

人工智能（双语）： **Artificial Intelligence**

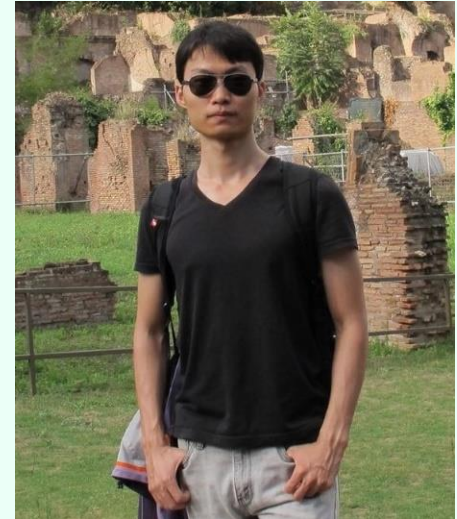
Introduction (Chapter 1)

Instructor: Qiang Yu (于强)

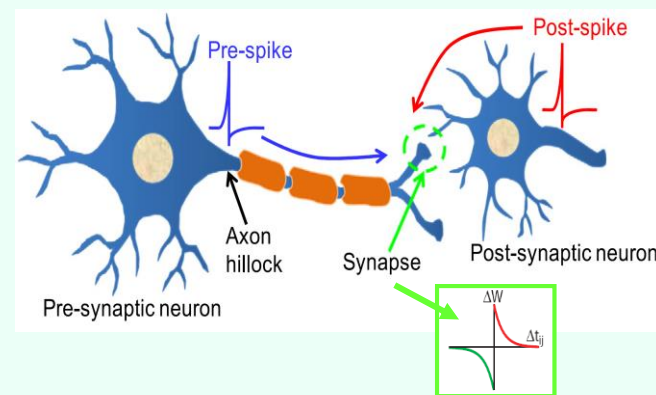
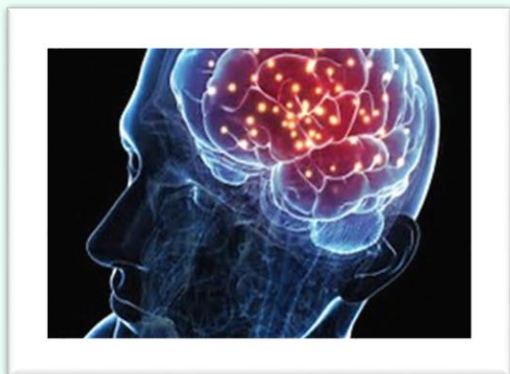
About Me

- 简介-于强

- 天津市海外青年人才，北洋青骨学者
- 研究方向：类脑智能计算、人工智能
- 教育及任职：
 - 哈尔滨工业大学，学士，2006-2010；
 - 新加坡国立大学，博士，2010-2014；
 - 德国马普所，马普学者，2014-2016；
 - 新加坡I2R，项目负责人，2016-2017；
 - 天津大学，2017年底-至今。
- 联系方式：yuqiang@tju.edu.cn



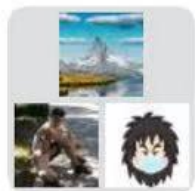
类脑智能计算



- Human Brain is **cognitively very powerful** and yet very **efficient**.
- Deep Learning encountered several bottlenecks.
- Read more:

