Artificial Intelligence

Lecture 1： Introduction

Xiaojin Gong

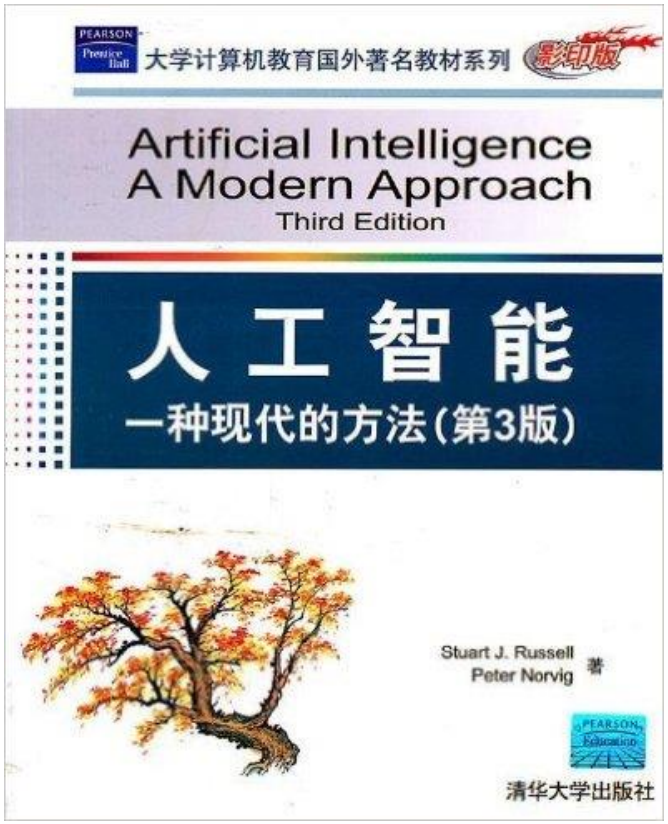
2022-02-21

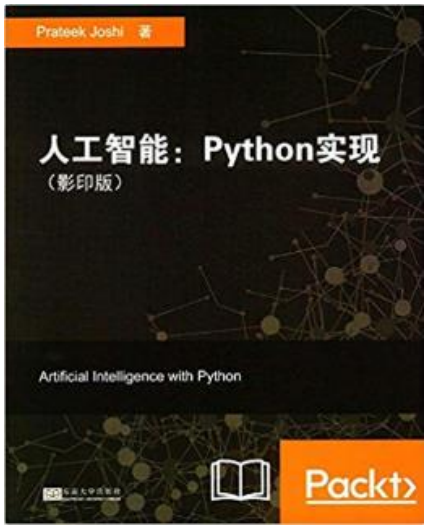
Outline

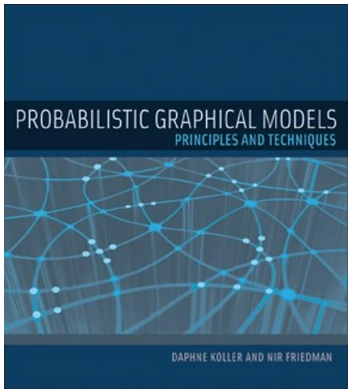
- Course Overview
- What is AI?
- AI History
- Course Topics
- Intelligent Agents

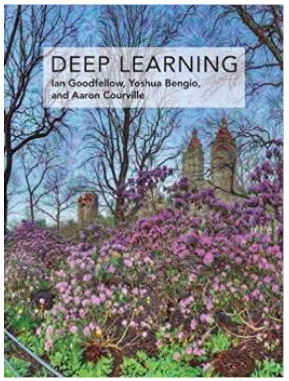
Course Overview

- Course Website:
  - 学在浙大 http://course.zju.edu.cn/
- Textbook: http://aima.cs.berkeley.edu/index.html



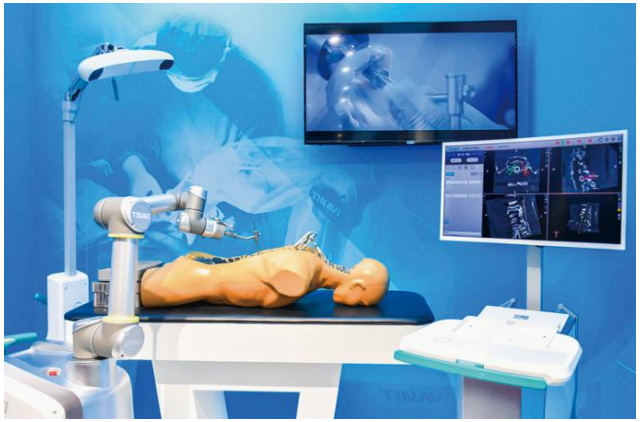






What is AI?

全球首个获得公民身份的机器人
2017年 沙特索菲亚

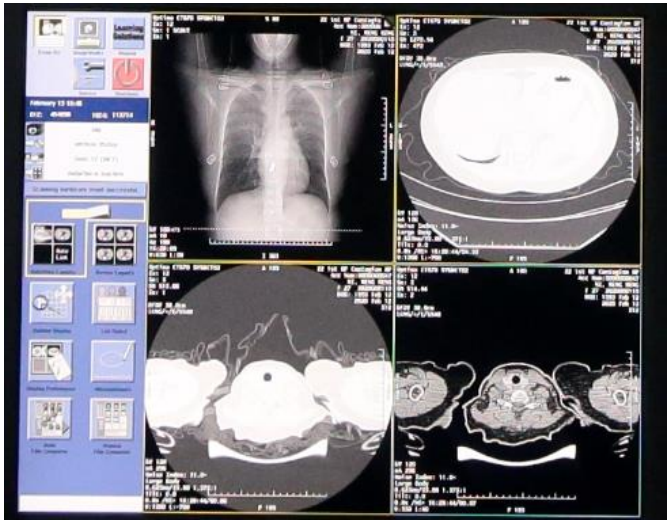


国产骨科手术机器人“天玑”
2018年

无人车

元宇宙

What is AI?



中科天玑 “智疫通”

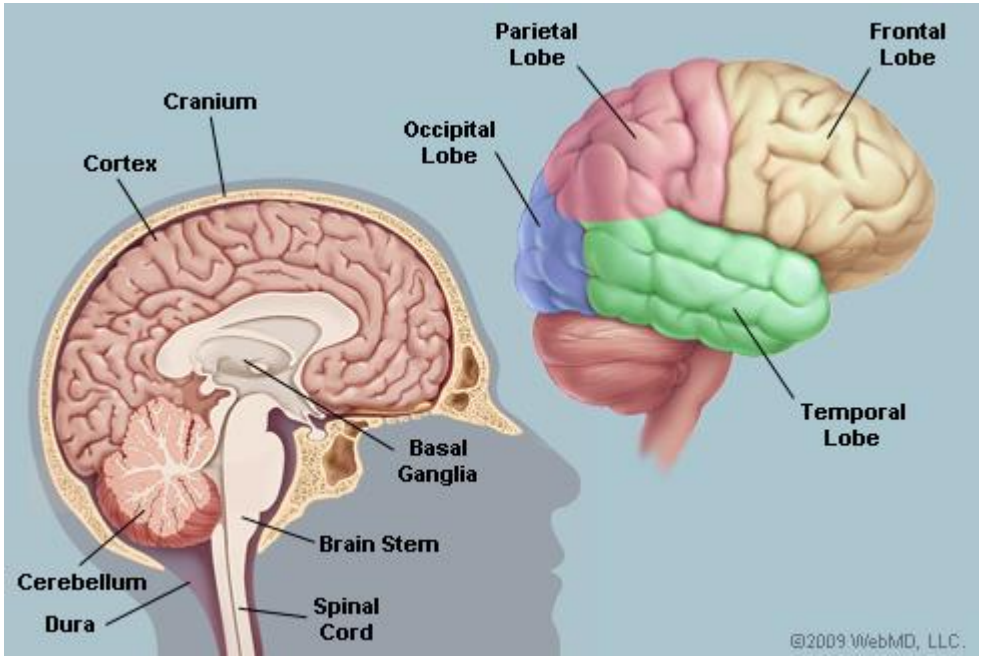
阿里达摩院AI诊断系统

What is AI?

