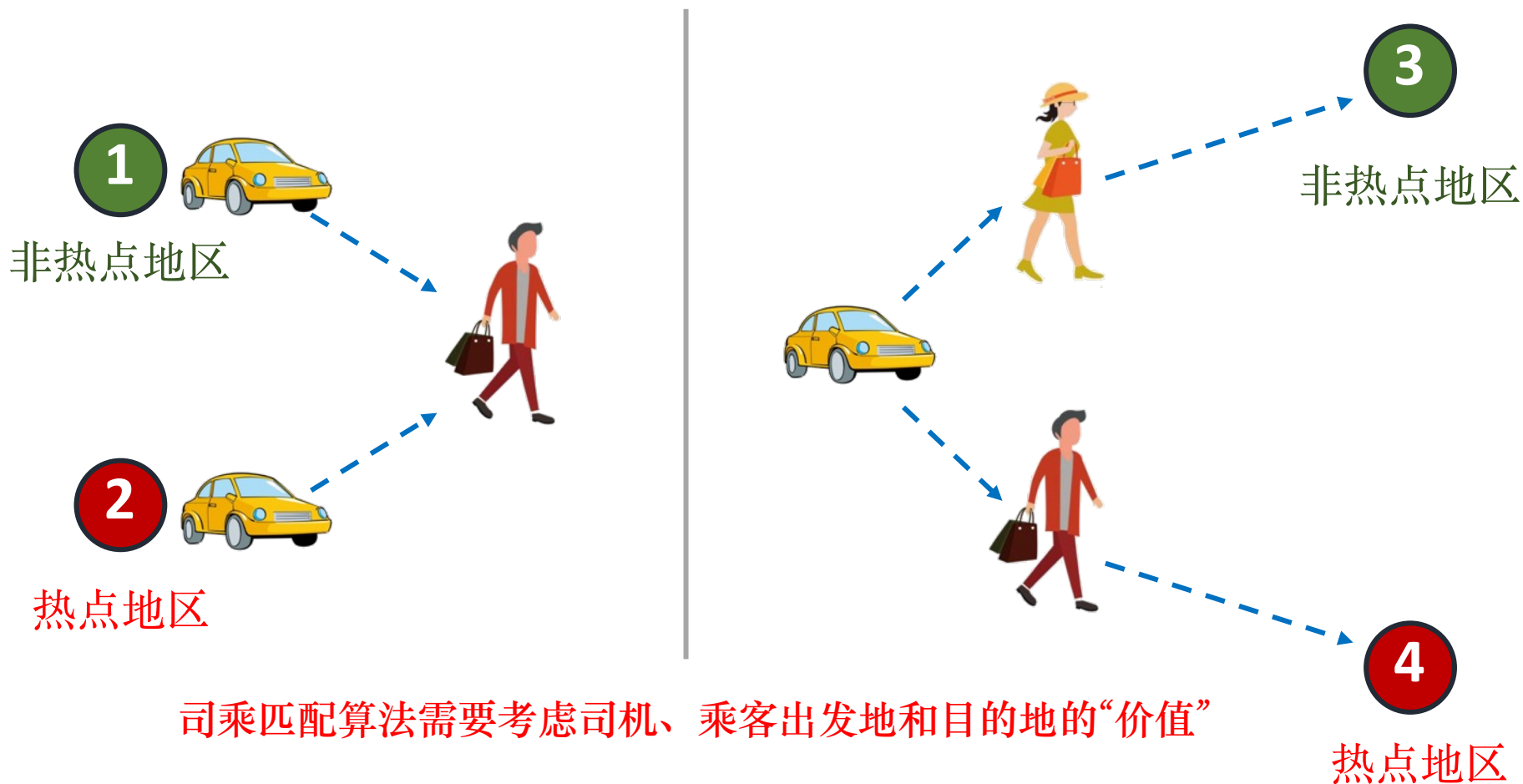
2. 数字经济—共享出行

共享出行平台决策问题：

- 司乘匹配
- 动态定价
- 空车调度
- 拼车算法