<https://tns-yuq.github.io/>

班级微信群



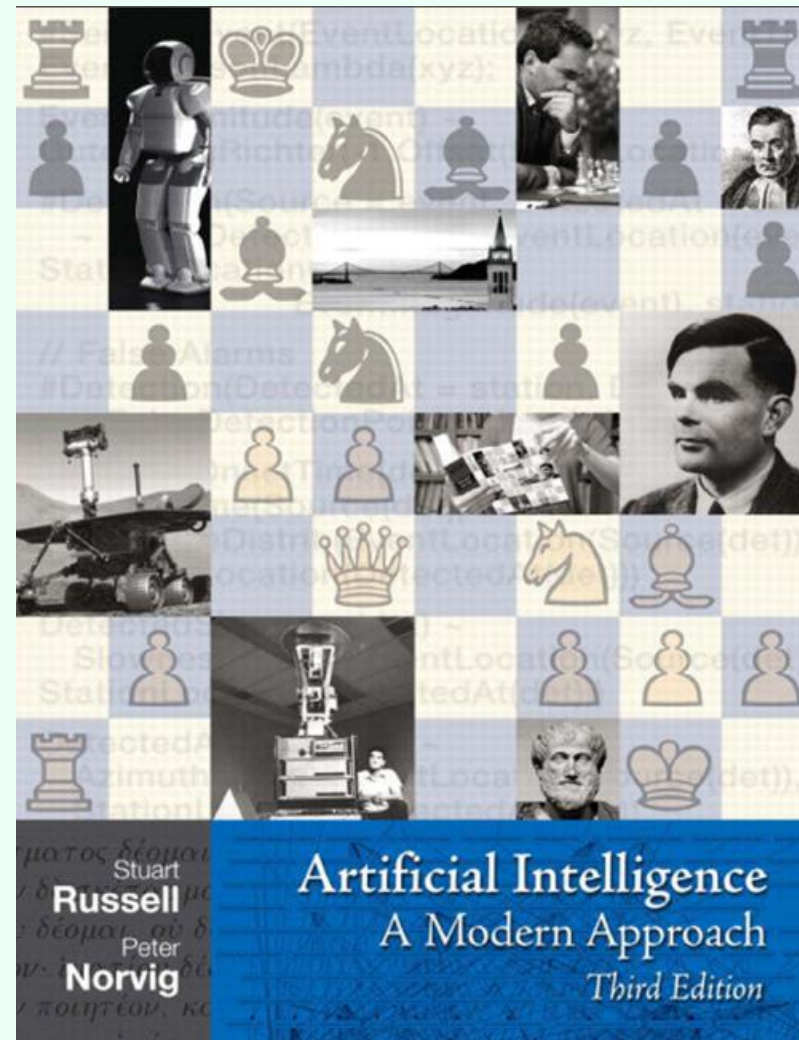
AI双语 212201



该二维码7天内(9月20日前)有效, 重新进入将更新

Basic information about course

- MoWe, Building 55-A308
- Two experiments (4 sessions in total 2 weeks)
- Text Book: Artificial Intelligence: A Modern Approach



Prerequisites

- Comfortable **programming** in general-purpose programming language – we'll do experiments in Python
- Some knowledge of **algorithmic concepts** such as running times of algorithms
- Ideally, some familiarity with **probability** (we will go over this from the beginning but we will cover the basics only briefly)
- Not scared of **mathematics**;

Grading

- Experiments: 40% (20% + 20%)
 - May discuss with another person; experimental reports and codes must be your own
- Final exam: 60%

Some highly visible recent AI successes in games



Watson defeats
Jeopardy champions
(2011)



DeepMind
achieves human-
level performance
on many Atari
games (2015)



AlphaGo defeats Go
champion (2016)



CMU's Libratus defeats
top human poker
players (2017)

Typical picture in news articles



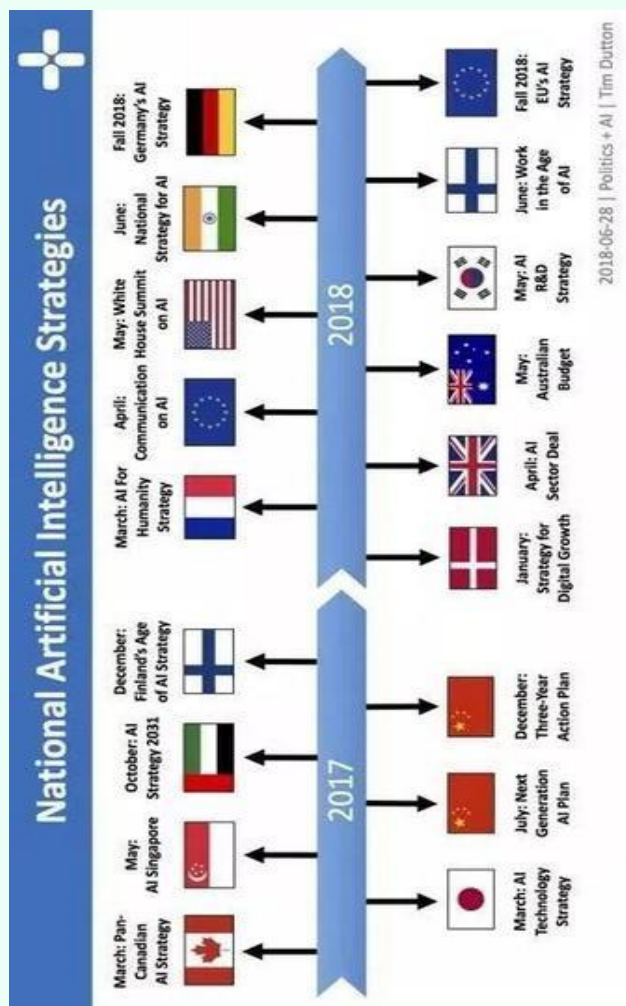
BusinessInsider reporting on the poker match...

Why AI gets so popular...

- 为何人工智能如此“火”
 - 产业变革的拐点
 - 蒸汽时代的蒸汽机
 - 电气时代的发电机
 - 信息时代的计算机和互联网
 - 智能时代的人工智能
 - 战略意义大
 - 世界主要发达国家抢滩布局人工智能
 - 我国也将发展AI提升到国家战略层面-抓住机遇
 - 影响范围广
 - 各行各业，方方面面

Why AI gets so popular...

- AI正在掀起新一轮产业革命与造富运动



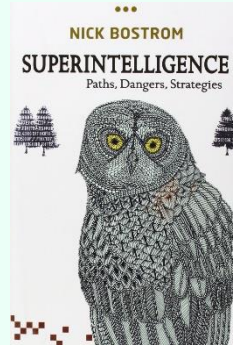
2021年全球公司市值排名，AI相关企业占据9席

- 1、美国苹果公司，2.296万亿美元
- 2、美国微软，1.827万亿美元
- 3、沙特阿美石油公司，1.81万亿美元
- 4、亚马逊，1.688万亿美元
- 5、Alphabet（谷歌母公司），1.416万亿美元
- 6、腾讯控股，9145.89亿美元
- 7、特斯拉，8079.29亿美元
- 8、脸书，7651.57亿美元
- 9、阿里巴巴集团，7188亿美元
- 10、台积电，6626亿美元