Thinking Humanly	Thinking Rationally
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

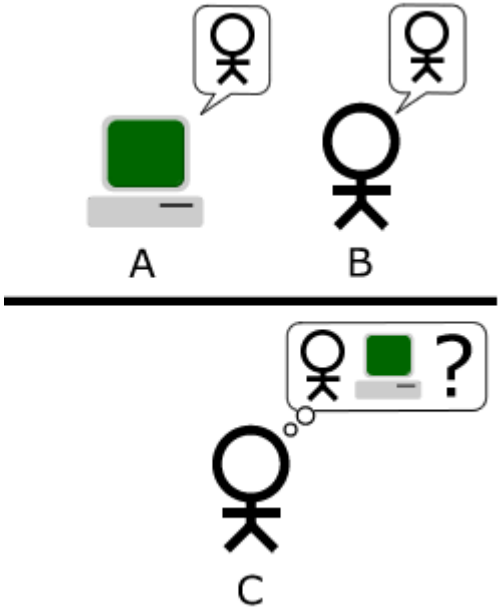
Thinking Humanly

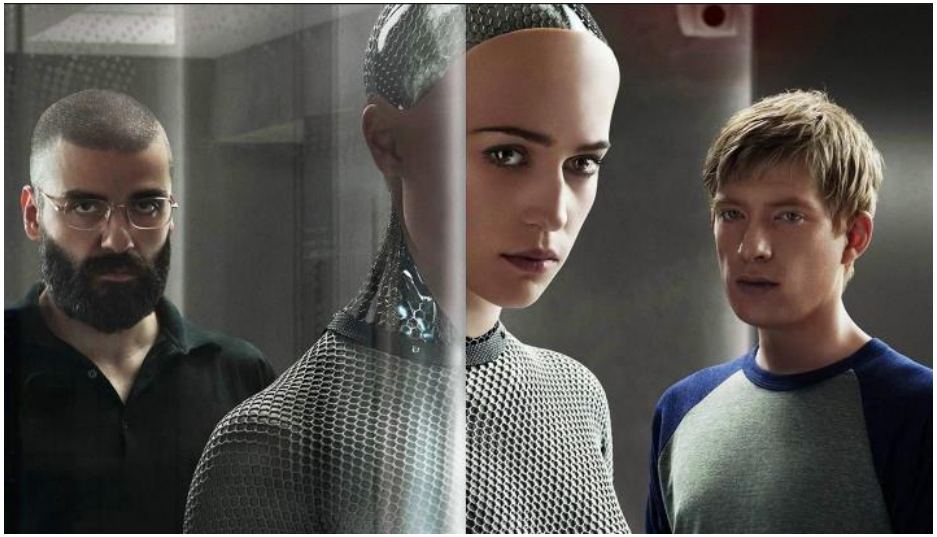
- Cognitive science
  - Introspection
  - Psychological experiments
  - Brain imaging



Acting Humanly

- Turing test: (Alan Turing, 1950)
  - Natural language processing
  - Knowledge representation
  - Automated reasoning
  - Machine learning
  - Computer vision
  - Robotics

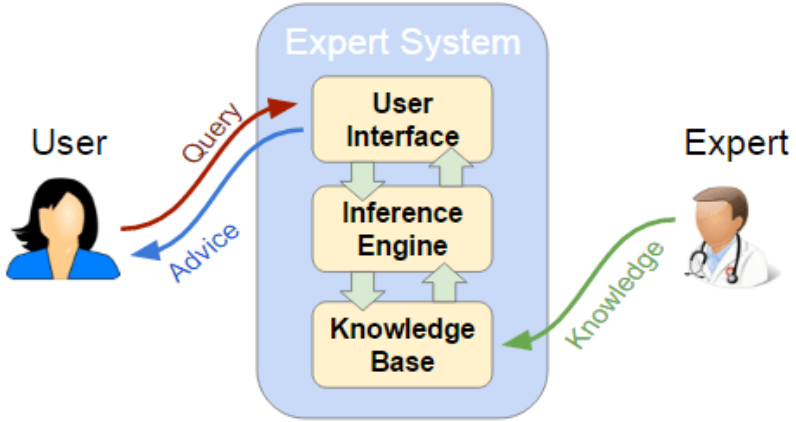


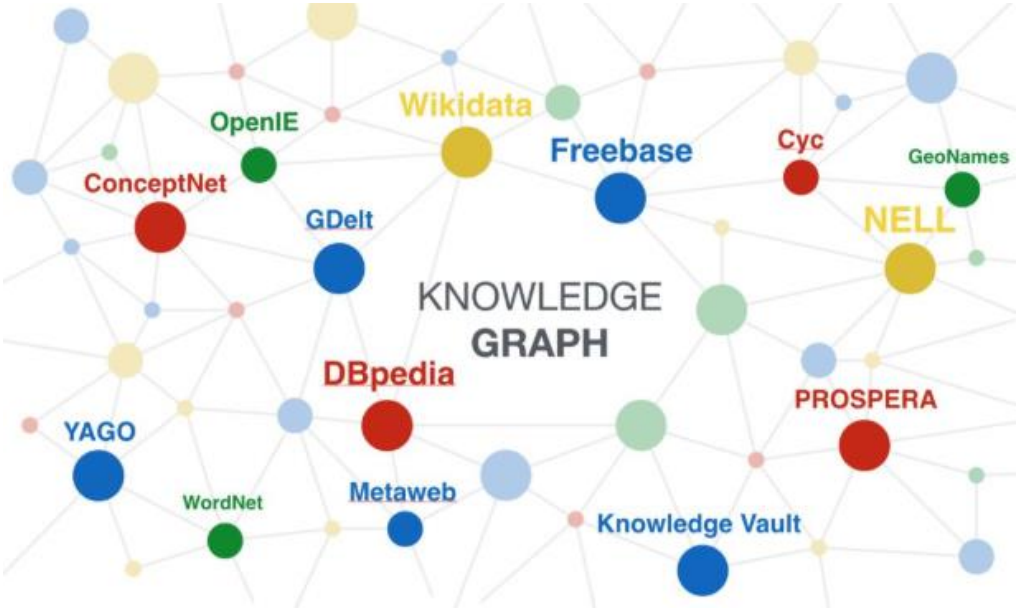


Ex Machina (2015)