- 路径规划
- 出勤激励

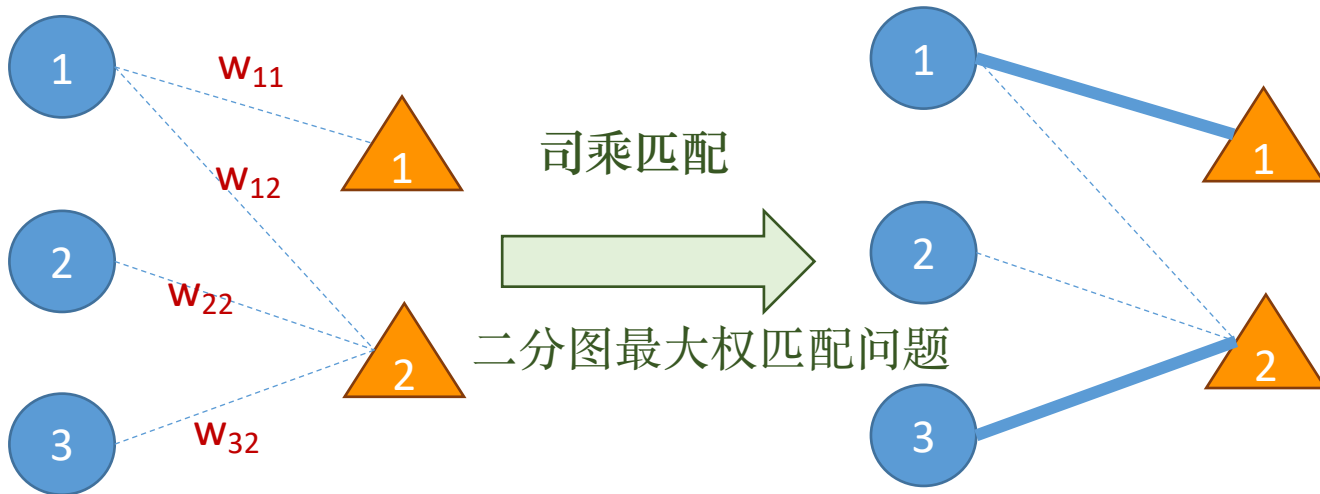
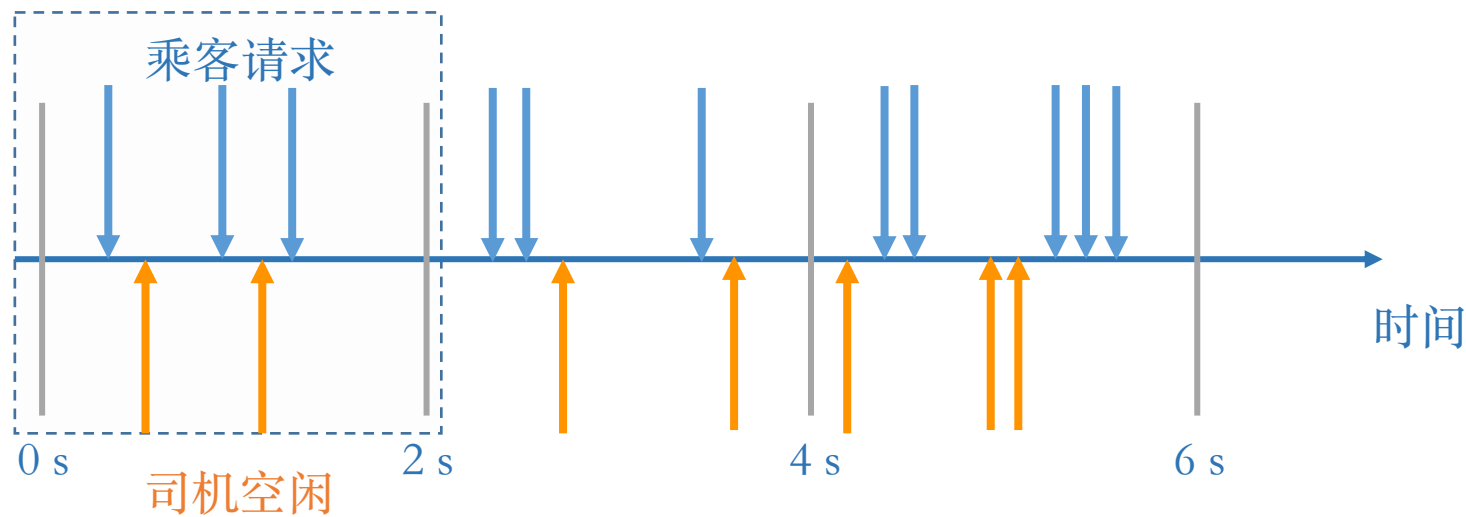