Worries about AI - superintelligence



→
writes



→
influences

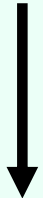


Elon Musk

→
donates to

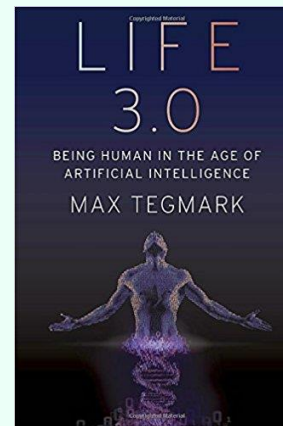


*is co-
founded by*



Max Tegmark

←
writes



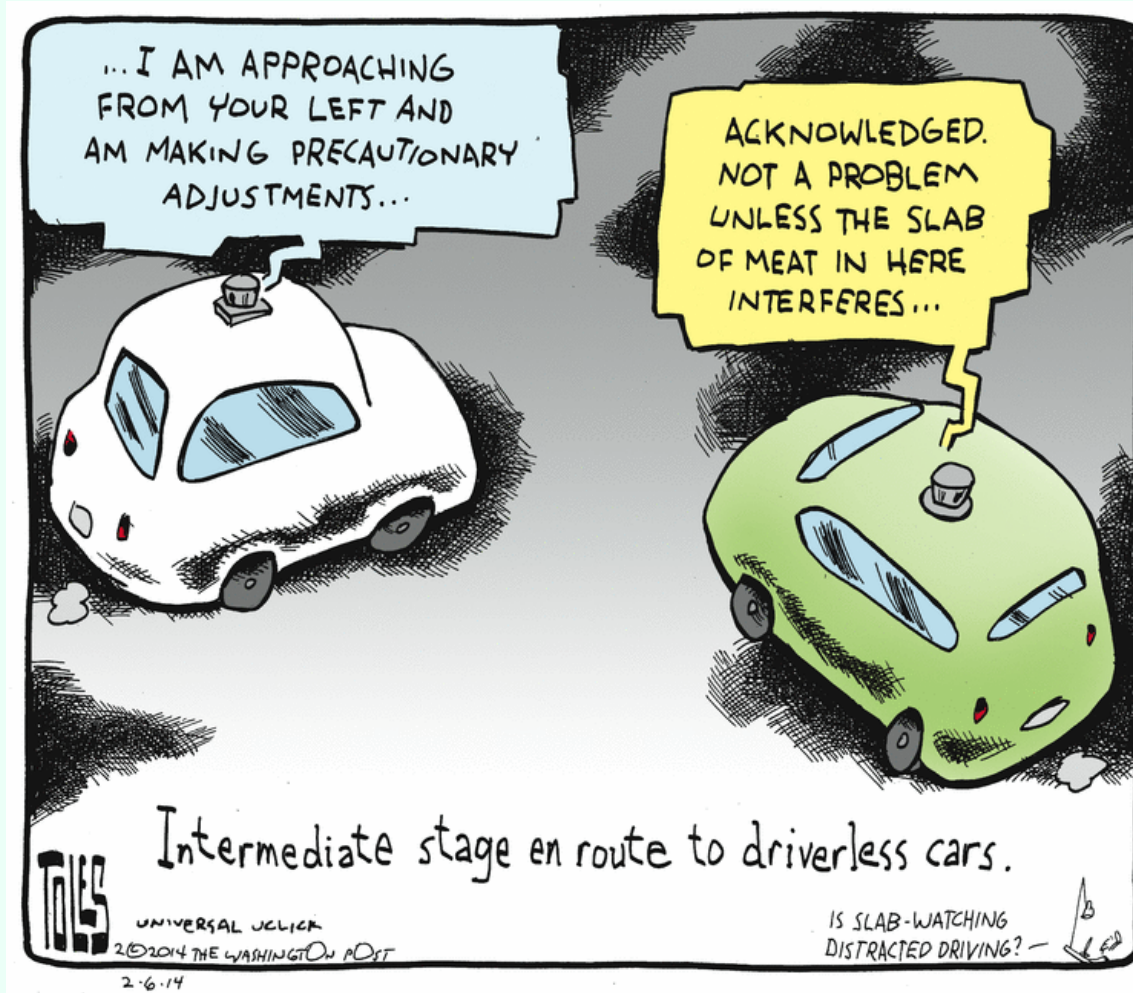
人工智能伦理



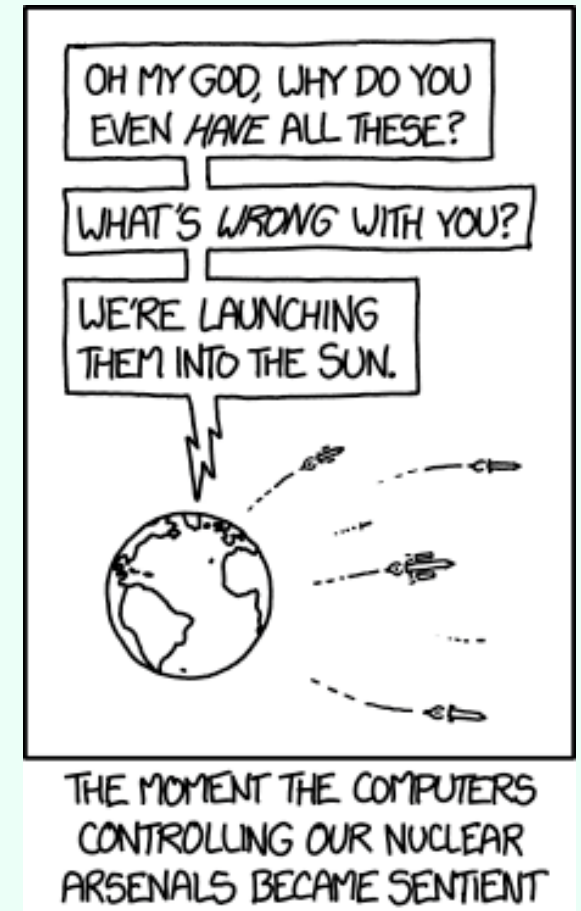


technological unemployment

Worries about AI - near term



autonomous vehicles – legal and other issues



autonomous
weapon systems

...

