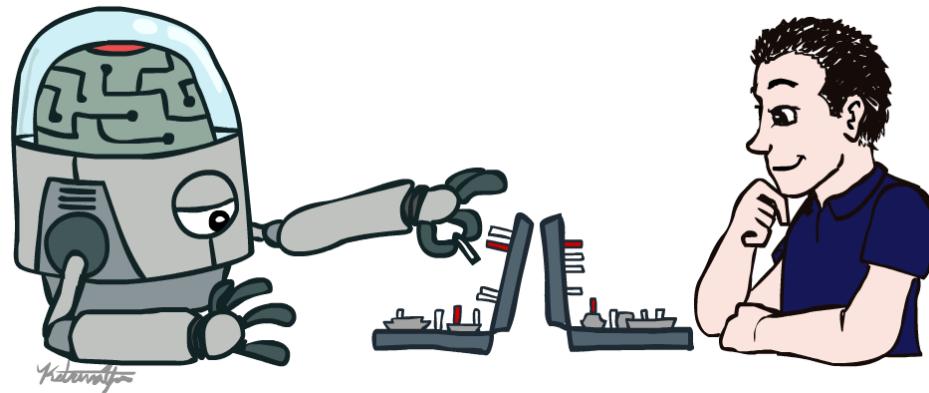


计算机科学选修课： 人工智能导论

齐琦



课程信息

- 课程网页: qiqi789.github.io/teaching/AI/

Qi Qi Web portal

Home English page Research 教课 Teaching Contact

人工智能 (2014年秋季学期)

更新通知

课程信息

上课时间地点
第1, 3至17周, 每周一, 5-7节课; 地点在 4-511.

课程形式:
以下安排可能会有变化。

- 讲课: 总共16周, 每周3个课次
- 课堂参与: 回答问题, 考勤。
- 课后作业: 6次
- 课堂测验: 3至5次
- 考试: 1个期中, 和1个期末考试。

评分标准:
以下安排可能会有变化。

课后作业	25%
课堂测验	15%
期中考试	20%
期末考试	40%
总共	100%

准则规范
作业都应该通过自己的思考和努力完成, 作业抄袭、考试作弊都不允许。作业晚交将被扣掉该次作业总分数的15%。

课本参考:

- 人工智能技术导论 (第三版), 廉师友, 西安电子科技大学出版社, 2007。
- Artificial Intelligence: A Modern Approach, 3rd edition, Stuart J. Russell and Peter Norvig. Prentice Hall, 2009.

课程安排

随课程进展, 以下安排可能会有变化。

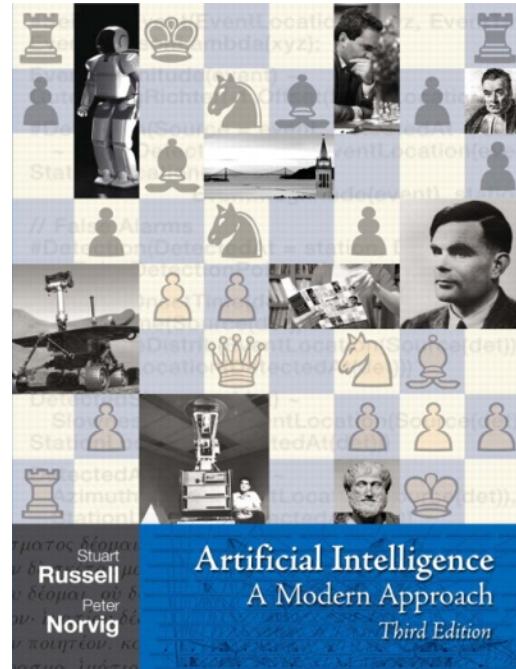
周期	日期	内容	幻灯片	作业	注释
1	9/1	人工智能简介; 课程内容介绍			
2	9/8	中秋节, 没课			
3	9/15	智能行为体和环境; 基础搜索法			
4	9/22	A*搜索和启发式法; 局部搜索; 基于搜索的智能行为体			
5	9/29	博弈; 约束满足问题			

课程信息

- 课后笔答作业：~6次 (25%)
 - 会在网页上贴出，包括完成周期等。
 - 要求独立完成。
- 课堂测验： 3~5次 (15%)
 - 20-30分钟每次，目的检查学习掌握程度。
- 期中考试1次 (20%)，期末考试1次 (40%)
- 课堂考勤和回答问题可以帮助提分
- 学校纪律准则

课本参考

- 人工智能技术导论（第三版），廉师友，西安电子科技大学出版社, 2007。
- Artificial Intelligence: A Modern Approach, 3rd edition, Stuart J. Russell and Peter Norvig. Prentice Hall, 2009.