2. 数字经济—共享出行

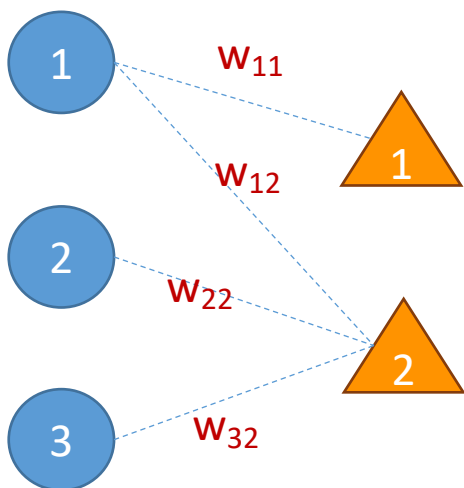


本节内容源自秦志伟博士报告“深度强化学习在网约车交易市场中的应用”

2. 数字经济—共享出行

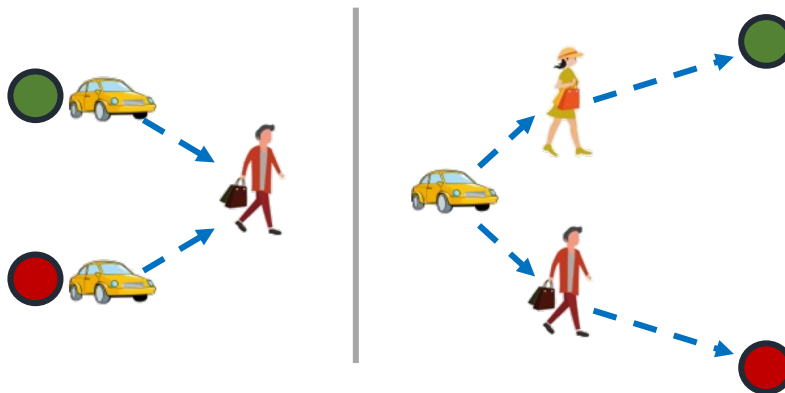


2. 数字经济—共享出行



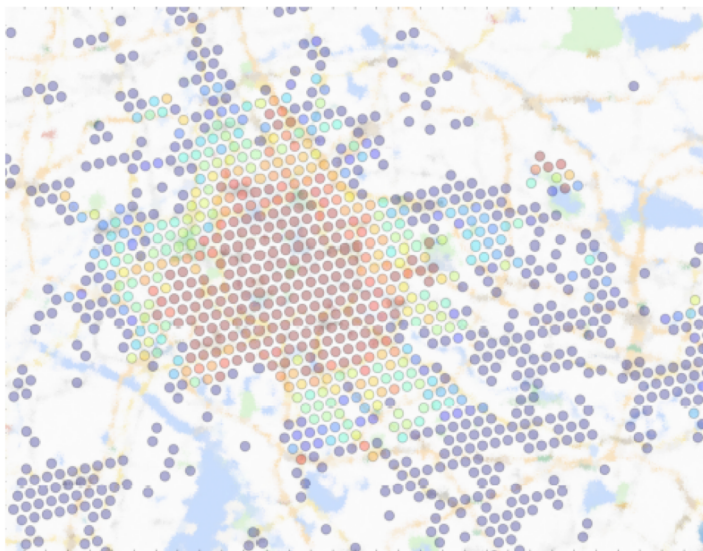
如何定义每一对乘客-司机的权重 W_{ij}

$$W_{ij} = \text{订单金额} + \text{乘客}i\text{目的地价值} - \text{司机}j\text{出发地价值}$$

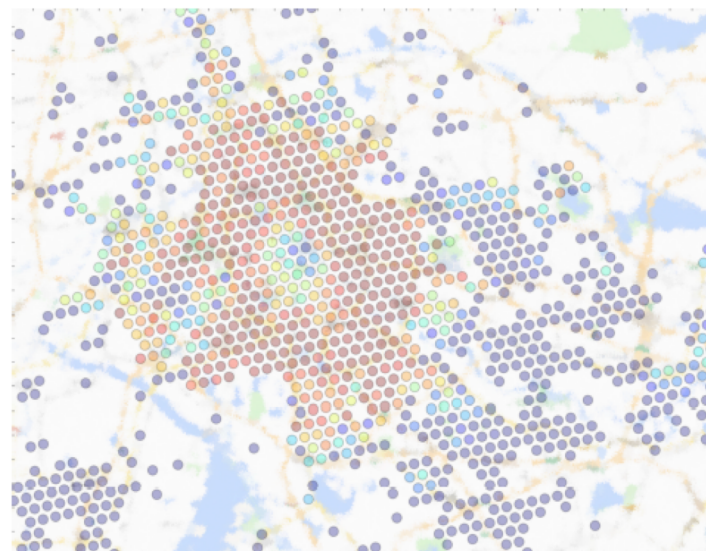


2. 数字经济—共享出行

$$W_{ij} = \text{订单金额} + \text{乘客}i\text{目的地价值} - \text{司机}j\text{出发地价值}$$



(a) 18:00-18:10, evening peak



(b) 09:00-09:10, after morning peak

具体方法可见Z Xu, et al., “Large-scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach,” KDD 2018.