What is artificial intelligence?

- Popular conception driven by science fiction
 - Robots good at everything except emotions, empathy, appreciation of art, culture, ...
 - ... until later in the movie.
- Current AI is also bad at lots of simpler stuff!
- There is a lot of AI work on thinking about what other agents are thinking



Real AI

- A serious science.
- **General-purpose AI** like the robots of science fiction is incredibly hard
 - Human brain appears to have lots of special and general functions, integrated in some amazing way that we really do not understand (yet)
- **Special-purpose AI** is more doable (nontrivial)
 - E.g., chess/poker/Go playing programs, logistics planning, automated translation, speech and image recognition, web search, data mining, medical diagnosis, keeping a car on the road,

Definitions of AI

if our system can be
more rational than
humans in some
cases, why not?

focus on action

Systems that think
like humans

Systems that think
rationally

sidesteps philosophical
issues such as “is the
system conscious” etc.

Systems that act
like humans

Systems that act
rationally

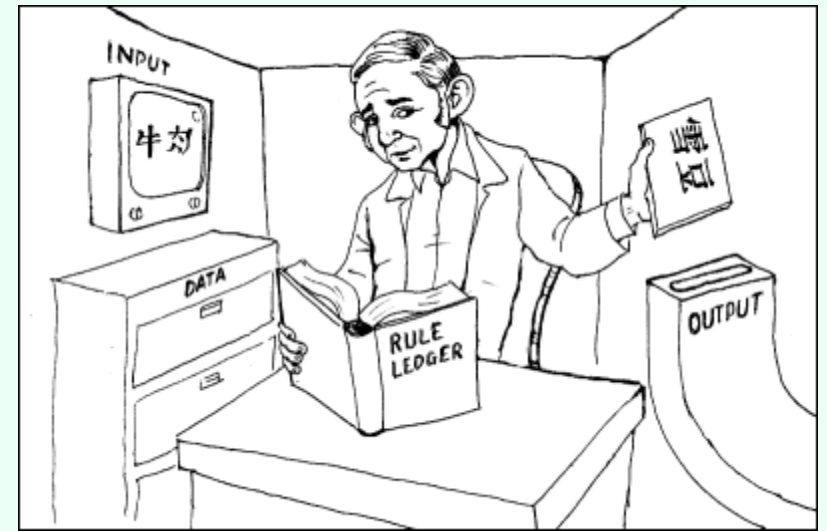
- We will mostly follow “act rationally” approach
 - Distinction may not be that important
 - acting rationally/like a human presumably requires (some sort of) thinking rationally/like a human,
 - humans much more rational anyway in complex domains

人工智能三大流派

- 符号主义（Symbolicism）
 - 逻辑主义(Logicism)、心理学派(Psychologism)或计算机主义(Computerism)
 - 基本思想：物理符号系统(符号操作系统)假设和有限合理性原理
- 连接主义（Connectionism）
 - 又称仿生学派(Bionicsism)或生理学派(Physiologism)
 - 基本思想：神经网络及其连接机制和学习算法
- 行为主义（Actionism）
 - 又称进化主义(Evolutionism)或控制论学派(Cyberneticsism)
 - 基本思想：控制论及感知-动作系统

“Chinese room” argument [Searle 1980]

image from <http://www.unc.edu/~prinz/pictures/c-room.gif>



- **Person** who knows English but not Chinese sits in room
- Receives notes in Chinese
- Has systematic English **rule book** for how to write new Chinese characters based on input Chinese characters, returns his notes
 - Person=CPU, rule book=AI program, really also need lots of paper (storage)
 - Has no understanding of what they mean
 - But from the outside, the room gives perfectly reasonable answers in Chinese!
- Searle's argument: the room has no intelligence in it!

Turing Test

- (Human) judge communicates with a human and a machine over text-only channel,
- Both human and machine try to act like a human,
- Judge tries to tell which is which.
- Numerous variants
- Loebner prize
 - 2016-2018 winner: <http://www.squarebear.co.uk/mitsuku/nfchat.htm>

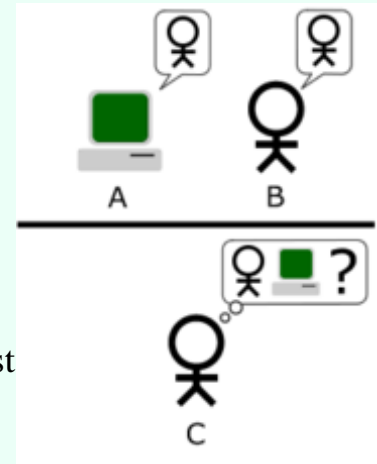


image from http://en.wikipedia.org/wiki/Turing_test

Turing Test on unsuspecting judges

- It is possible to (temporarily) fool humans who do not realize they may be talking to a bot
- ELIZA program [Weizenbaum 66] rephrases partner's statements and questions (~psychotherapist)

Is Turing Test the right goal?

