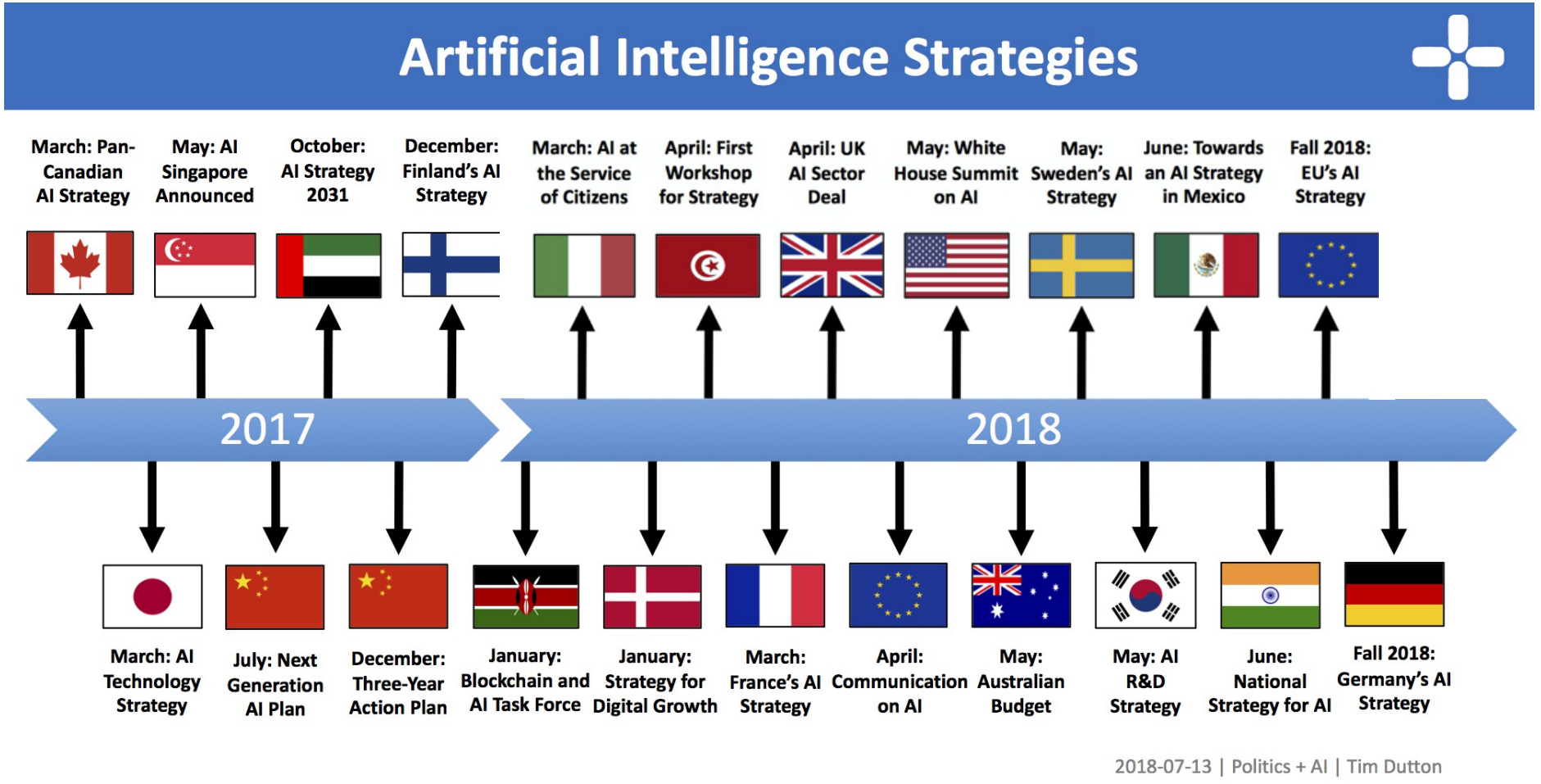
浅谈人工智能的下一个十年

Jie Tang
Computer Science
Tsinghua University

PPT可以在主页下载: http://keg.cs.tsinghua.edu.cn/jietang/ 或者： Google “Jie Tang”

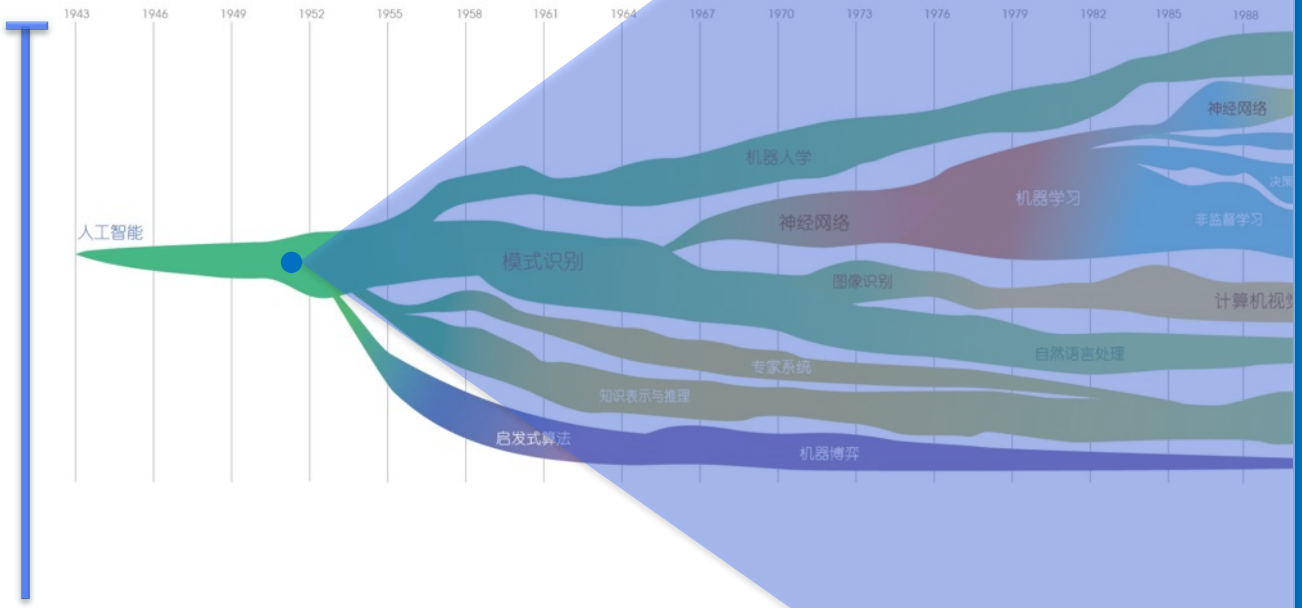
人工智能的第三次浪潮


Artificial Intelligence Strategies
March: Pan-Canadian AI Strategy
May: AI Singapore Announced
October: AI Strategy 2031
December: Finland's AI Strategy
March: AI at the Service of Citizens
April: First Workshop for Strategy
April: UK AI Sector Deal
May: White House Summit on AI
May: Sweden's AI Strategy
June: Towards an AI Strategy in Mexico
Fall 2018: EU's AI Strategy
2017
2018
March: AI Technology Strategy
July: Next Generation AI Plan
December: Three-Year Action Plan
January: Blockchain and AI Task Force
January: Strategy for Digital Growth
March: France's AI Strategy
April: Communication on AI
May: Australian Budget
May: AI R&D Strategy
June: National Strategy for AI
Fall 2018: Germany's AI Strategy
2018-07-13 | Politics + AI | Tim Dutton

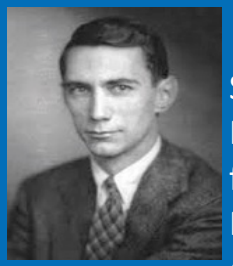
人工智能历史

SEARCH

人工智能领域发展趋势 Pc



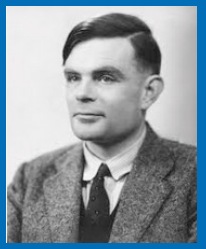
1950计算机象棋博弈



Claude Shannon

Shannon, Claude E. "XXII. Programming a computer for playing chess." Philosophical magazine 41.314 (1950): 256-275.

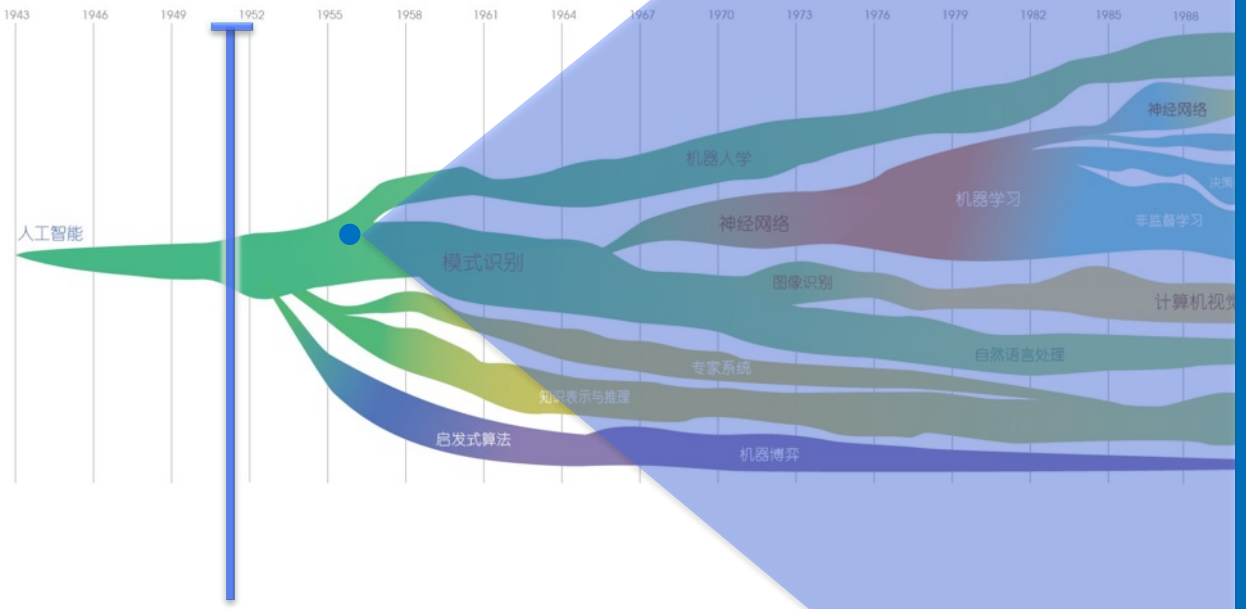
1954图灵测试



Alan Turing

Turing, Alan M. "Solvable and unsolvable problems." Science News-ens. fr 39 (1954).