Thinking Rationally

- Logic
  - Aristotle: Syllogisms



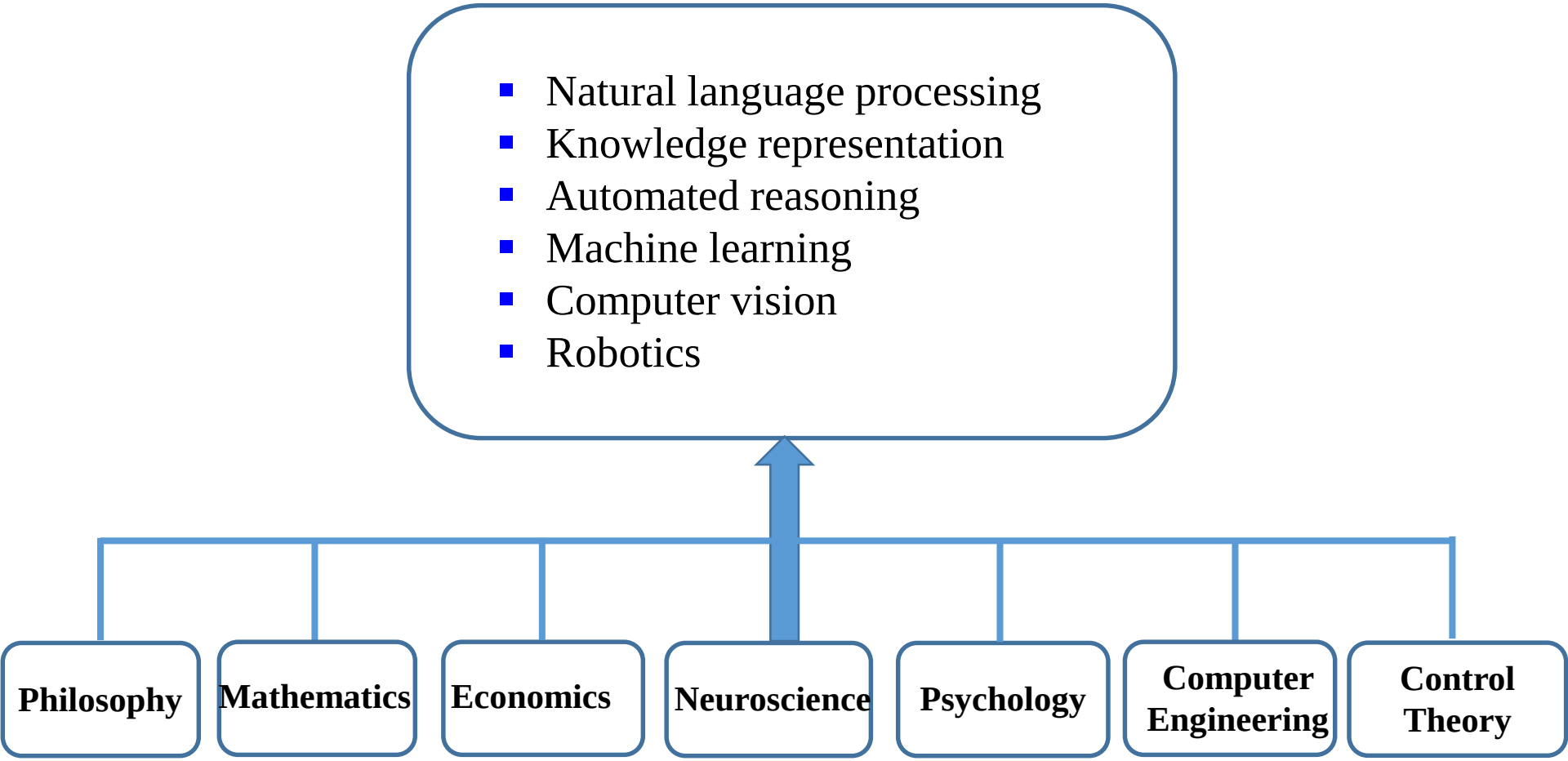


Acting Rationally

- Rational agent
  - Acts so as to achieve the best outcome, or when there is uncertainty, the best expected outcome.

This course concentrates on general principles of rational agents and on components for constructing them.

The Foundations of AI



Philosophical Foundations of AI?

- Strong AI
  - Can machines really think?
  - Conscious machines, self-awareness
  - Little progress, little serious activity, controversial
- Weak AI
  - Can machines act intelligently?
  - Focused on specific tasks
  - Get machines to do things that people can do which machines currently cannot
  - Most AI research on this topic
  - Lots of progress recently

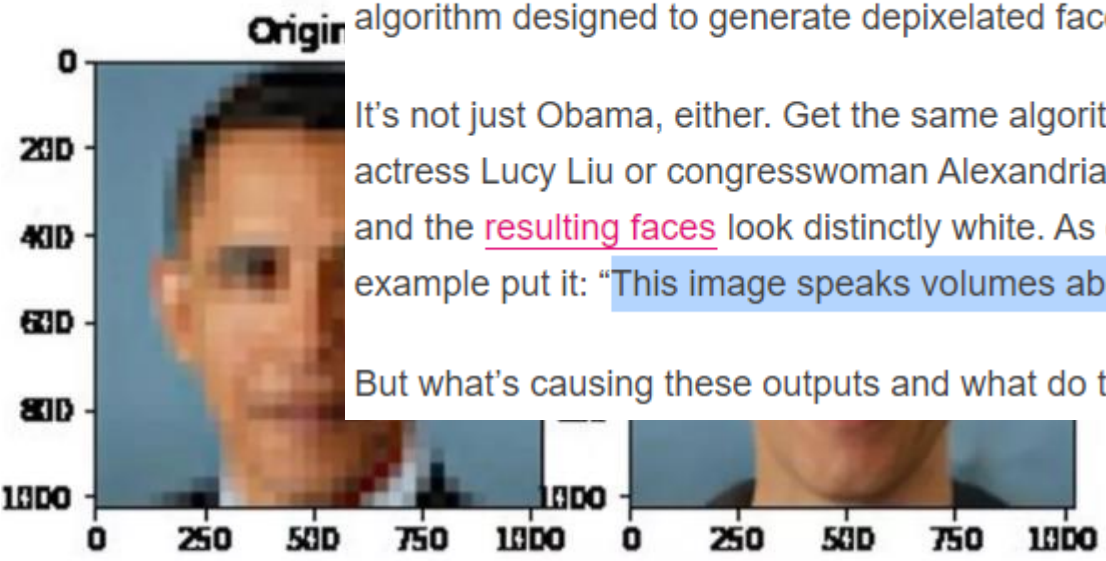
Philosophical Foundations of AI?

- AI Safety, Ethics, and Policy

Bill Gates and Alicia Vikander in Ex Machina, a film that looks at how AI affects humans

Philosophical Foundations of AI?

- AI Safety, Ethics, and Policy
  - Fairness, bias, inequality
  - Privacy and National Security
  - Rights and moral consideration for AI and robots?
  - Ethics and AI: teaching machines to be moral?



It's a startling image that illustrates the deep-rooted biases of AI research. Input a low-resolution picture of Barack Obama, the first black president of the United States, into an algorithm designed to generate depixelated faces, and the output is a white man.

It's not just Obama, either. Get the same algorithm to generate high-resolution images of actress Lucy Liu or congresswoman Alexandria Ocasio-Cortez from low-resolution inputs, and the resulting faces look distinctly white. As one popular tweet quoting the Obama example put it: "This image speaks volumes about the dangers of bias in AI."

But what's causing these outputs and what do they really tell us about AI bias?

Philosophical Foundations of AI?

- AI Safety, Ethics, and
  - Fairness, bias, inequali
  - Privacy and National S
  - Rights and moral consi
  - Ethics and AI: teaching

杭州市智能网联汽车开放测试道路公示

根据杭州市经济和信息化局、杭州市公安局、杭州市交通运输局《关于印发杭州市智能网联车辆道路测试管理实施细则（试行）的通知》（杭经信联推进〔2018〕121号）要求，经第三方测试机构提供道路环境勘察报告，专家评估，杭州市自动驾驶测试管理联席工作小组研究，现就杭州市余杭区未来科技城文二西路及周边道路开放用于智能网联汽车测试路段进行公示。

一、道路范围

1. 文二西路（西：绿汀路，东：高教路）
2. 葛巷路（西：绿汀路，东：景兴路）
3. 绿汀路（北：文一西路，南：文二西路）
4. 创明路（北：葛巷路，南：文二西路）
5. 景兴路（北：葛巷路，南：文二西路）