- “Aeronautical engineering texts do not define the goal of their field as making ‘machines that fly so exactly like pigeons that they can fool even other pigeons.’” [Russell and Norvig]

Lessons from AI research

- **Clearly-defined** tasks that we think require intelligence and education from humans tend to be doable for AI techniques
 - Playing chess, drawing logical inferences from clearly-stated facts, performing probability calculations in well-defined environments, ...
 - Although, **scalability** can be a significant issue
- **Complex, messy, ambiguous** tasks that come naturally to humans (in some cases other animals) are much harder...
- ... though recent years have seen remarkable progress, especially in machine learning for narrow domains
 - Image recognition, speech recognition, reinforcement learning in computer games, self-driving cars
- AI systems still lack: broad understanding of the world, common sense, ability to learn from very few examples, truly out-of-the-box creativity...
- We don't understand consciousness. (Does it matter for AI?)

Some areas where humans shine

- Coming up with **reasonably good** solutions in complex messy environments
- **Adapting/self-evaluation/creativity** (“My usual approach to chess is getting me into trouble against this person... Why? Is there something entirely different I can do?”)
- **Analogical reasoning, transfer learning** (applying insights from one domain to another)
- **Explaining** our reasoning
- Tasks that require a **broad understanding** of the (human) world
- Knowing **what it's like to be human**
- **Humor**
- ...

Early history of AI

- 50s/60s: Early successes! AI can draw logical conclusions, prove some theorems, create simple plans... Some initial work on neural networks...
- Led to **overhyping**: researchers promised funding agencies spectacular progress, but started running into difficulties:
 - **Ambiguity**: highly funded translation programs (Russian to English) were good at syntactic manipulation but bad at disambiguation
 - “The spirit is willing but the flesh is weak” becomes “The vodka is good but the meat is rotten”
 - **Scalability/complexity**: early examples were very small, programs could not scale to bigger instances
 - Limitations of **representations** used

History of AI...

- 70s, 80s: Creation of **expert systems** (systems specialized for one particular task based on experts' knowledge), wide industry adoption
- Again, overpromising...
- ... led to **AI winter(s)**
 - Funding cutbacks, bad reputation