人工智能领域发展趋势 Po



1956达特茅斯会议


John McCarthy


Marvin Minsky


Nathan Rochester


Claude Shannon

McCarthy, J., et al. "Dartmouth Conference." Dartmouth Summer Research Conference on Artificial Intelligence. 1956

1959一般问题解决器

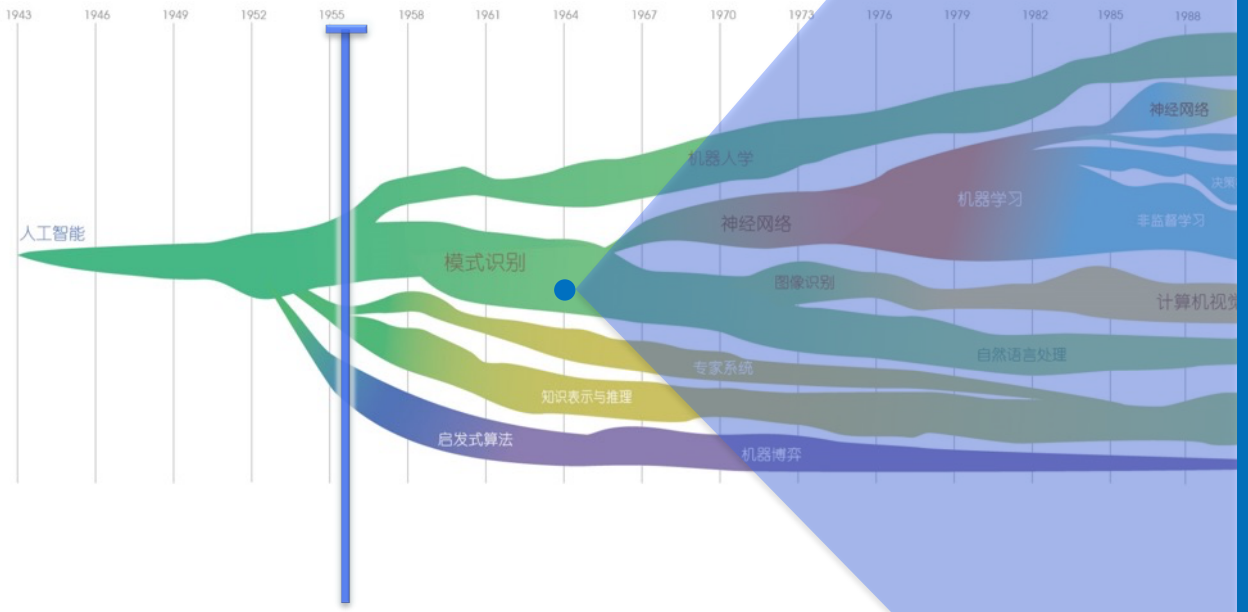

Herbert Simon


J.C. Shaw


Allen Newell

Newell, A.; Shaw, J.C.; Simon, H.A. (1959). Report on a general problem-solving program. Proceedings of the International Conference on Information Processing. pp. 256–264.

人工智能领域发展趋势 Po



1964 理解自然语言输入

Daniel Bobrow

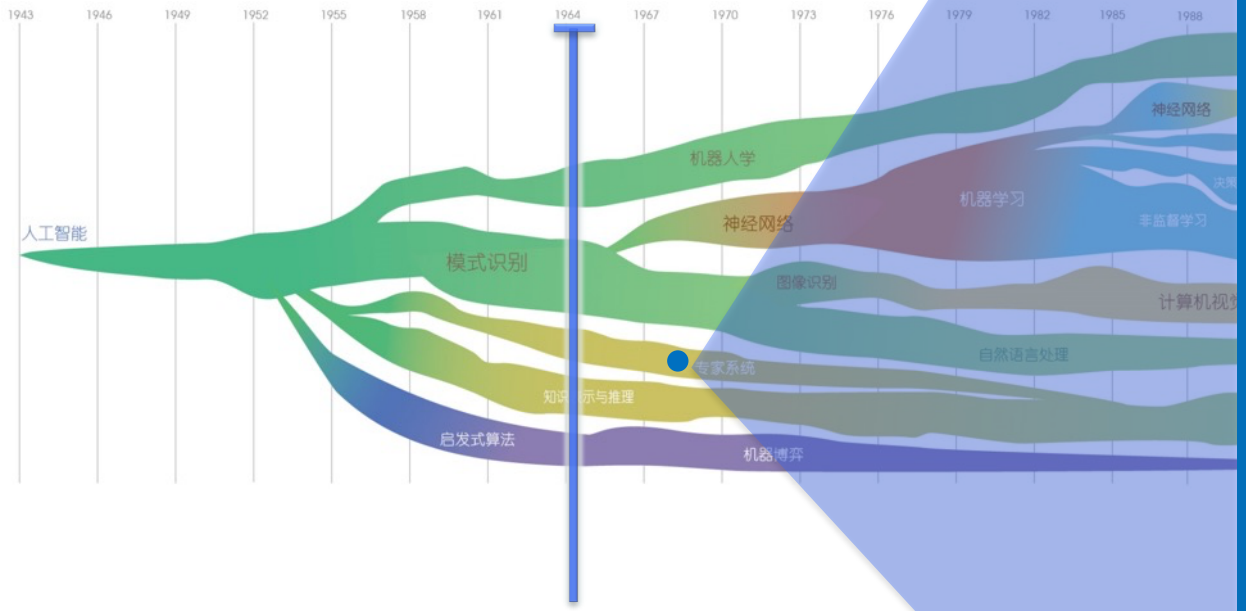
Bobrow, Daniel G. "Natural language input for a computer problem solving system." (1964)

1966 ELIZA人机对话

Joseph Weizenbaum

Weizenbaum, Joseph. "ELIZA—a computer program for the study of natural language communication between man and machine." Communications of the ACM 9.1 (1966): 36-45.

人工智能领域发展趋势 Po

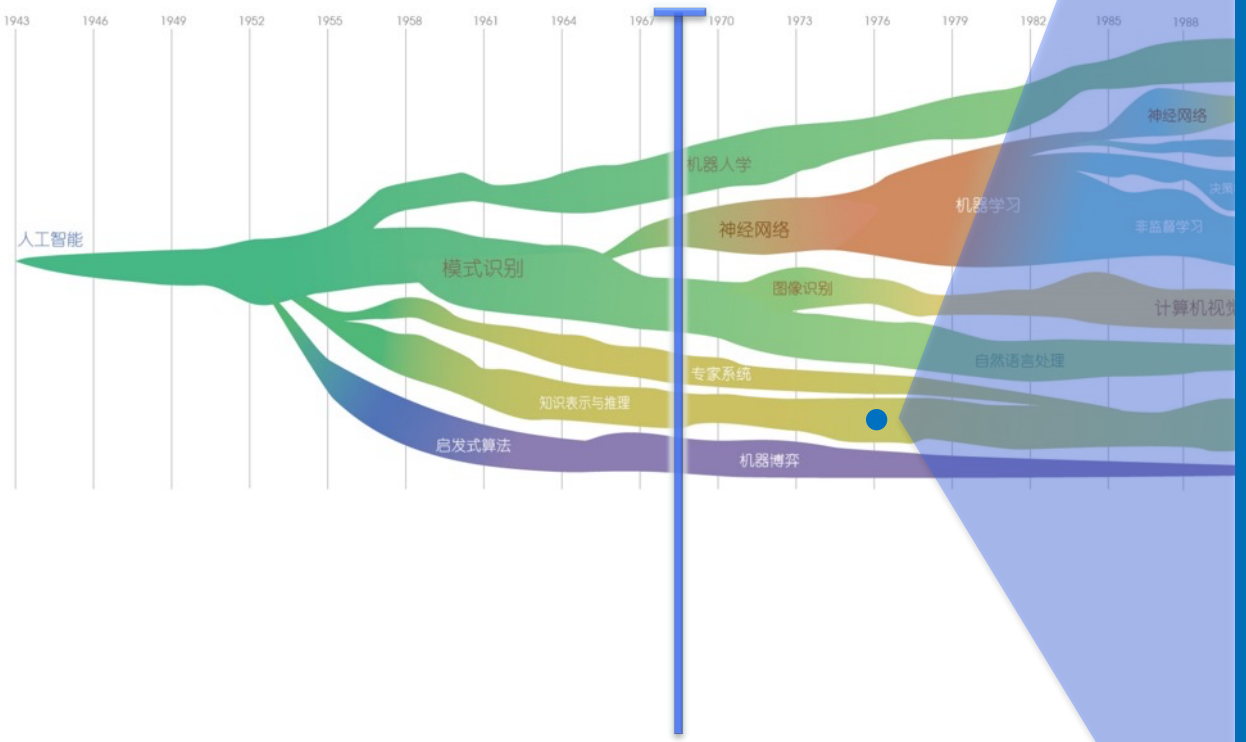


1968 世界首个专家系统DENDRAL

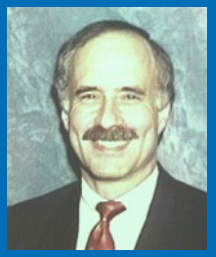
Edward Feigenbaum

Buchanan, Bruce, Georgia Sutherland, and Edward A. Feigenbaum. Heuristic DENDRAL: a program for generating explanatory hypotheses in organic chemistry. Defense Technical Information Center, 1968.