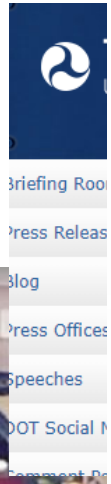
二、测试时间

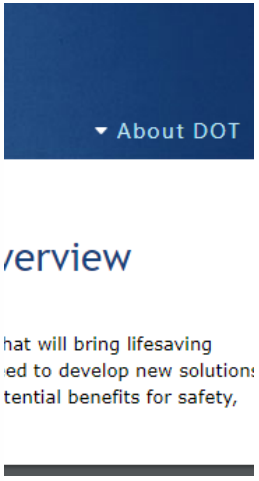
许可测试时间为10:00-16:00、21:00-00:00。且无天气预报等级为暴雨、暴雪、能见度低于100米的大雾等极端天气情况。

公示期为：2019年5月28日到2019年6月3日。

杭州市自动驾驶测试管理联席工作小组

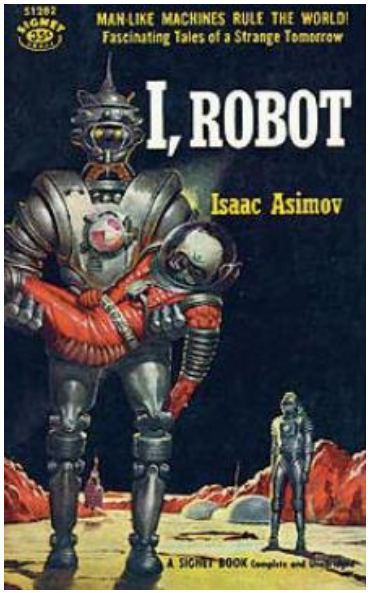
2019年5月28日





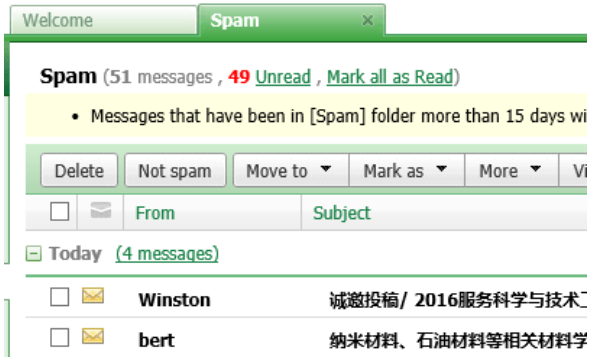
Philosophical Foundations of AI?

- AI Safety, Ethics, and Policy
  - Isaac Asimov‘s “Three Laws of Robotics”
    - A robot may not injure a human being or, through inaction, allow a human being to come to harm.
    - A robot must obey orders given it by human beings except where such orders would conflict with the First Law.
    - A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.



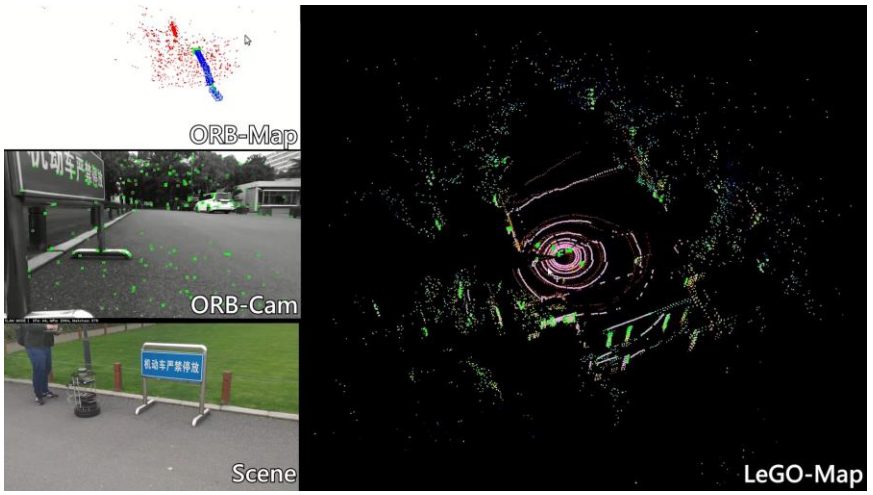
Natural Language Processing

- Automatic speech recognition
- Text-to-speech synthesis
- Dialog systems
- Machine translation
- Spam filtering
- …



Computer Vision

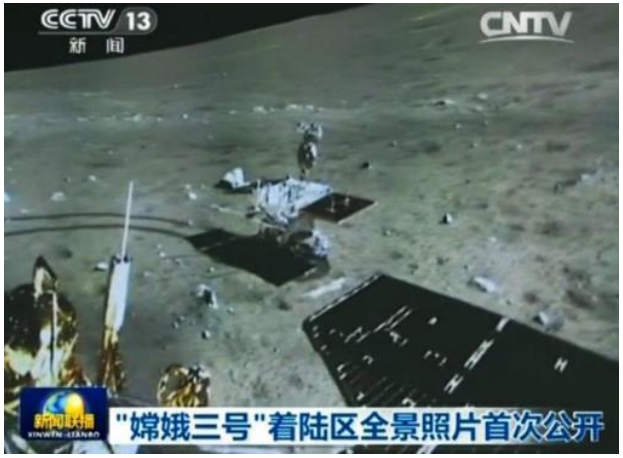
- Object detection/recognition
- Scene understanding
- Image classification
- Pose recognition
- …

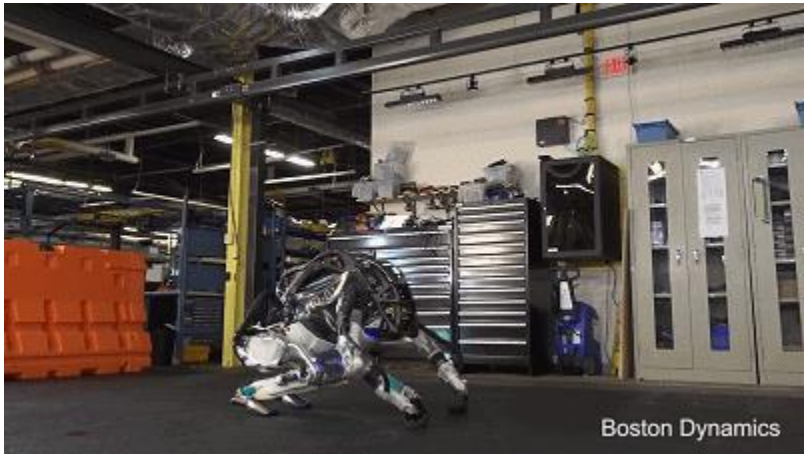


Robotics

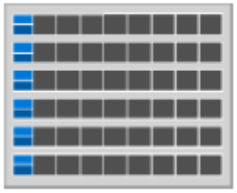
- Autonomous driving
- Unmanned ground vehicles
- Autonomous underwater vehicles
- …

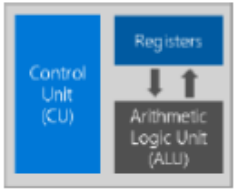
- Environment perception
- Map building
- Path planning