今天内容

- 什么是人工智能?
- 人工智能发展的历史
- 当今发展的状况、应用
- 未来发展

人工智能是什么？



人工智能是什么？

- 计算机科学与工程中的一个科学研究领域
- 研究目标：不仅是理解，而且是去建造智能体
- 兴起于二战结束后；“Artificial Intelligence”名词出现于1956
- 仍在前沿发展中，许多未解决问题

人工智能研究角度

思考过程，
推理角度

像人一样去
思考

合理地思考

行为角度

像人一样去
行动

合理地行动

和人表现比较
(经验学科，对
人行为的观察和
假设)

以理想化的表现
为尺度
(涉及数学和工
程的组合)

“像人一样去思考”（对认知建模）

- 需要理解人如何思考
- 交叉学科：认知科学（Cognitive science），心理学，神经系统科学
 - 计算机模型 ← 人工智能
 - 实验技术 ← 心理学
 - 研究人的心智（大脑）
 - 依赖于对人或动物的实验
 - 人工智能实验对象主要是计算机

“像人一样去行动”（图灵测试）

- 图灵测试 (Turing Test, by Alan Turing, 1950)
 - 判断一个计算机是否具有智能性
 - A computer passes the test if a human interrogator, after posing some written questions, cannot tell whether the written responses come from a person or from a computer.
- 计算机需具备这些能力，以通过测试
 - 自然语言处理
 - 知识表达（储存）
 - 自动推理（用已知推新结论）
 - 机器学习（在新环境中学习）
- 全面图灵测试 (Total Turing Test, by Harnad, 1991)
 - “The candidate must be able to do, in the real world of objects and people, everything that real people can do”
 - 计算机视觉，语音识别，机器人技术

“合理地思考”（思维法则）

- 希腊哲学家，亚里士多德
 - 三段论，演绎推理 (syllogism)
- 逻辑学：研究思维的法则
- 用逻辑规则建造智能系统首先面临的阻碍
 - 自然界中许多知识很难用逻辑表示法来表述；有时知识不是100%确定的
 - 原则理论上可以解决的问题，在实际中未必能轻松做到
- 很难对思维进行编码；最终关键的不是如何思考，而是如何行动。

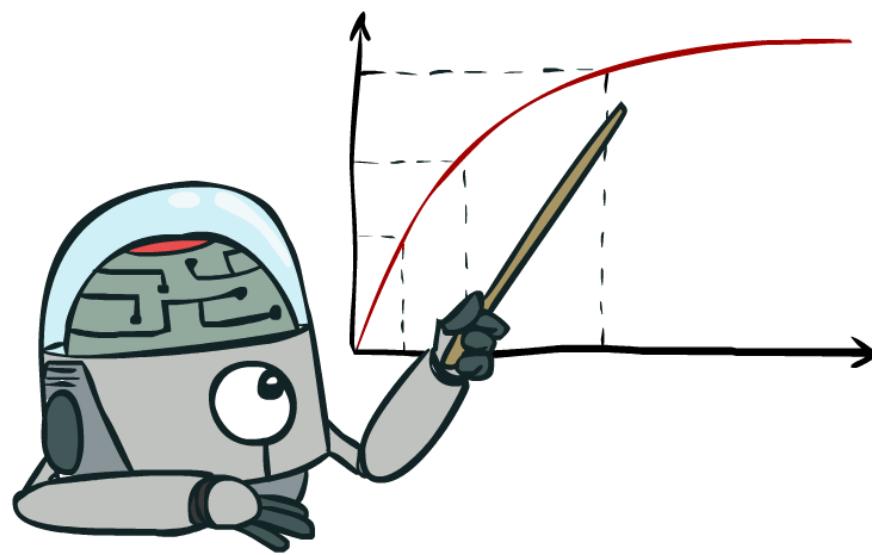
“合理地行动”

- 智能体 (agent)
 - 独立行动，感知环境，适应变化，追求目标
- 合理行动的智能体 (rational agent)
 - 去努力追求最好的结果，或最好的期望结果
- 合理性 (rationality) 不光都是指需要进行逻辑意义上正确的推理后才能得到的
 - 例如有些反射行为
- 基于合理性智能体方式进行人工智能研究的优势
 - 更具普遍意义(不仅只是依赖于逻辑推理)
 - 更便于实际开发 (数学计算)
 - 完美合理性 (总是做对的事) 在复杂情况下难以实现 (计算量要求)

合理的决策

- 合理的：在给定信息情况下，最大化的达到预先制定的目标
- 这里的合理性，主要关注决策本身（要做的行动），而不是决策形成背后的思维过程。
- 用行动结果的功效(utility)来衡量目标
- 合理的决策，是指最大化行动所能带来结果的功效。
 - 计算化的合理性

- 最大化行动结果的功效期望值



人工智能之前

- 哲学（从亚里士多德开始）
- 数学（逻辑，概率）
- 神经系统科学（神经元）
- 经济学（合理性，博弈论）
- 控制原理（反馈）
- 心理学（学习、认知模型）