人工智能领域发展趋势 Po



1976 大规模知识库构建与维护



Randall Davis

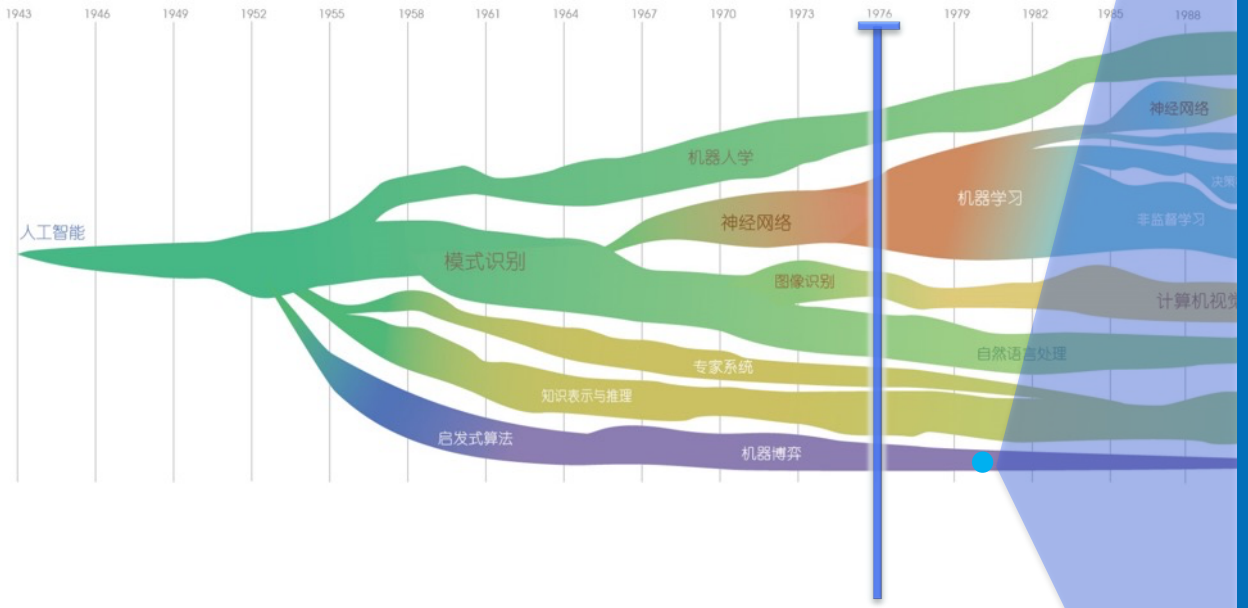
Applications of meta level knowledge to the construction, maintenance and use of large knowledge bases[M]. Stanford University, Computer Science Department, AI Laboratory, 1976.

1980 非单调逻辑

Drew McDermott, Jon Doyle

McDermott D, Doyle J. Non-monotonic logic I[J]. Artificial intelligence, 1980, 13(1): 41-72.

人工智能领域发展趋势 Po

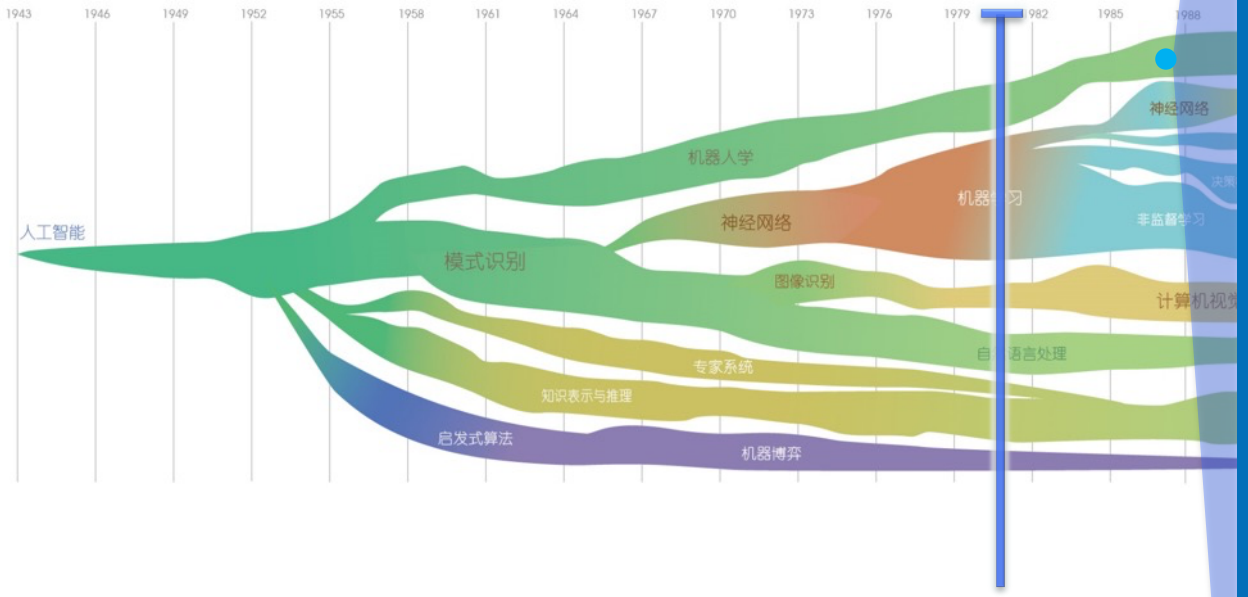


1980 计算机战胜双陆棋世界冠军

Hans Berliner

Berliner H J. Backgammon computer program beats world champion[J]. Artificial Intelligence, 1980, 14(2): 205-220.

人工智能领域发展趋势 Po

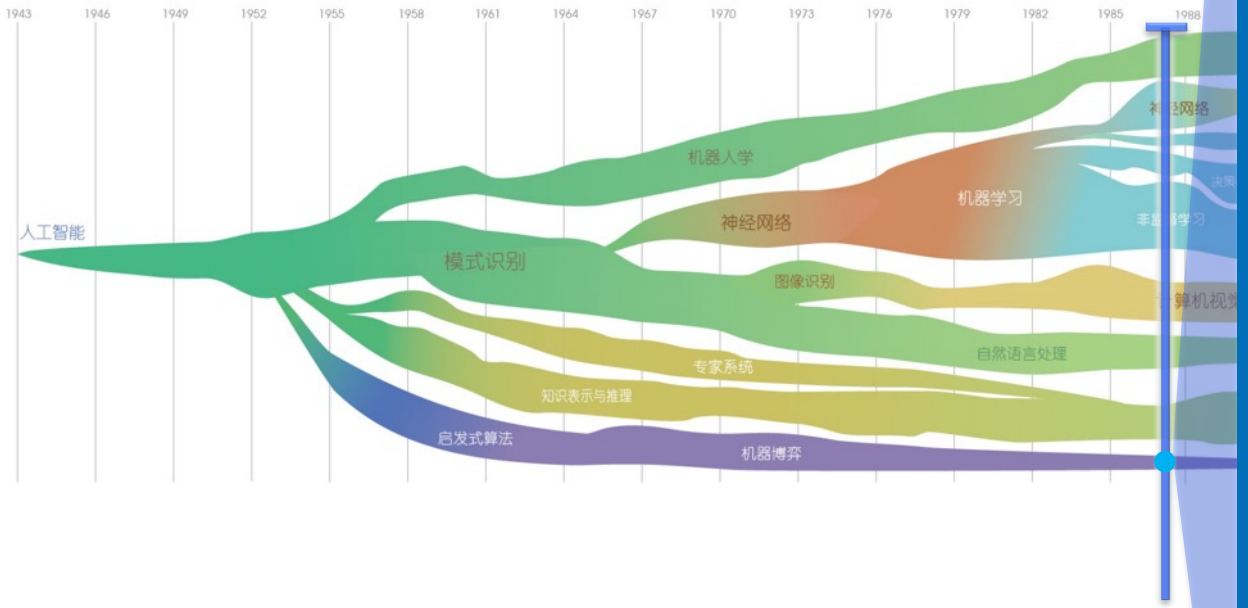


1987 基于行为的机器人学

Rodney Brooks

Brooks R. A robust layered control system for a mobile robot[J]. Robotics and Automation, IEEE Journal of, 1986, 2(1): 14-23

人工智能领域发展趋势 Po



1987 自我学习双陆棋程序

Gerry Tesauro

Tesauro G. TD-Gammon, a self-teaching backgammon program, achieves master-level play[J]. Neural computation, 1994, 6(2): 215-219.

1998 语义互联网路线图

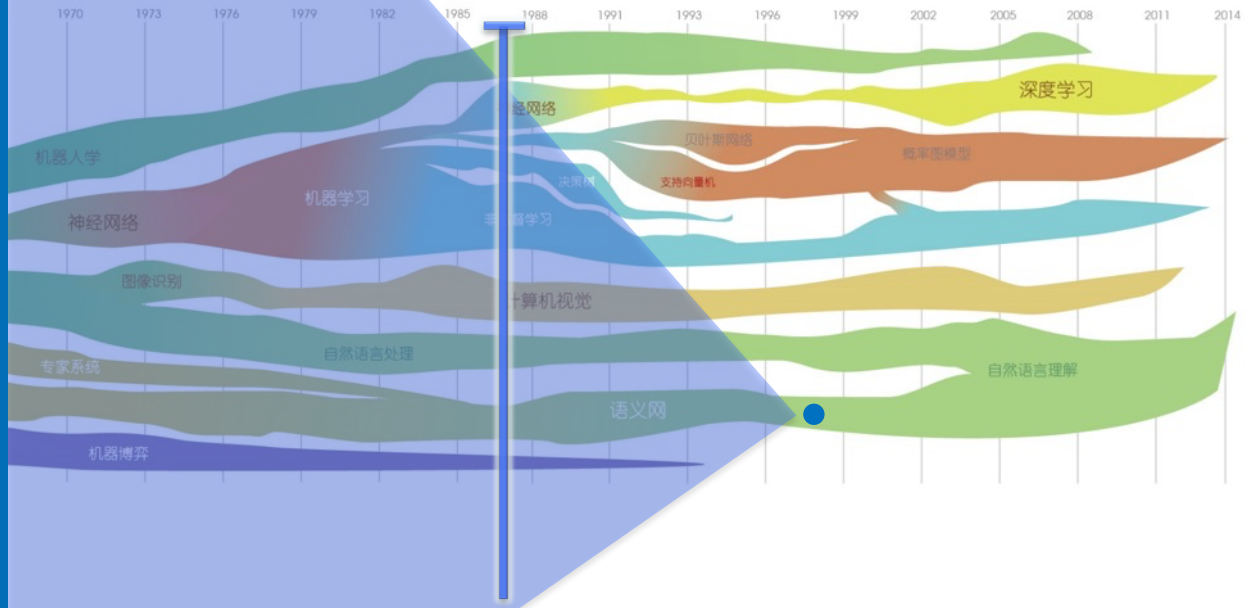
Tim Berners-Lee

Berners-Lee, Tim. "Semantic web road map." (1998).

2004 OWL语言

McGuinness, Deborah L., and Frank Van Harmelen. "OWL web ontology language overview." W3C recommendation 10.2004-03 (2004): 10.