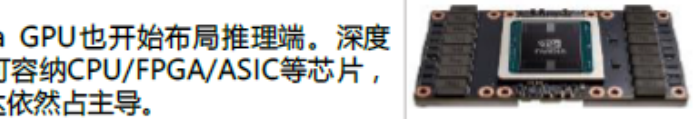


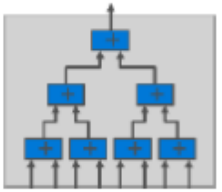
AI Chips

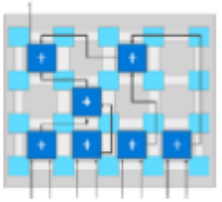


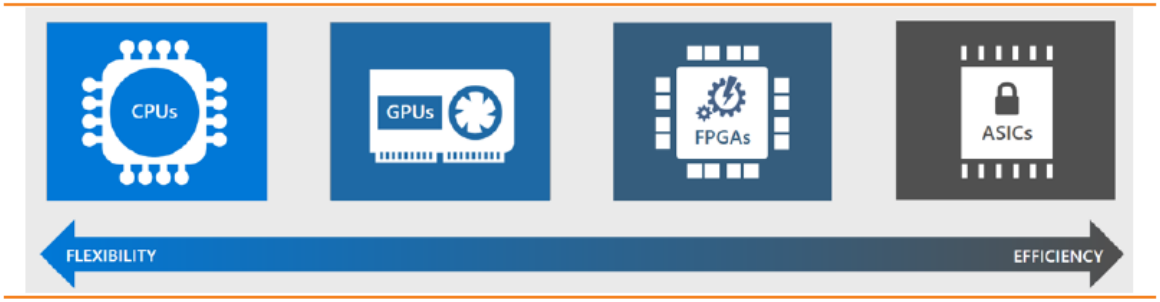


训练端	推理端
GPU：以英伟达为主，AMD为辅标榜通用性，多维计算及大规模并行计算架构契合深度学习的需要。在深度学习上游训练端（主要用在云计算数据中心里），GPU是当仁不让的第一选择。	**GPU**：英伟达Volta GPU也开始布局推理端。深度学习下游推理端虽可容纳CPU/FPGA/ASIC等芯片，但竞争态势中英伟达依然占主导。
ASIC：以谷歌的TPU、英特尔的Nervana Engine为代表，针对特定框架进行深度优化定制。但开发周期较长，通用性较低。比特币挖矿目前使用ASIC专门定制化矿机。	**ASIC**：下游推理端更接近终端应用，需求也更加细分，英伟达的DLA，寒武纪的NPU等逐步面市，将依靠特定优化和效能优势，未来在深度学习领域分一杯羹。
CPU：通用性强，但难以适应于人工智能时代大数据并行计算工作。	**FPGA**：依靠可编程性及电路级别的通用性，适用于开发周期较短的IoT产品、传感器数据预处理工作以及小型开发试错升级迭代阶段等。但较成熟的量产设备多采用ASIC。









Credits: 天风证券研究所

AI Chips


Cambricon
寒 武 纪
首页
产品技术
开放平台
开发者
新闻中心
加入我们
关于我们
云端AI加速卡
为企业级数据中心提供AI计算加速
了 解 更 多
智能处理器IP
采用定制化的低功耗处理器架构，广泛应用于各类智能终端
了 解 更 多
边缘AI加速卡
面向超紧凑/低功耗的边缘端AI推理部署方案
了 解 更 多
Cambricon

The History of AI

A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College
M. L. Minsky, Harvard University
N. Rochester, I.B.M. Corporation
C.E. Shannon, Bell Telephone Laboratories

August 31, 1955

1956 Dartmouth Conference: The Founding Fathers of AI

The History of AI

1956
The term “artificial intelligence” is coined at a Dartmouth conference and AI is founded as an academic discipline.

1956–1974
The golden years of AI er
funding in promising, logi
problem-solving approac

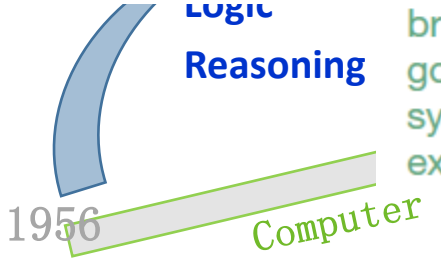
1974–1980
Overly high expectations
the limited capacities of A
leads to the first “AI winte
funding and interest in AI

1980–1987
The rise of knowledge-based e
systems brings new successes
change in the focus of researc
funding toward this form of AI.

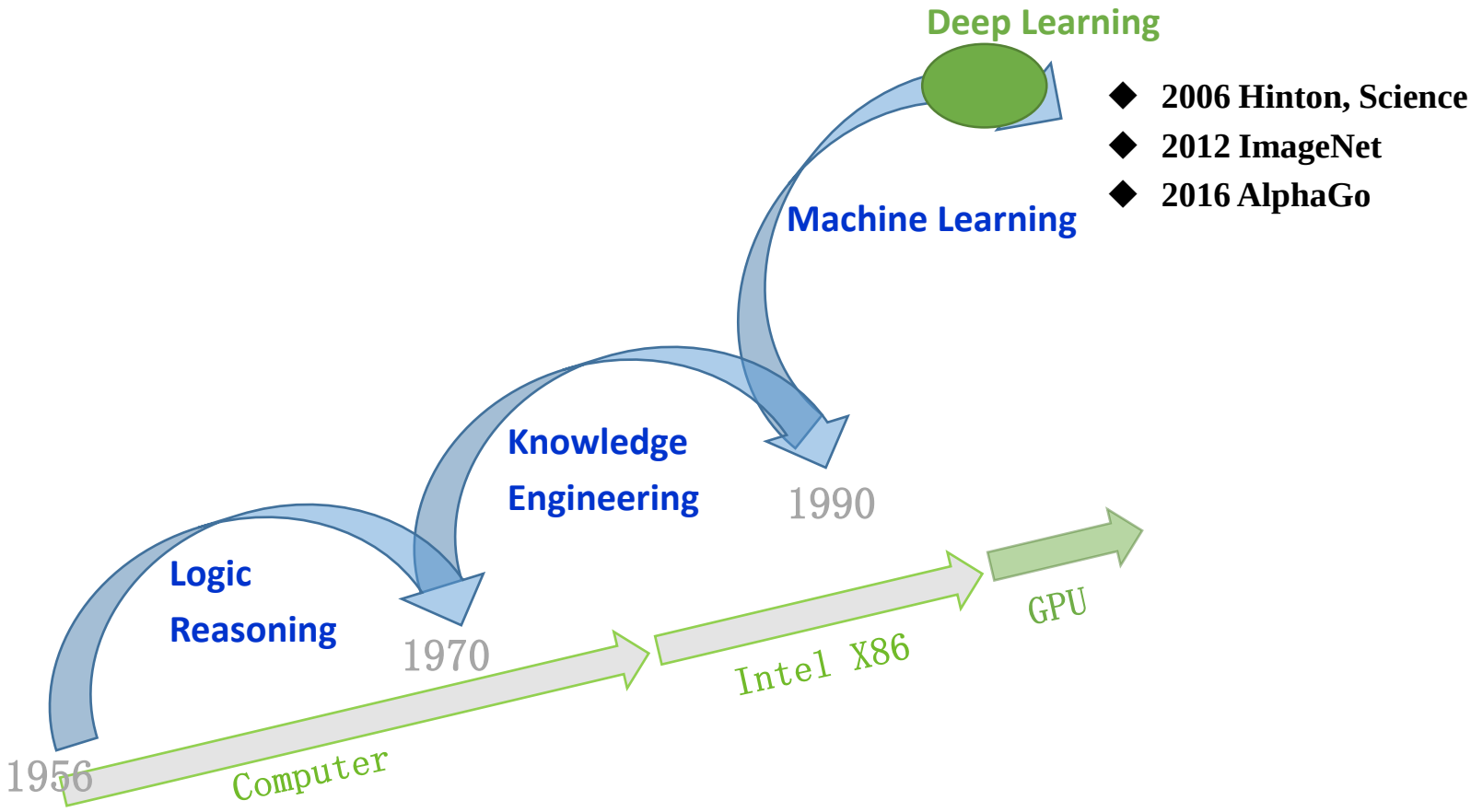
1987–1993
The second “AI winter” starts
sudden collapse of the special
hardware industry in 1987. The
brings with it negative percept
governments and investors, as
systems show their limitations
expensive to update and maint

1993–2011
Optimism about AI returns and increases. New successes are marked with the help of increased computational power and AI becomes data-driven. In 1997, IBM’s DeepBlue beats world champion Kasparov at chess. In 2002, Amazon uses automated systems to provide recommendations. In 2011, Apple releases Siri and IBM Watson beats two human champions at the TV quiz Jeopardy.