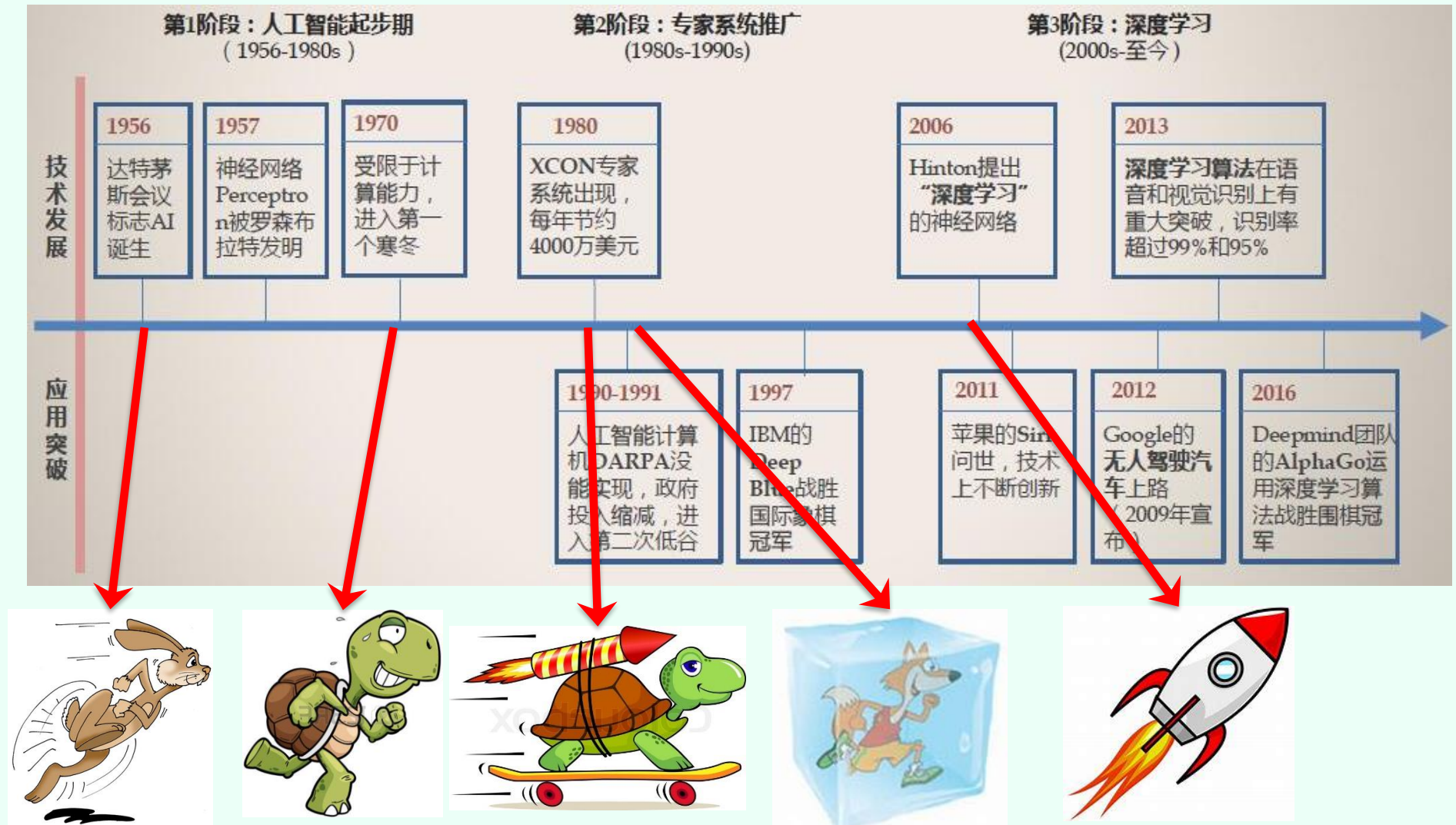
Modern AI

- More rigorous, scientific, formal/mathematical
- Fewer grandiose promises
- Divided into many subareas interested in particular aspects
- More directly connected to “neighboring” disciplines
 - Theoretical computer science, statistics, economics, operations research, biology, psychology/neuroscience, ...
 - Often leads to question “Is this really AI”?
- Some senior AI researchers are calling for re-integration of all these topics, return to more grandiose goals of AI
 - Somewhat risky proposition for graduate students and junior faculty...

History of AI

- 1943 McCulloch & Pitts: Boolean circuit model of brain
- 1950 Turing's "Computing Machinery and Intelligence"
- 1956 Dartmouth meeting: "Artificial Intelligence" adopted
- 1950s Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
- 1965 Robinson's complete algorithm for logical reasoning
- 1966-73 AI discovers computational complexity
Neural network research almost disappears
- 1973 - 1980 The first winter of AI
- 1969-1979 Early development of knowledge-based systems
- 1980 -1987 The rise of expert system (Lisp machine)
- 1986 -- Neural networks return to popularity
- 1987 - 1993 The second winter of AI
- 1993- present The emergence of intelligent agents (Multiagent Systems)
- 2000 - present Deep Learning (2010) + Reinforcement Learning (2015)
Multiagent Deep Reinforcement Learning

History of AI



Some AI examples

- 微软诗人小冰

太阳不吝惜光的施予
撒向天空的笑声
和着妙境的音乐
水中清澈的灵魂
一如你纯然的美
在这徘徊无助的人间
有了她们就像有了全世界
挥动着透明的羽翼
飞进了梦里的梦
在从容的天空中起舞
繁星闪现



机器人作诗能否打动人类

后来，我考上了大学
有了第一台爱笑的相机
也有姐姐一样的漂亮牙齿
我抱着相机
给姐姐看铺进村子的铁轨
父母的夕阳婚纱照
和社团里一起爬山的男孩
姐姐笑着凑到我耳边
悄悄讲了个故事



机器人作诗能否打动人类

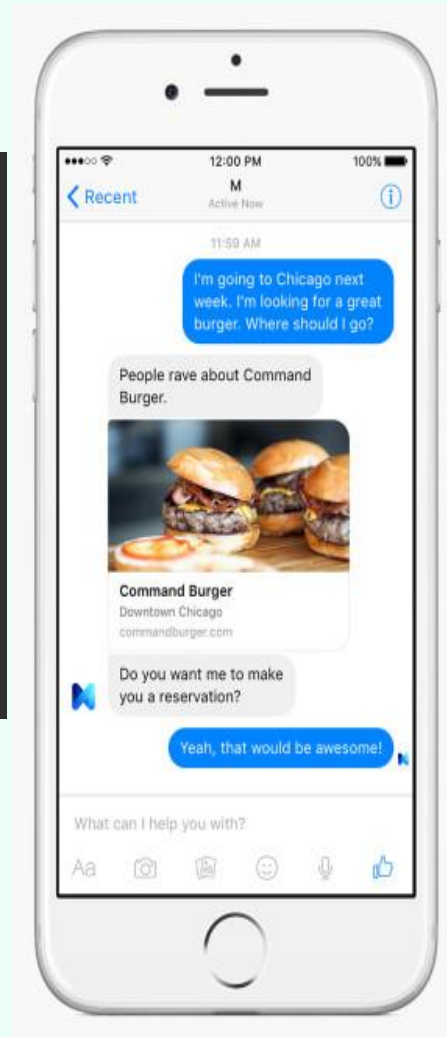
她借他们相机，如赠白马
驮着山里的孩子去远方
现在
要再送他一名天堂的银匠
将没有名字的悲伤
锻造成闪亮的马鞍
那嘹亮笑容里
有对贫瘠生活
精致的反抗
在柔软又带刺的山林里
像只喜鹊一样地活着