域发展趋势 Powered by AMiner



2006 深度学习

Geoffrey Hinton

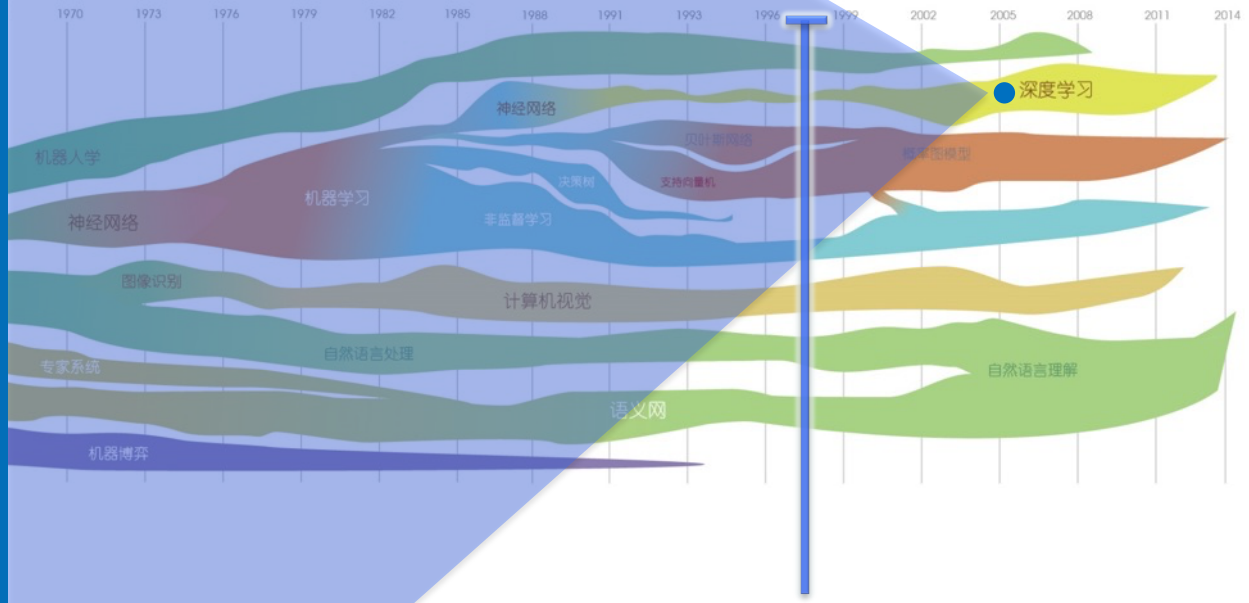
Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh. "A fast learning algorithm for deep belief nets." Neural computation 18.7 (2006): 1527-1554.

2011 高层抽象特征构建



Le, Quoc V., et al. "Building high-level features using large scale unsupervised learning." arXiv preprint arXiv:1112.6209 (2011).

域发展趋势 Powered by AMiner

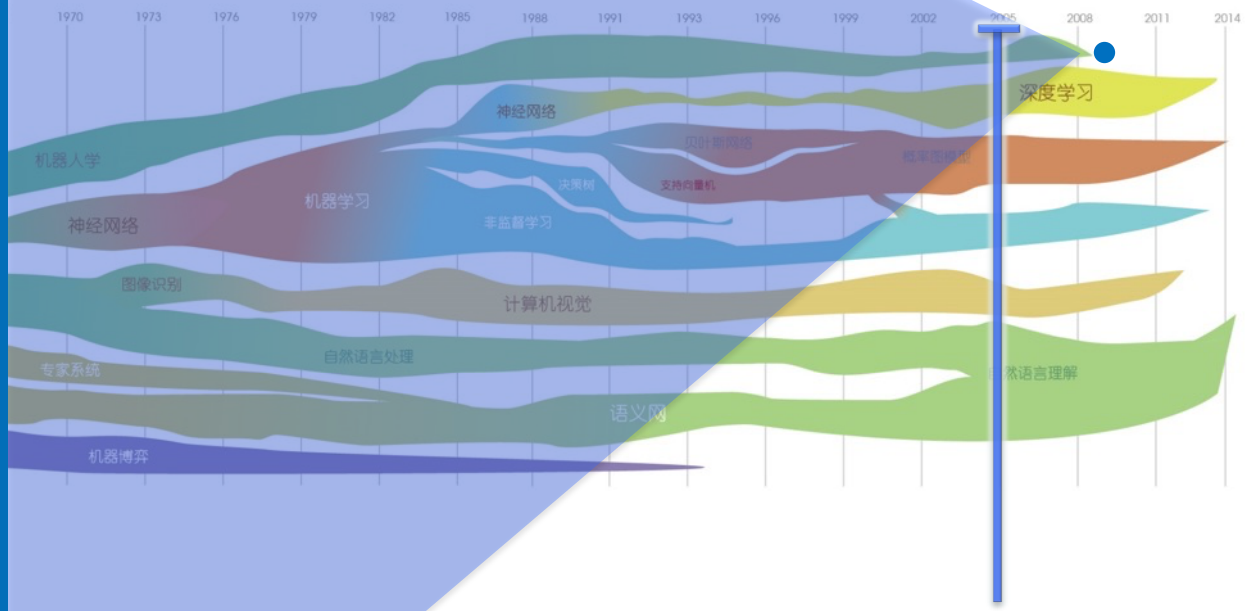


2009 谷歌自动驾驶汽车

Sebastian Thrun

Markoff, John. "Google cars drive themselves, in traffic." The New York Times 10 (2010): A1.

域发展趋势 Powered by AMiner



2011 沃森获得Jeopardy冠军

IBM’s Watson

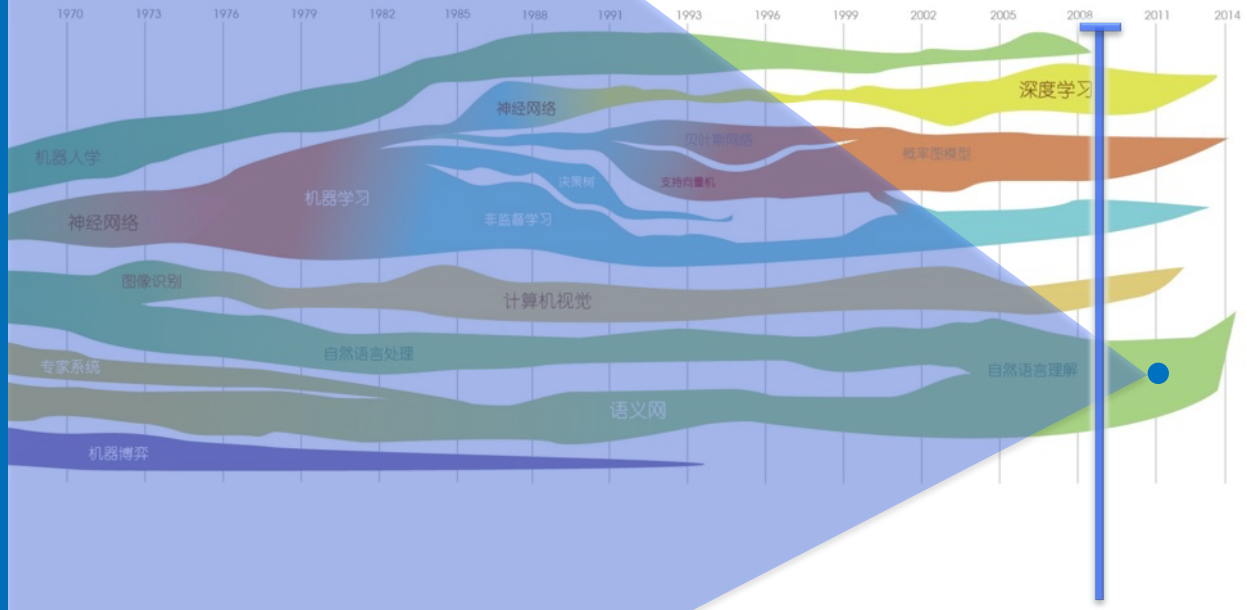
Markoff, John. "Computer program to take on ‘Jeopardy!’." The New York Times (2009).

2011 自然语言问答

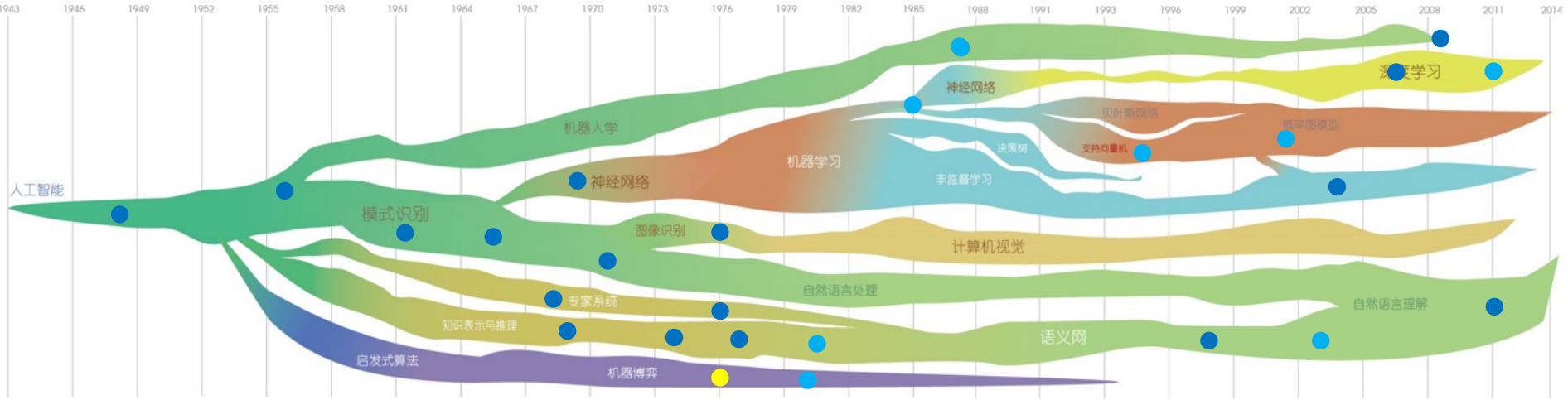
Apple’s Siri

Sadun, Erica, and Steve Sande. Talking to Siri: Learning the Language of Apple's Intelligent Assistant. Que Publishing, 2013.