2012–today
Increased availability of data, connectedness and computational power allow for breakthroughs in machine learning, mainly in neural networks and deep learning, heralding a new era of increased funding and optimism about the AI potential. In 2012, Google driverless cars navigate autonomously and in 2016 Google AlphaGo beats a world champion in the complicated board game Go.



The History of AI



The Present and Future of AI

http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm

中华人民共和国中央人民政府

（三）战略目标。

分三步走：

第三步，到2030年人工智能理论、技术与应用总体达到世界领先水平，成为世界主要人工智能创新中心，智能经济、智能社会取得明显成效，为跻身创新型国家前列和经济强国奠定重要基础。

——形成较为成熟的新一代人工智能理论与技术体系。在类脑智能、自主智能、混合智能和群体智能等领域取得重大突破，在国际人工智能研究领域具有重要影响，占据人工智能科技制高点。

——人工智能产业竞争力达到国际领先水平。人工智能在生产生活、社会治理、国防建设各方面应用的广度深度极大拓展，形成涵盖核心技术、关键系统、支撑平台和智能应用的完备产业链和高端产业群，人工智能核心产业规模超过1万亿元，带动相关产业规模超过10万亿元。

——形成一批全球领先的人工智能科技创新和人才培养基地，建成更加完善的人工智能法律法规、伦理规范和政策体系。

The Present and Future of AI

中华人民共和国教育部
Ministry of Education of the People's Republic of China

（四）主要目标

到2020年，基本完成适应新一代人工智能发展的高校科技创新体系和学科体系的优化布局，高校在新一代人工智能基础理论和关键技术研究等方面取得新突破，人才培养和科学研究的优势进一步提升，并推动人工智能技术广泛应用。

到2025年，高校在新一代人工智能领域科技创新能力和人才培养质量显著提升，取得一批具有国际重要影响的原创成果，部分理论研究、创新技术与应用示范达到世界领先水平，有效支撑我国产业升级、经济转型和智能社会建设。

到2030年，高校成为建设世界主要人工智能创新中心的核心力量和引领新一代人工智能发展的人才高地，为我国跻身创新型国家前列提供科技支撑和人才保障。

Course Topics

This course concentrates on general principles of rational agents and on components for constructing them.

Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Course Topics

- Designing rational agents
  - An agent is an entity that perceives and acts
  - A rational agent selects actions that maximize its expected utility
  - Characteristics of the percepts, environment, and action space dictate techniques for selecting rational actions

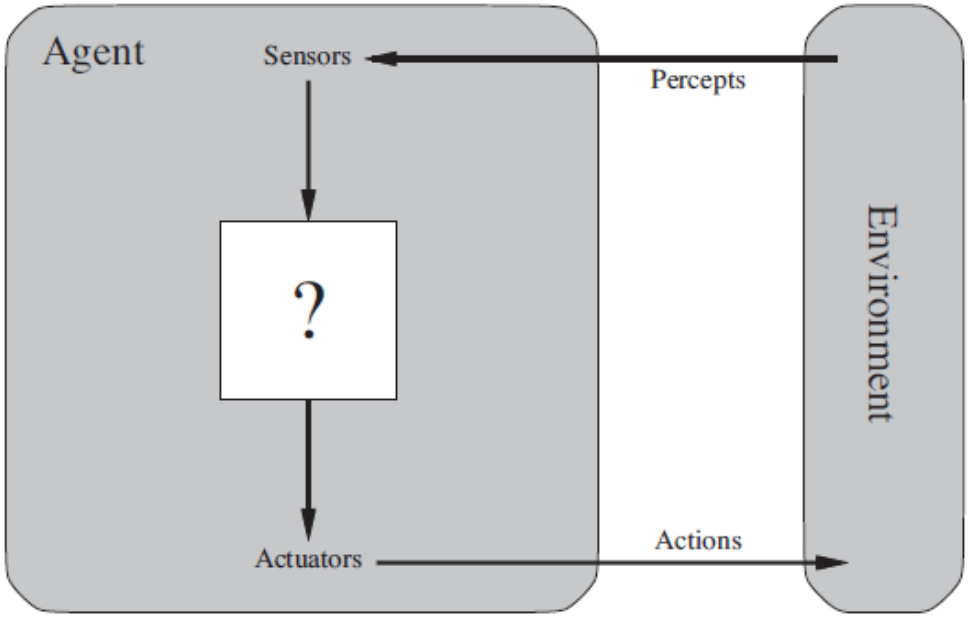
- This course is about
  - General AI techniques for a variety of problems
  - Learning to recognize when and how a new problem can be solved with an existing technique

Course Topics

L1	Feb. 21		Introduction to AI
L2	Feb. 28	Problem -Solving	Search I: Uninformed Search + Heuristic Search
L3	Mar. 7		Search II: Adversarial Search + Monte Carlo Tree Search *Recitation I: Python Tutorial*
L4	Mar. 14		Constraint Satisfaction Problems
L5	Mar. 21	Logic	Logic + Inference
L6	Mar. 28	Uncertain Reasoning	Probability *Recitation II:*
L7	Apr. 4		Bayesian Network I: Representation
L8	Apr. 11		Bayesian Network II: Inference
L9	Apr. 18		Markov Network: Representation + Inference
L10	Apr. 25		Lab *Recitation III:*
L11	May 2	Learning	Learning for PGM
L12	May 9		Classification
L13	May 16		Deep Learning
L14	May 23		Markov Decision Process
L15	May 30		Reinforcement Learning *Recitation VI:*
L16	Jun. 6		*Final Project Presentation*
E1	TBD		Final Exam

Intelligent Agents

- An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators.
- For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has.



Task Environments

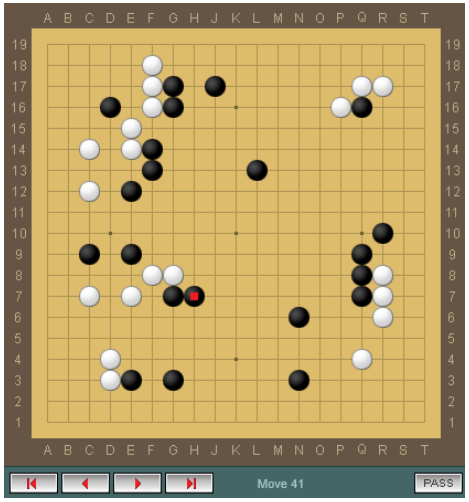
- Performance
- Environment
- Actuators
- Sensors

Agent Type	Performance Measure	Environment	Actuators	Sensors
Taxi driver	Safe, fast, legal, comfortable trip, maximize profits	Roads, other traffic, pedestrians, customers	Steering, accelerator, brake, signal, horn, display	Cameras, sonar, speedometer, GPS, odometer, accelerometer, engine sensors, keyboard

Figure 2.4 PEAS description of the task environment for an automated taxi.