2. 数字经济



互联网公司实时对资源调度、对资源定价



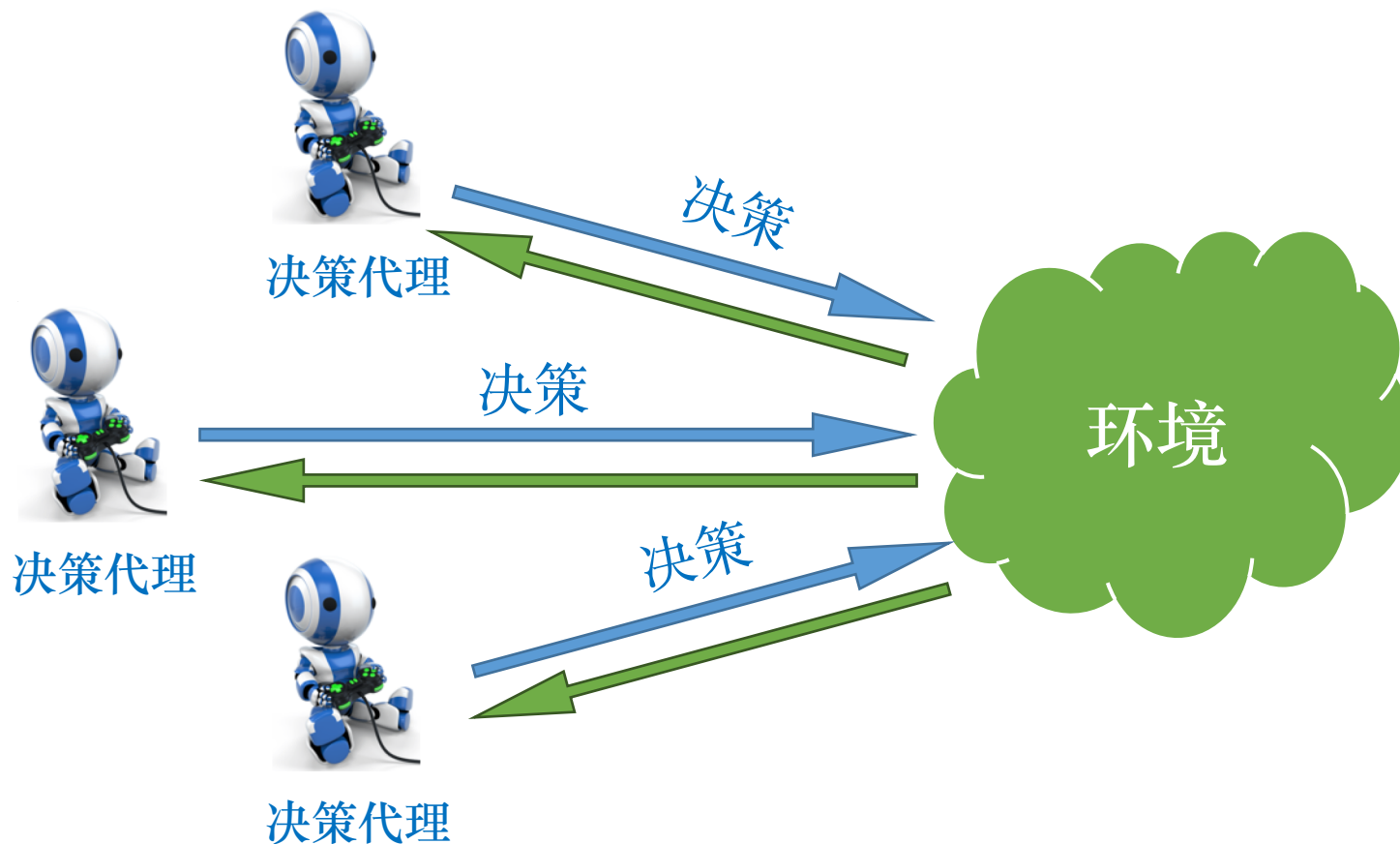
策略空间、状态空间大，环境动态变化、未知信息多



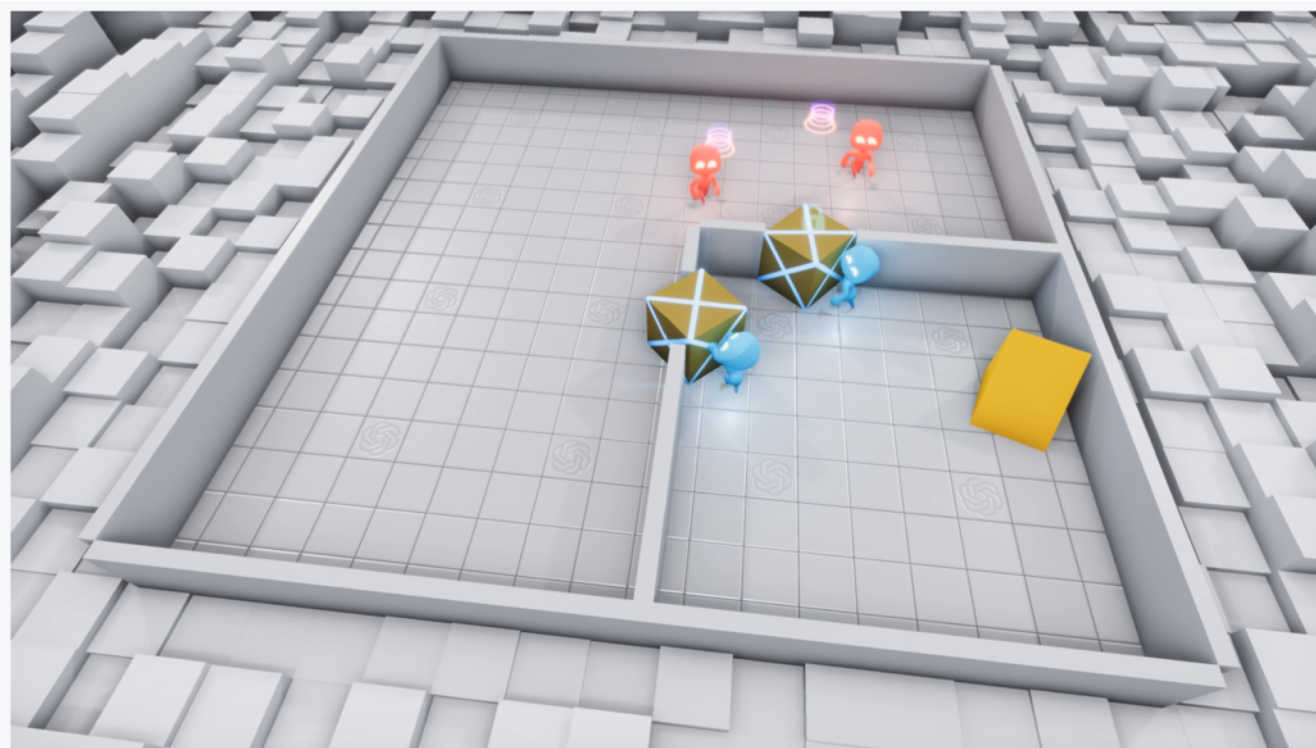
需要对市场价格-需求曲线、用户行为进行预测

3. 多智能体协作

多个智能体分别学习环境、分别决策、实现整体最优
(多智能体强化学习 Multi-Agent Reinforcement Learning)