域发展趋势 Powered by AMiner



人工智能领域发展趋势 Powered by AMiner



人工智能近10年

SEARCH

人工智能

博弈对策

无人驾驶

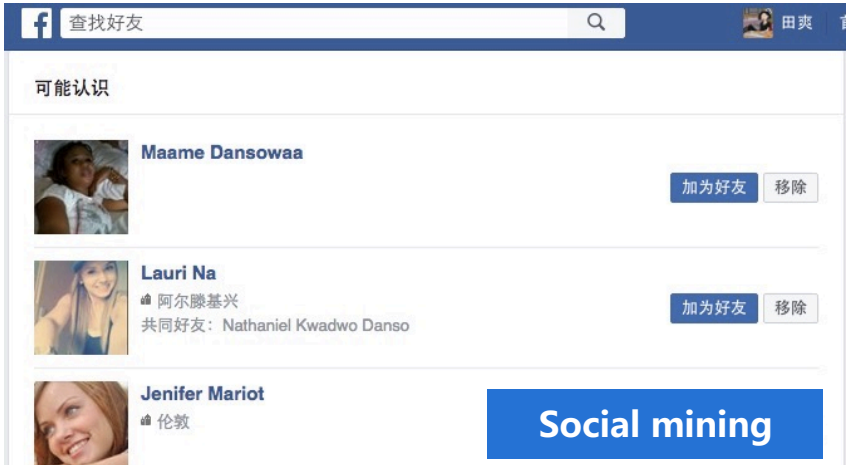
- dog + cat =

图像识别

Look broader

QA

Recommendatio



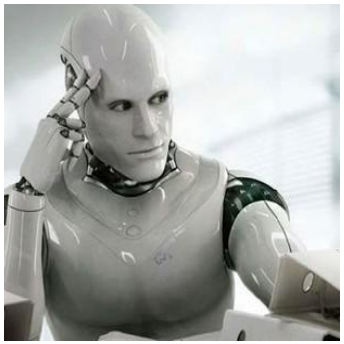
Social mining

Search engine

National AI Strategy

国务院印发《新一代人工智能发展规划》

“要把人工智能发展放 在国家战略层面系统布局、主动谋划，牢牢把握人工智能发展新 阶段国际竞争的战略主动，打造竞争新优势、开拓发展新空间， 有效保障国家安全。” 2017, July



Intelligent

AI Education

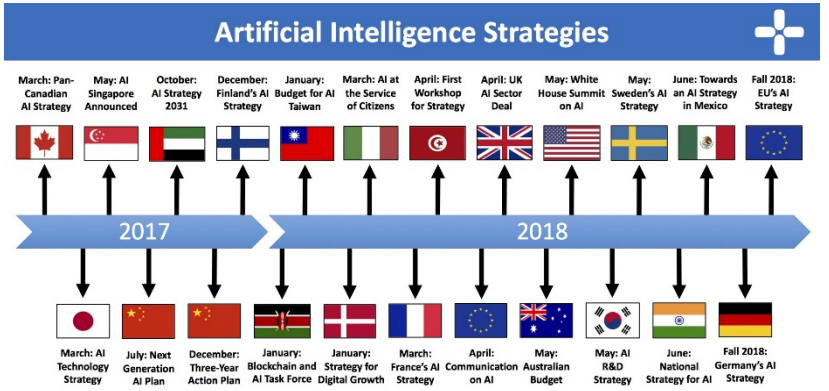
AI Agriculture

AI Everything…

....

National AI Strategies

- Example Countries
  - In 2017, Canada launched Pan-Canadian Artificial Intelligence Strategy
    - AMII, Vector, Mila
  - This year, UK is putting AI and Data at the heart of its Industrial Strategy
  - In March 2018, France announces his AI strategy with investing €1.5 billion over 5 yrs on AI research and innovation
  - US recently also release a national strategy for artificial intelligence. Moreover,
    - OpenAI
    - TensorFlow



AI开放平台

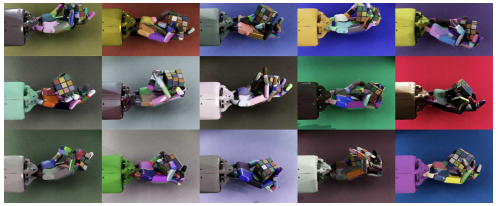
- 部分国际AI开放平台
  - OpenAI (Sutskever)
  - Mila (Bengio)
  - Vector Institute (Hinton)

OpenAI

- Research Director: Ilya Sutskever
- Sponsored by
  - 马斯克
  - Y Combinator总裁阿尔特曼
  - 天使投资人彼得•泰尔等

OpenAI

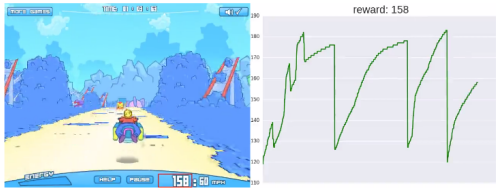
- 目标：制造通用机器人和使用自然语言的聊天机器人



单个机械手还原魔方：Dactyl机器人

击败Dota2游戏国际职业选手：OpenAI Five

强化学习算法游戏模拟平台：Gym



通用强化学习算法测评平台：Universe

模型	总参数量	网络层数	训练数据量级
GPT	1.25 亿	12	十亿词
GPT-2	15 亿	48	百亿词
GPT-3	1750 亿	96	千亿词

全球第一个通用预训练语言模型，启发了BERT，当前参数最大的模型，在多项NLP任务中表现优异

Mila (Bengio)

- Founder: Yoshua Bengio
- 以人为本的AI平台
  - 培养年轻人
  - 专注Deep Learning
- 创新创业

Vector Institute (Hinton)

- Founder: Geoffrey Hinton
- 以研究为主
  - 专注Deep Learning
- 创新创业

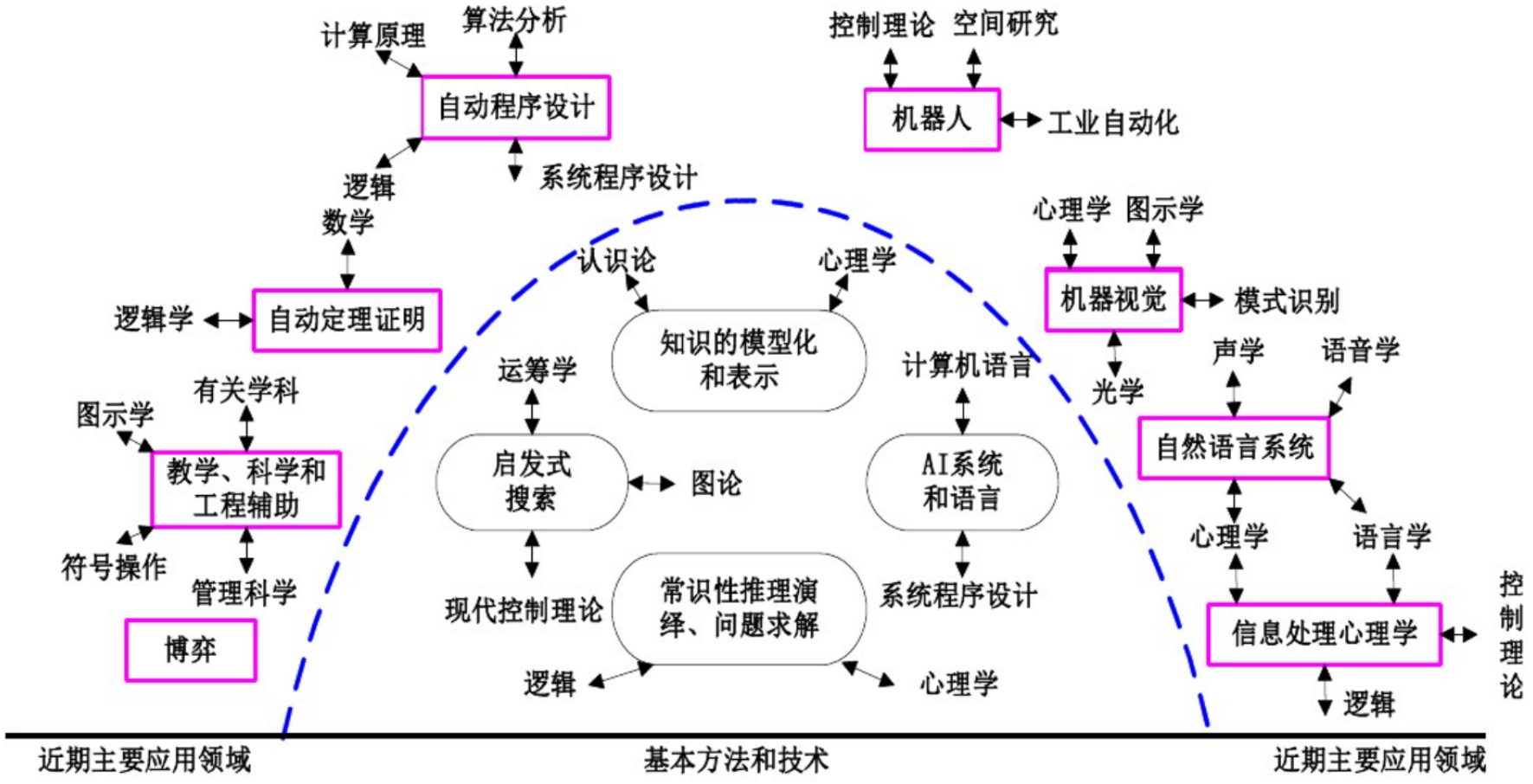
AI in Tsinghua

- In September, we established Tsinghua AI Institute
  - By Prof. Bo Zhang
- Focus
  - Fundamental theories and methodologies
  - Interdisciplinarity