机器人作诗能否打动人类

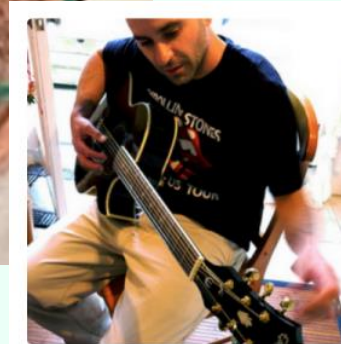
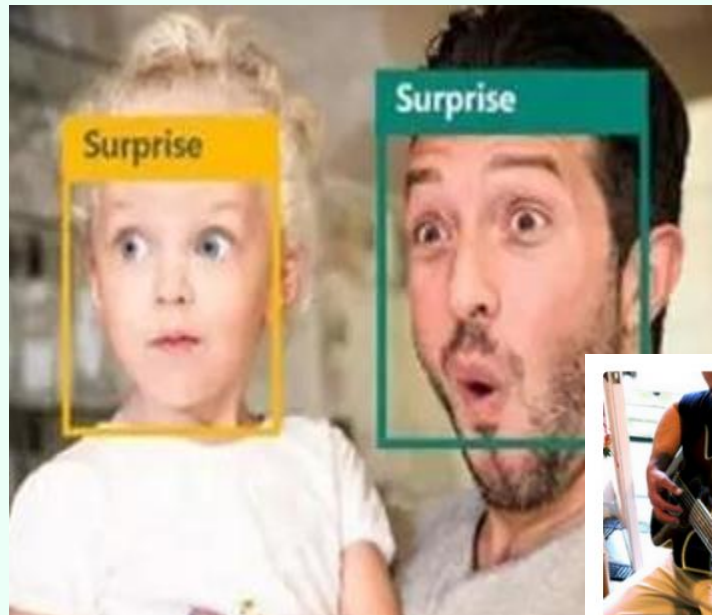
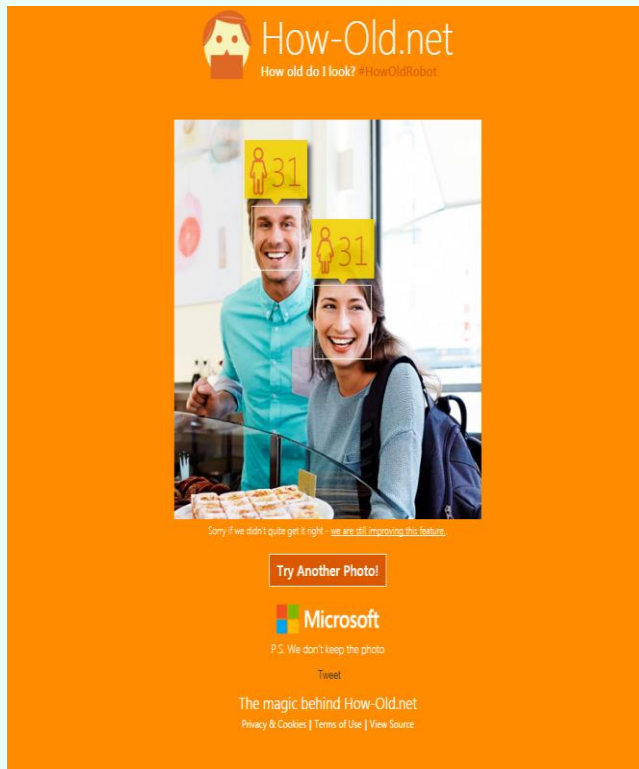
Some AI examples

- Personal Assistant



Some AI examples

- Image



"man in black shirt is playing guitar."



"construction worker in orange safety vest is working on road."



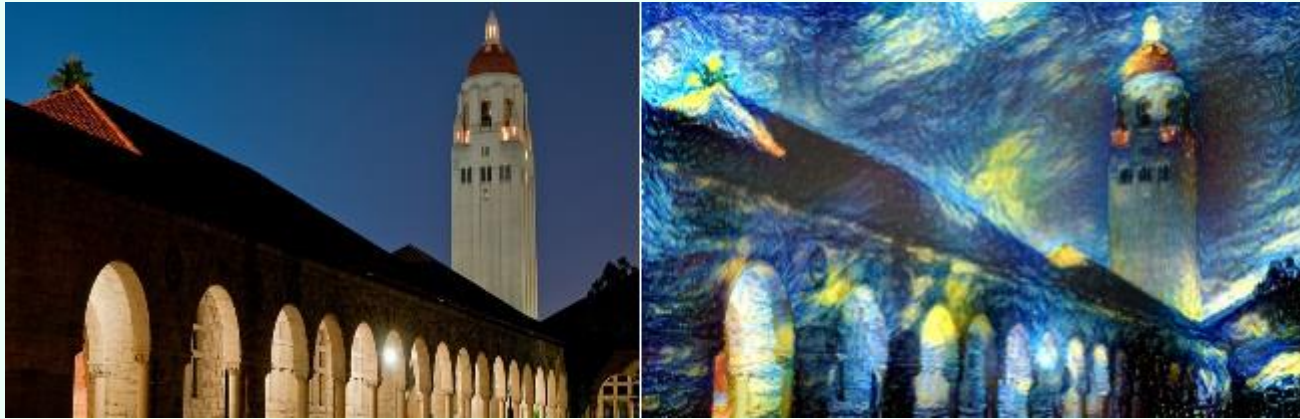
"girl in pink dress is jumping in air."



"black and white dog jumps over bar."