人工智能发展简史

- 1940-50：早期
 - 布尔逻辑电路模拟大脑， McCulloch & Pitts 1943
 - 1950, Alan Turing, 发表计算机器和智能文章，介绍了图灵测试，机器学习，基因算法，增强式学习。
- 1950-70：人工智能领域的诞生和兴起
 - 1956 夏在Dartmouth， John McCarthy在一个会议上提出人工智能领域。
 - Newell和Simon编制了一个推理程序，可证明一些数学定理。进而推出了另一个普遍问题解决器(General Problem Solver)。
 - 其他的程序用于证明几何定理，或下跳棋等。
 - 1958， McCarthy在MIT定义了高级编程语言Lisp。
 - 1960s， 神经网络也同时发展。
 - 也面临问题：语言翻译还需要背景知识；复杂问题求解；计算复杂度

人工智能发展简史

- 1970-80：知识推理系统
 - 利用领域特定知识帮助推理，专家系统
 - 知识表达和推理用编程语言，Prolog, Planner等
- 1980--：人工智能开始在工业界应用
 - 1982，第一个商用专家系统R1，帮助配置计算机系统配件
 - 1988-93，专家系统在工商业应用中不能达到预期成果，“人工智能的冬天”
- 1986--：神经网络的再次兴起
 - 误差反向传播算法；多层神经网络

人工智能发展简史

- 1987--: 拥抱其他领域的科学成果
 - 概率方法再次兴起，专注于不确定条件下的推理
 - 例如语音识别，利用隐式马可夫模型（数学理论）
 - 贝叶斯网络（Judea Pearl, 1988）主导不确定推理和专家系统
 - 统计学，机器学习，数据挖掘
- 2001--: 大数据时代
 - 从大量数据中学习模式规律
 - 数据量增大，算法结果可以提高
 - 从数据中学习知识 vs 事先输入所有知识

人工智能当今能做到的（应用）

- 无人驾驶汽车
 - STANLEY, DARPA沙漠大奖赛, 2005
 - CMU's BOSS, 城市挑战赛, 2006
 - Google's
- 语音识别
 - 自动语音识别；文字到语音合成
 - 航空公司电话自动订票；计算机语音指令
- 自主规划和调度
 - 宇航器自动控制

人工智能当今能做到的（应用）

- 博弈比赛
 - IBM's 深蓝(DEEP BLUE)击败象棋冠军Kasparov, 1997
 - 特制硬件，每秒分析200万步棋局
 - IBM's 华生(Watson)赢得Jeopardy知识问答比赛的冠军, 2011
- 垃圾电子邮件自动过滤
 - 邮件中有80%-90%是垃圾邮件
- 后勤规划
 - 1991, 波斯湾危机, 美国军方使用一个动态分析和规划工具协调后勤车辆、物资、人员的输送调配（数量50,000次）。规划时间从以前的几周减到几小时。

人工智能当今能做到的（应用）

- 机器人
 - iRobot公司已卖出2百万Roomba真空吸尘器机器人；还有用于美军的PackBot机器人，用于处置危险品和排爆等任务。
- 机器自动翻译语言
 - 计算机程序自动把阿拉伯语翻译成英语；利用统计方法和机器学习算法。
 - 网页搜索
 - 文件分类，邮件过滤等

人工智能机器人视频播放

- 机器人足球比赛
- 机器人叠衣服
- 行走机器人
- 无人驾驶汽车
- 仓库物流机器人
- 无人送货飞行器

嵌入应用

- 人工智能也存在于其他许多有用的系统里：
 - 调度系统，例如航班路线，军队调度
 - 路线规划，例如Google地图
 - 医疗诊断
 - 自动监控系统
 - 互联网搜索引擎
 - 垃圾邮件分类
 - 电话自动答复工台
 - 信用卡欺诈侦测
 - 在线商店商品推荐
 - 等，非常多！

人工智能的未来

- 人工智能研究的目的
 - 创造越来越智能的系统
 - 更好地理解人类智能
 - 扩大智能带给人类的益处

人工智能的未来

- 人工智能的发展在加速，部分原因是工业界应用的竞赛
- 性能一旦达到应用的最低标准后，每提高1%都能带来数十亿的回报
 - 语音识别
 - 文字理解
 - 物体识别
 - 自动驾驶车辆
 - 民用机器人

人工智能的未来—超级人工智能

- 超越人的人工智能有可能实现
 - 量子计算机可大幅提高计算能力
 - 但并不是一个直接的目标
- 既有益处，也有忧虑
 - 所有人类文明都来自于智能
 - 有可能超越人类理解和控制（电影“Transcendence”提示的）
- 超级人工智能在当今难以实现
 - 计算能力的提升和让一个机器有意识之间没有关系
 - 研究人大脑机制还处在初级阶段
- 引发人工智能科学家预先思考
 - 类似思考存在于其他领域，例如核物理，基因工程。
 - 发挥有利的一面，避免不利的一面。