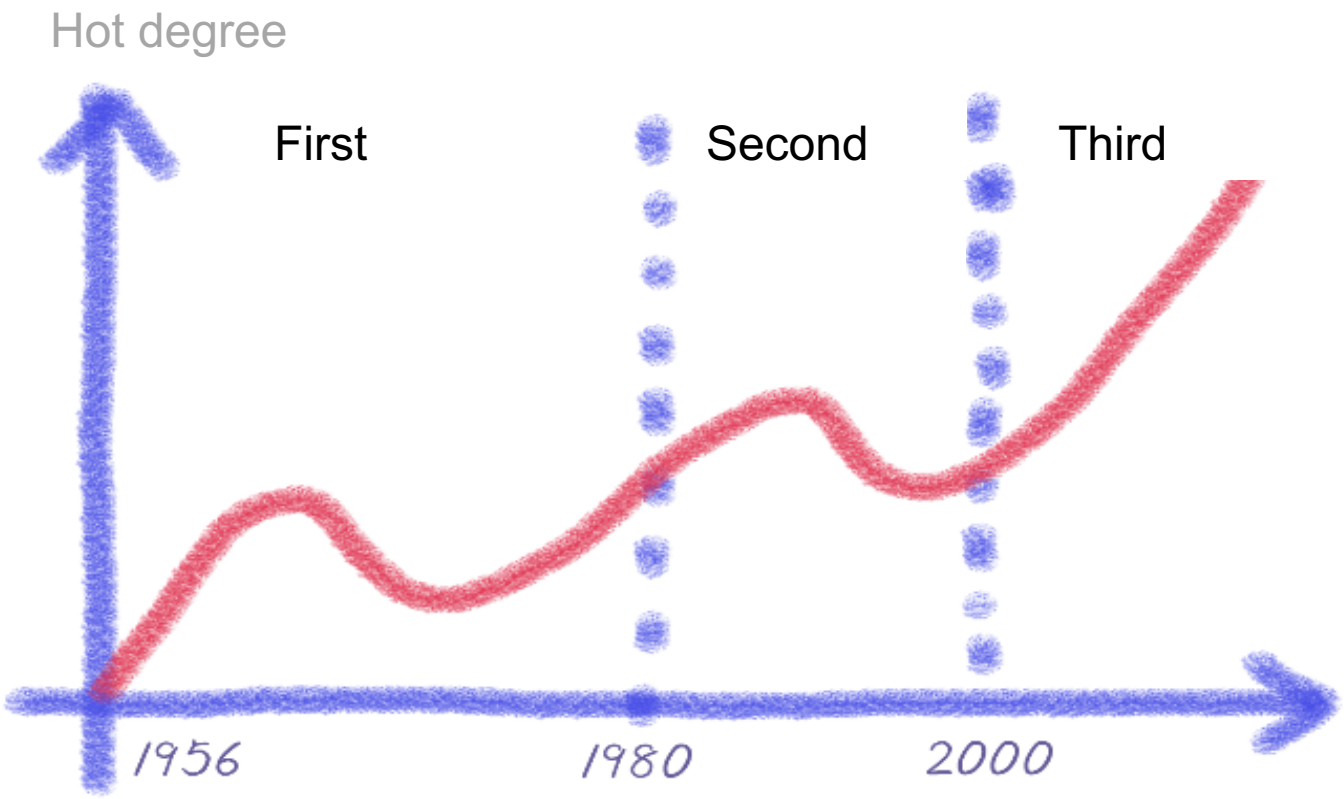
Beijing AI Institute

- In November, Beijing launched Beijing AI Institute
  - By Dr. Hongjiang Zhang
- Focus
  - ground-breaking research

学科领域交叉与渗透


计算原理
算法分析
自动程序设计
逻辑
数学
系统程序设计
逻辑学
自动定理证明
图示学
有关学科
教学、科学和
工程辅助
符号操作
管理科学
博弈
控制理论
空间研究
机器人
工业自动化
心理学
图示学
机器视觉
模式识别
光学
声学
语音学
自然语言系统
心理学
语言学
信息处理心理学
逻辑
控制理论
认识论
心理学
知识的模型化
和表示
运筹学
启发式
搜索
图论
现代控制理论
常识性推理演
绎、问题求解
逻辑
心理学
计算机语言
AI系统
和语言
系统程序设计
近期主要应用领域
基本方法和技术
近期主要应用领域

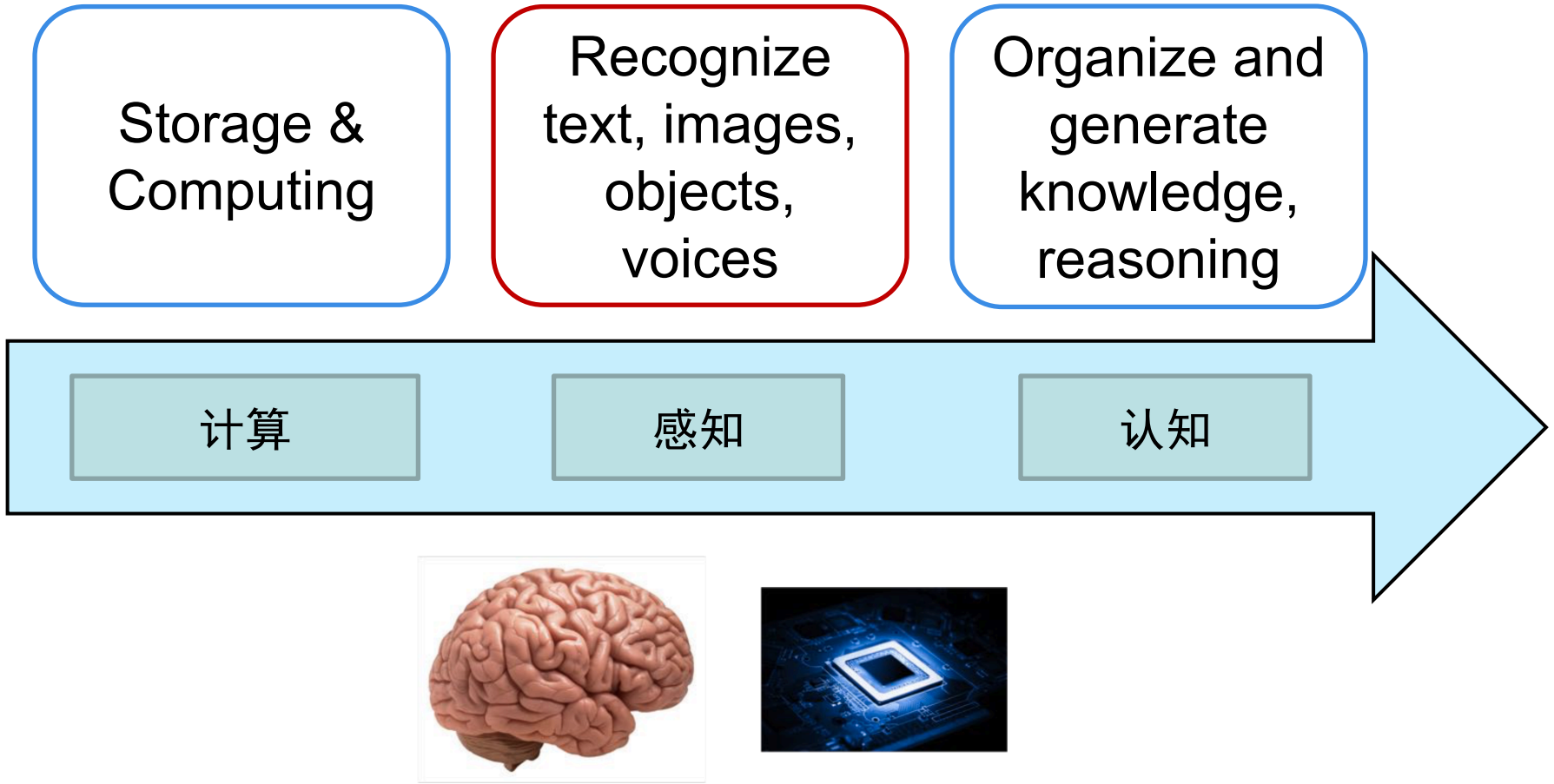
AI History



• Logic Theory Machine • Perceptron	• Expert System • Decision Tree	• Big Data • Deep Learning

AI趋势：从感知到认知

- 从感知到认知



Stochastic vs Deterministic

Uncertainty!

机器学习快速发展

Classification Models

Decision tree
Bayesian Classifier
Perceptron
Neural Networks

Maximal Margin

SVM
(Vapnik)

Reinforcement Learning

Q-Learning
Policy Gradient
Actor-Critic
TRPO

Sequential Learning

HMM, MEMM, CRF,
Voted Perceptron

Deep Learning

DNN, CNN, ResNet

Deep RL

DQN, VIN, A3C

Graphical Models

Factor Graph,
Exponential Model,
Topic Model (PLSI, LDA)

Recurrent Nets

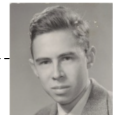
RNN, LSTM, GRU

Deep Generative Models

DBN, AutoEncoder,
VAE, GAN, WGAN

For more from http://aminer.cn/

DL算法

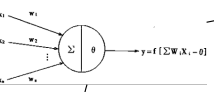


Frank Rosenblatt
Cornell University
psychologist

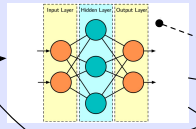
Geoffery Hinton
University of Toronto
deep learning

Yann Lecun
New York University
deep learning

Perceptron(1958)



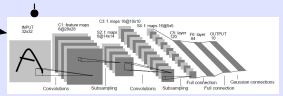
BPNN/MLP(1986)

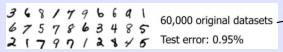


Neocognitron(1980)
[convolution & pooling]

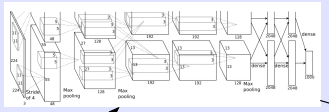
Hopfield Network(1982)
[recurrent & feedback]

LeNet/CNN(1998)





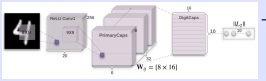
AlexNet(2012)

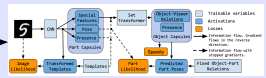


Relu, dropout & bigger

Stacked Capsule Autoencoders(2019)

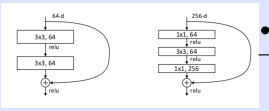
Capsule Nets(2017)



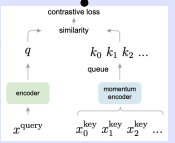


Kaiming He (何恺明)
MSRA => FAIR
computer vision

MoCo(2019)



constrastive learning



ResNet(2016)

VGG(2014)
GoogLeNet(2015)

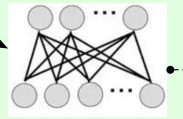
DenseNet(2017)
SENet(2018)

NAS(2018)

EfficientNet (2019)

RegNet (2020)

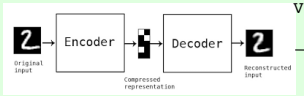
RBM(1986/2006)



stack

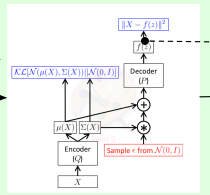
Deep Belief Nets(2006)

AutoEncoder(1989/2006)
Denosing Autoencoder(2008)



Variational Inference

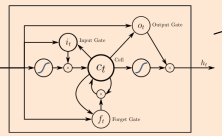
VAE(2013)



beta-VAE(2018)
VQ-VAE-2(2019)

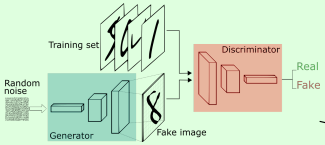
Max Welling
University of Amsterdam
statistical learning

RNN/LSTM(1997)



RNN in Speech Recognition(2013)

GAN(2014)