3. 多智能体协作—捉迷藏AI

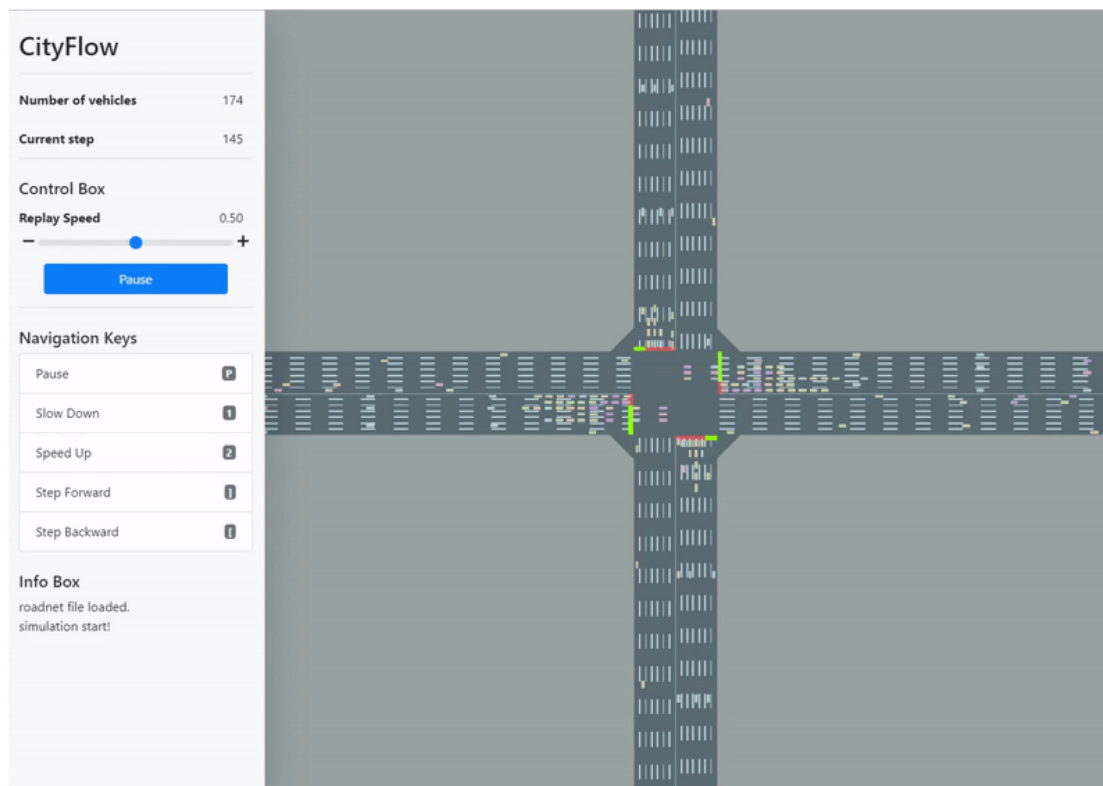


[OpenAI宣传视频](#)

[OpenAI项目主页](#)

更多信息可见吴翼博士报告“Curriculum Evolution and Emergent Complexity with Multi-Agent Reinforcement Learning”

3. 多智能体协作—交通灯控制



[交通灯控制说明视频](#)

[CityFlow项目主页](#)

更多信息可见张伟楠博士报告“大规模智能体强化学习：算法与系统”

3. 多智能体协作



对算力要求更高、算法收敛难度更大



其它应用场景：游戏AI、军事AI



决策智能与感知智能结合



资料介绍及经验分享

AI研究动态阅读渠道



微信公众号：AI科技评论



、机器之心



CCF官网（学术报告视频）



B站官方号（学术报告视频）

AI学习、竞赛渠道



网课平台：B站、Coursera



Andrew Ng, 《Machine Learning》



Andrew Ng et al., 《Deep Learning Specialization》



李宏毅, 《Machine Learning》（中文）



大数据竞赛平台：天池大数据竞赛、DataFountain、DataCastle

关于研究—书籍推荐



豆瓣阅读
优质类型小说平台

分类 搜索

原创

图书 > 文学 > “研”磨记

「研」磨记
一名博士生的回忆录

by
Philip J. Guo
Nalatieler 慕尼黑 地质妞
王辰宇 史庭训 译

The Ph.D. Grind
译言·古登堡计划

“研”磨记 英文原版名 **The Ph.D. Grind**
一名博士生的回忆录

作者 [美] Philip J. Guo
译者 王辰宇 史庭训 Nalatieler 慕尼黑 地质妞
类别 出版 / 非虚构
提供方 译言·古登堡计划
字数 约 67,000 字

★★★★★ 9.0 732 人评价

免费阅读

斯坦福大学计算机博士Philip Guo写的一本关于他读博经历的小书

关于本科阶段接触科研



需要主动搜索信息、主动思考、主动规划

——研究生学习和中学学习最大区别在于由被动变主动、由短期回馈变长期回馈，本科是过渡阶段

谢谢!