Some AI examples

- Painting



斯坦福校园

星空风格的斯坦福校园



梵高 - 《星空》

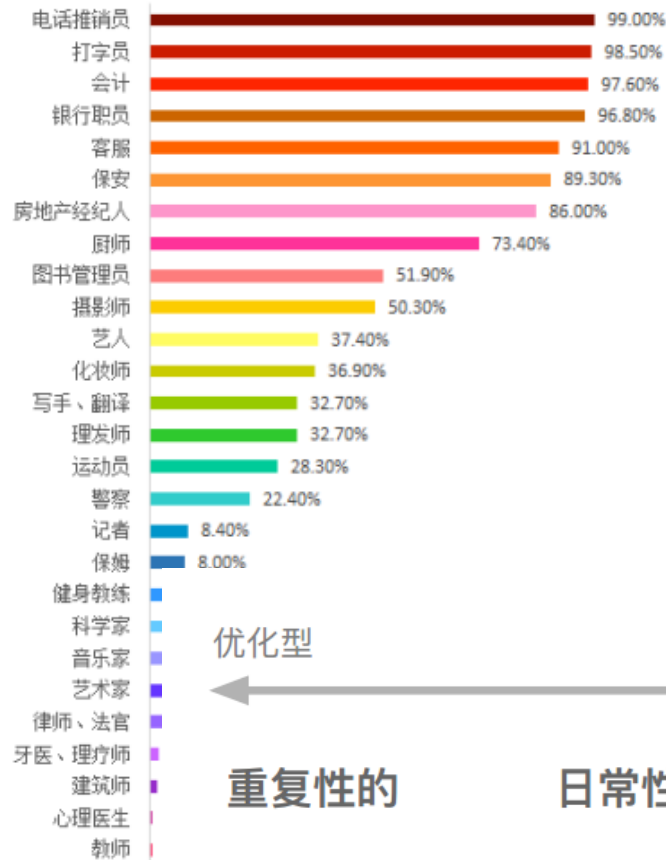
甘道夫

星空风格的甘道夫

New revolution



Are we prepared



翻译



记者



司机



客服



保姆



助理



保安



交易员

优化型

创意/决策型

重复性的

日常性的

优化性的

复杂的

创意的

电话销售

洗碗工

客服

AI

5

年

卡车司机

验血师

保安

AI

10

年

放射科医生

记者

研究分析师

AI

15

年

CEO

并购专家

经济学家

安全

专栏作家

科学家

艺术家

安全

A threat?

future of life INSTITUTE

Technology is giving life the potential to flourish like never before...

News: AI Biotech Nuclear Climate Partner Orgs

An Open Letter

RESEARCH PRIORITIES FOR ROBUST AND BENEFICIAL ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) research has explored a variety of problems and approaches since its inception, but for the last 20 years or so has been focused on the problems surrounding the construction of intelligent agents – systems that perceive and act in some

By 2014, both physicist Stephen Hawking and business magnate Elon Musk had publicly voiced the opinion that superhuman artificial intelligence could provide incalculable benefits, but could also end the human race if deployed incautiously.

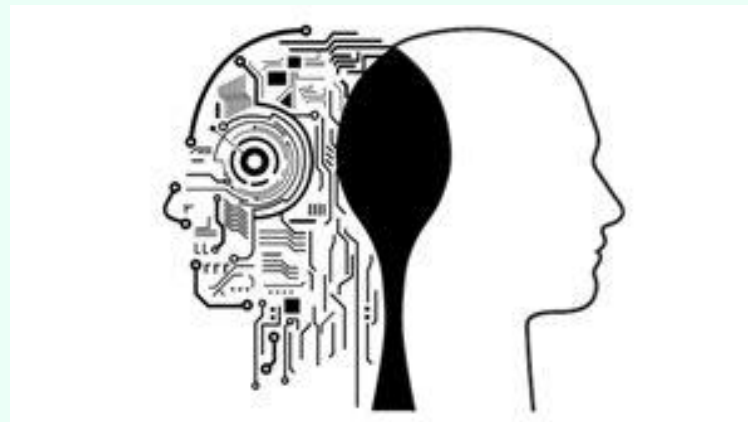
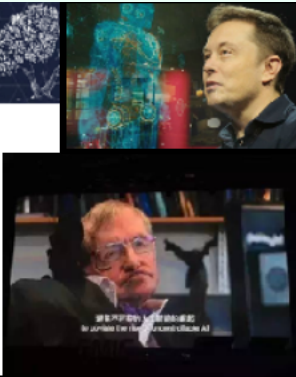
2014年开始，霍金和马斯克公开宣称超级人工智能的价值难以估计，但同时，人类也可能招致灭顶之灾

Hawking and Musk both sit on the scientific advisory board for the Future of Life Institute, an organization working to "mitigate existential risks facing humanity".

两人都是“生命未来（future of life）”研究所科学顾问

The institute drafted an open letter directed to the broader AI research community, and circulated it to the attendees of its first conference in Puerto Rico during the first weekend of 2015. The letter was made public on January 12.

生命未来研究所起草了一封公开信，2015年第一个周末的会议上首次发给与会者，2015年1月12日正式公开.....



Some AI videos

- Note: there is a lot of AI that is not quite this “sexy” but still very valuable!
 - E.g. logistics planning – DARPA claims that savings from a single AI planning application during 1991 Persian Gulf crisis more than paid back for all of DARPA’s investment in AI, ever. [Russell and Norvig]
- <https://www.youtube.com/user/aaaivideocompetition>
- <https://www.youtube.com/watch?v=1JJsBFiXGI0>
- <https://www.youtube.com/watch?v=s6VIWUDUHTa4>
- <https://www.youtube.com/watch?v=C5Xnxjq63Zg>
- <https://www.youtube.com/watch?v=ScXX2bndGJc>
- <https://www.youtube.com/watch?v=V1eYniJ0Rnk>

This course

- Focus on **general AI techniques** that have been useful in many applications
- Will try to **avoid application-specific techniques** (still interesting and worthwhile!)