Ian Goodfellow
Google Brain
deep adversarial learning

DCGAN(2014)
WGAN(2017)
BigGAN(2018)

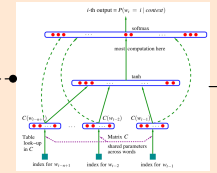
styleGAN(2018)
sinGAN(2019)

IRGAN(2017)
SeqGAN(2017)

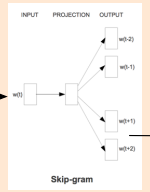
ELECTRA(2019)

Christopher Manning
Stanford University
natural languge process

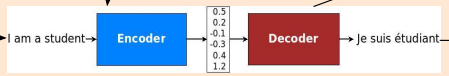
Neural Probabilistic Language Model(2003)



word2Vec(2013)

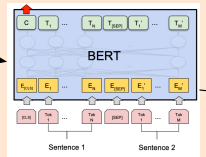


Seq2Seq(2014)



Character CNN(2015)
Transformer(2017)

BERT(2018)

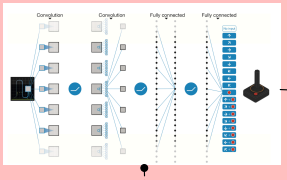


XLNet(2019)
Roberta(2019)
Albert(2019)
T5(2019)

Jürgen Schmidhuber
IDSIA
Universal AI

Yoshua Bengio
University of Montreal
Deep learning

Deep Q-learning(2013)

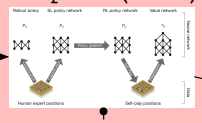


DDPG(2015)

Double DQN(2015)
Dueling Net(2016)

MCTS

AlphaGo(2016)



Alpha Zero(2017)

A3C(2016)

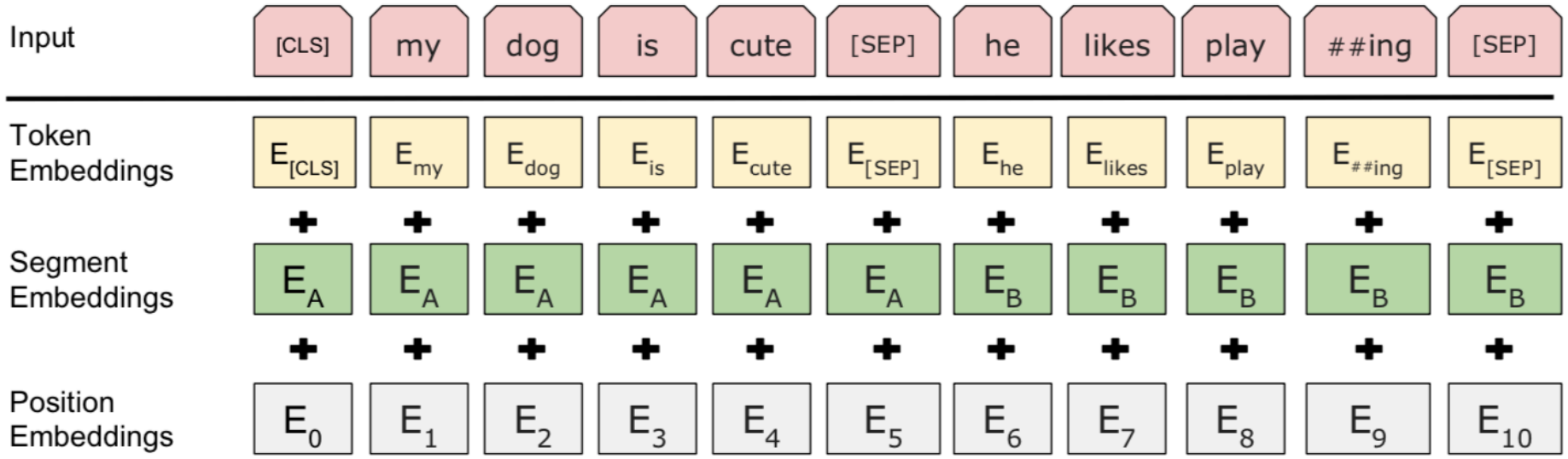
TD3(2018)
PPO(2018)

SAC(2019)

Sergey Levine
UC Berkeley
Robotics

David Silver
DeepMind
Reinforcement learning

预训练模型BERT

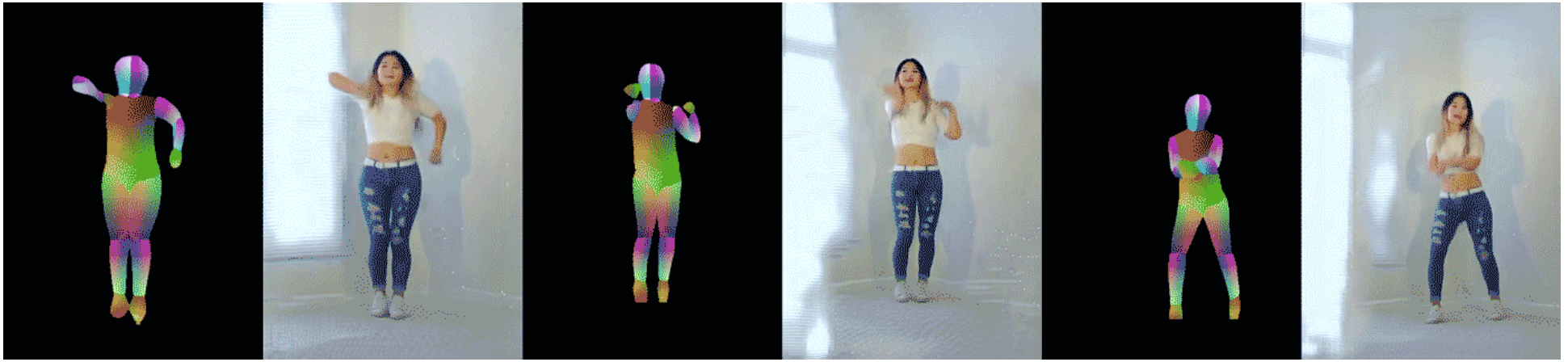


- 预训练Pre-train
- 微调Fine tune
- Beat all state-of-the-arts on 11 NLP tasks in 2018