Task Environments

- Fully observable vs. partially observable
- Deterministic vs. stochastic
- Episodic vs. sequential
- Static vs. dynamic
- Discrete vs. continuous
- Known vs. unknown

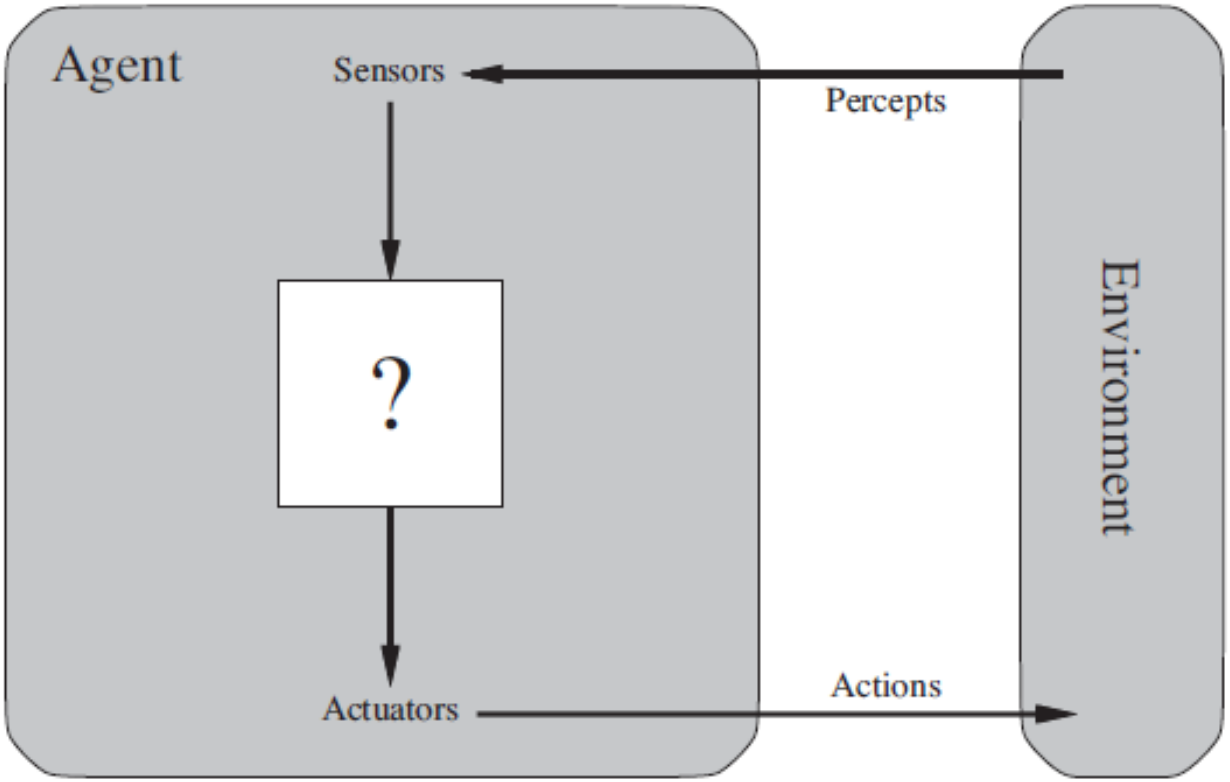


Task Environments

Task Environment	Observable	Agents	Deterministic	Episodic	Static	Discrete
Crossword puzzle	Fully	Single	Deterministic	Sequential	Static	Discrete
Chess with a clock	Fully	Multi	Deterministic	Sequential	Semi	Discrete
Poker	Partially	Multi	Stochastic	Sequential	Static	Discrete
Backgammon	Fully	Multi	Stochastic	Sequential	Static	Discrete
Taxi driving	Partially	Multi	Stochastic	Sequential	Dynamic	Continuous
Medical diagnosis	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
Image analysis	Fully	Single	Deterministic	Episodic	Semi	Continuous
Part-picking robot	Partially	Single	Stochastic	Episodic	Dynamic	Continuous
Refinery controller	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
Interactive English tutor	Partially	Multi	Stochastic	Sequential	Dynamic	Discrete

Figure 2.6 Examples of task environments and their characteristics.

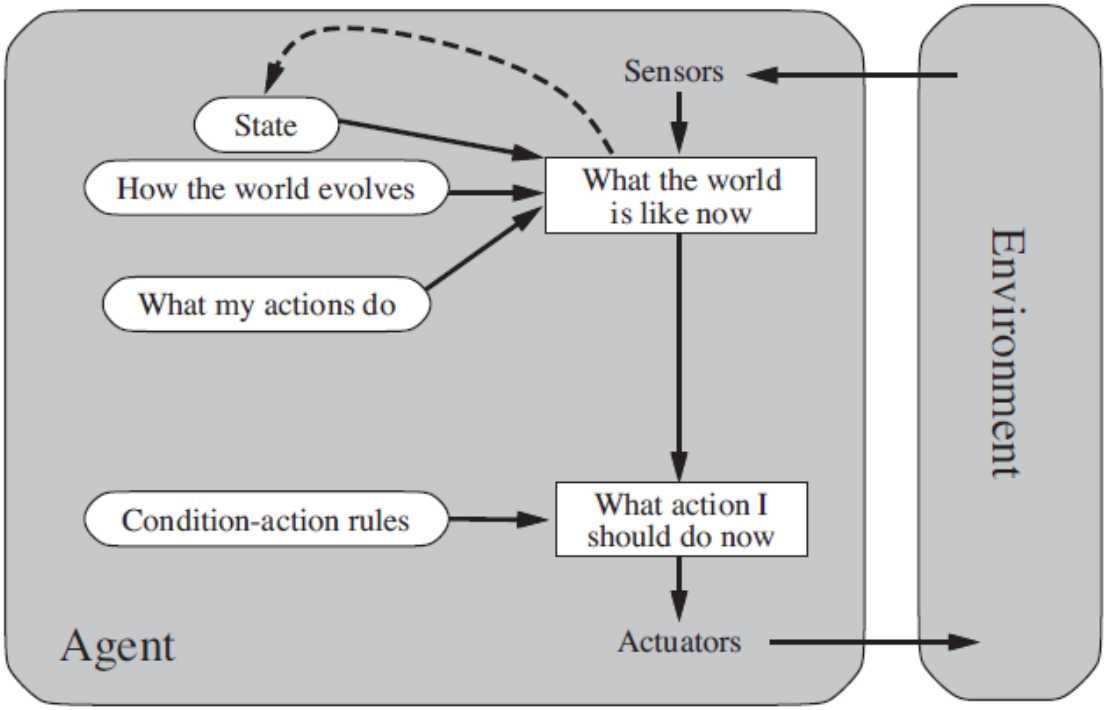
The Structure of Agents



Agent = Architecture + Program

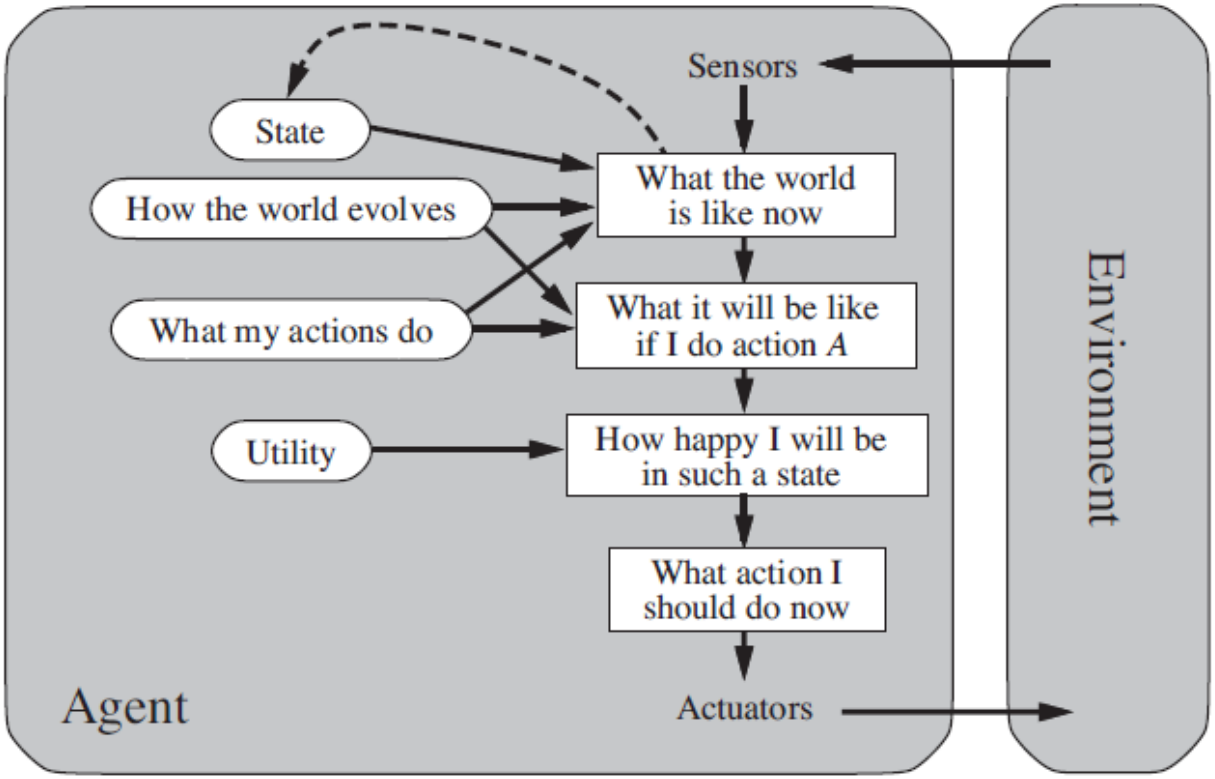
Reflex Agents

- Select action based on current percept (and maybe memory)
- May have memory or a model of the world's current state
- Do not consider the future consequences of their actions
- Consider how the world IS
- Can a reflex agent be rational?



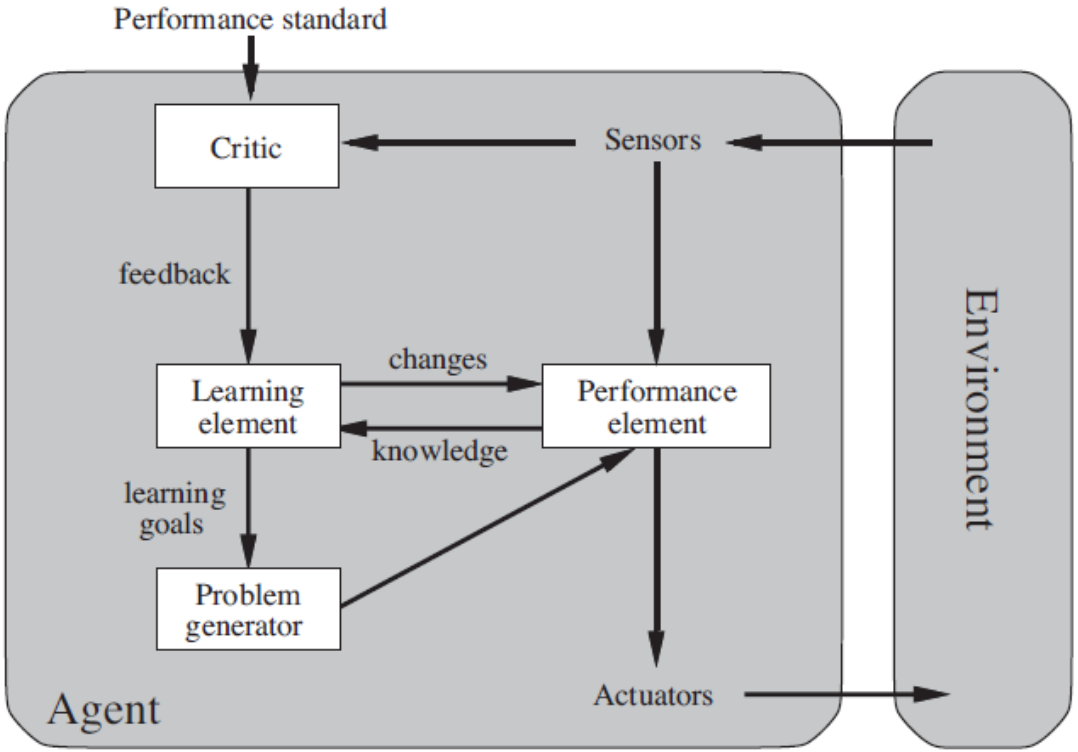
Planning Agents

- Decisions based on (hypothesized) consequences of actions
- Must have a model of how the world evolves in response to actions
- Must formulate a goal / utility
- Consider how the world WOULD BE



Learning Agents

- Learning element: making improvement
- Performance element: selecting external actions
- Problem generator: suggesting actions that will lead to new and informative experiences



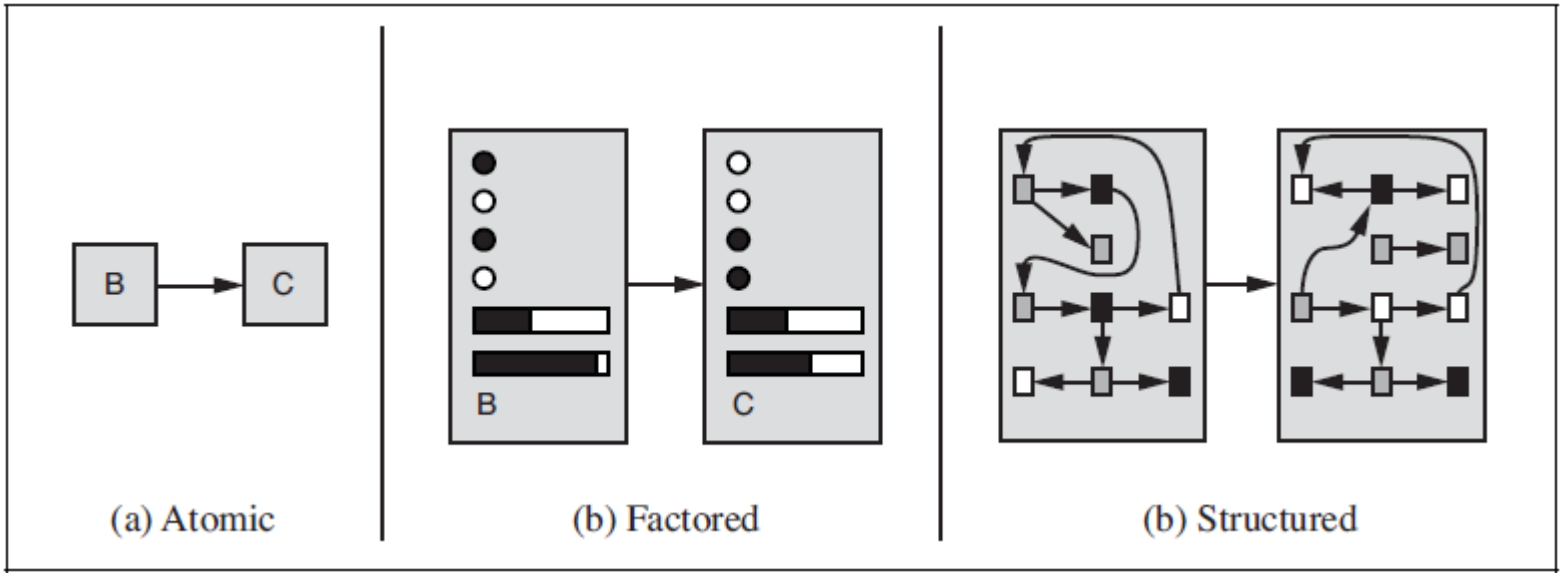
Representations of States

- Rational Agents
  - Representation of Environments

Search
Game-playing
Markov decision processes

Constraint satisfaction
Bayesian network
Machine learning

First-order logic
Knowledge-based learning
Natural language processing



Assignments

- Reading assignment:
  - Ch. 1-2, Ch. 26-27