https://arxiv.org/pdf/1810.04805.pdf

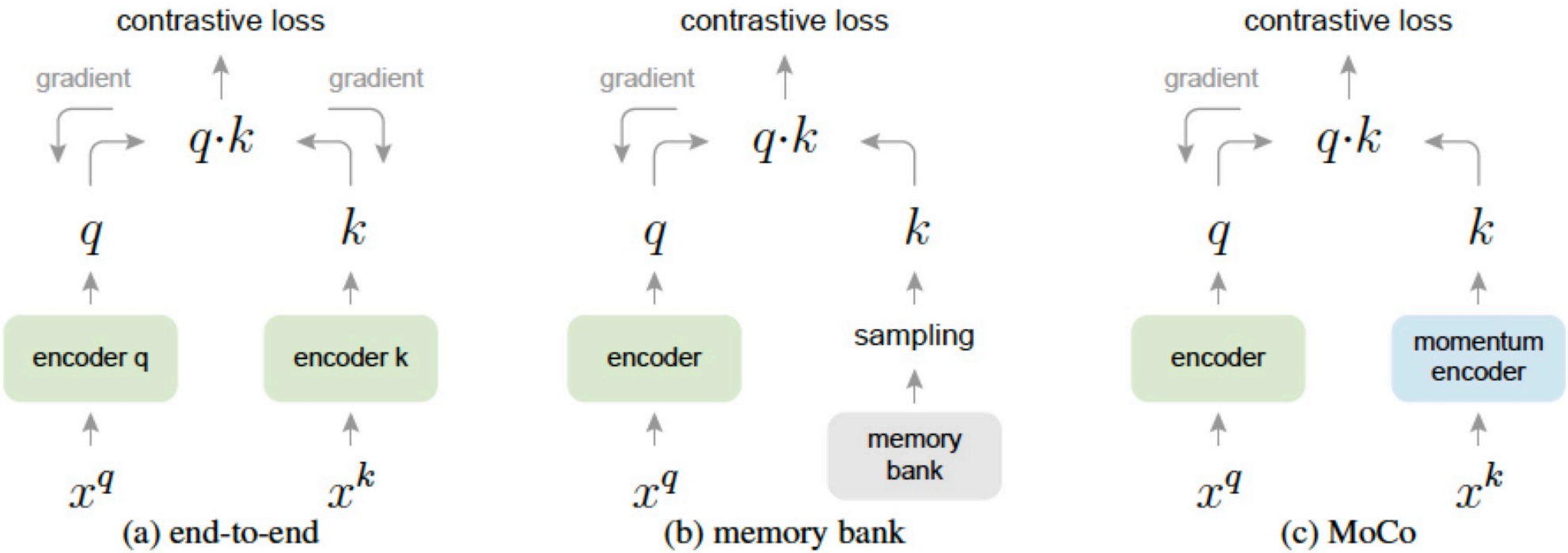
Video-to-Video Synthesis

- The best video synthesis performance



https://arxiv.org/abs/1808.06601

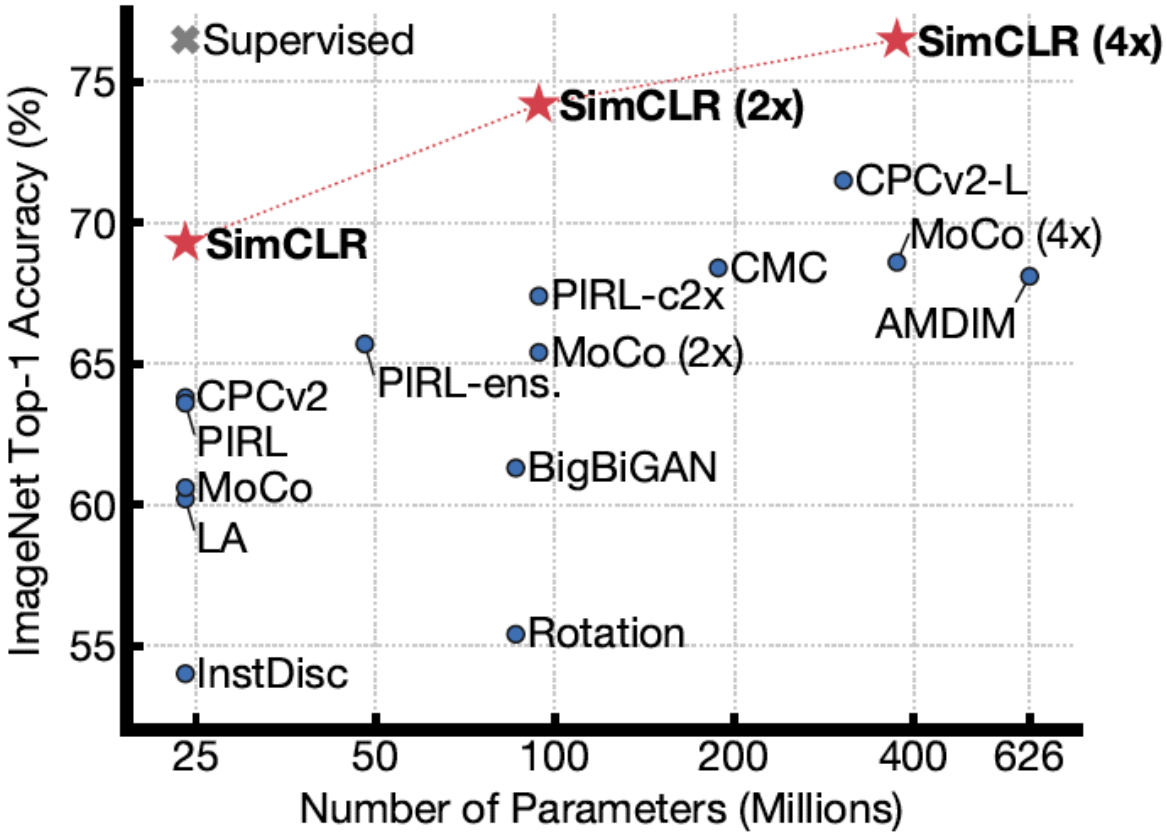
自监督学习— MoCo



- 无需标记样本，即可学习图形表示
- Momentum contrastive learning
- 效果甚至超过有监督学习结果

https://arxiv.org/abs/1911.05722

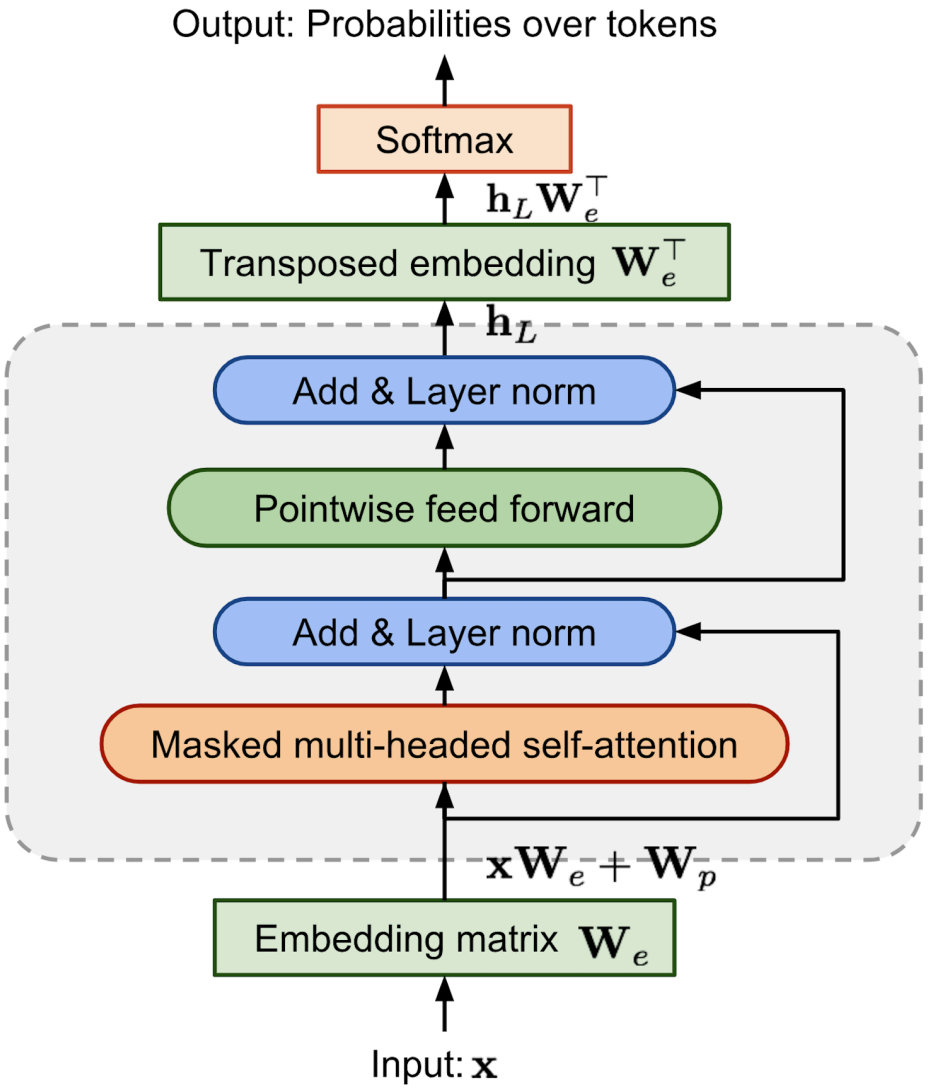
SimCLR



- Simplified contrastive learning framework
- Outperform previous self-supervised and semi-supervised methods on ImageNet

https://arxiv.org/abs/2002.05709

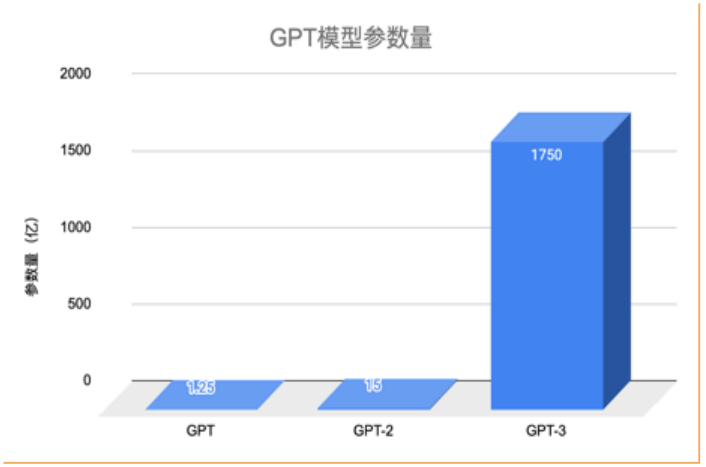
GPT-3



Transformer Block

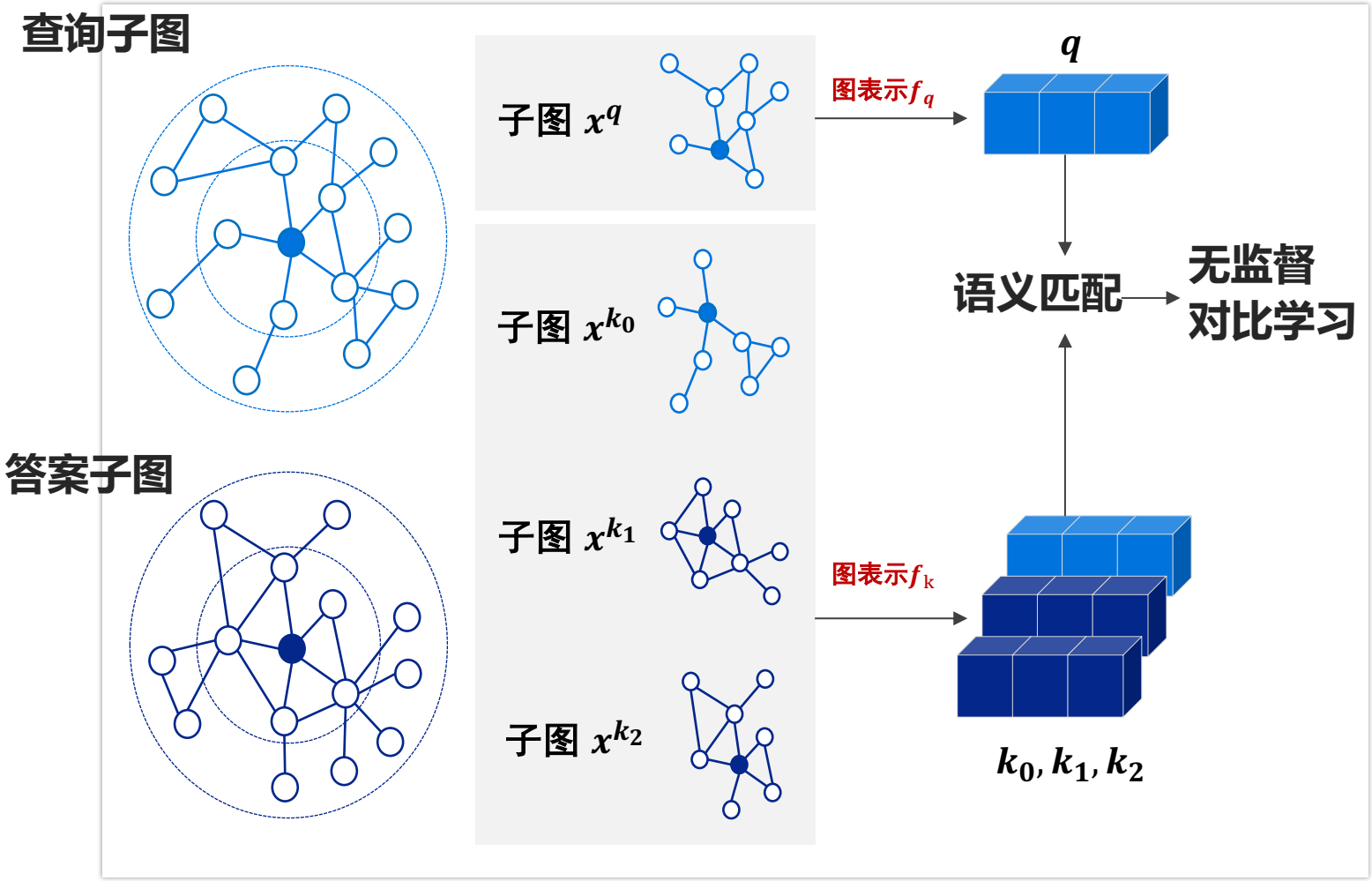
Repeat x L=12

$$\mathbf{h}_\ell = \text{transformer\_block}(\mathbf{h}_{\ell-1})$$
$$\ell = 1, \ldots, L$$



https://lilianweng.github.io/lil-log/2019/01/31/generalized-language-models.html

面向图数据的自监督学习



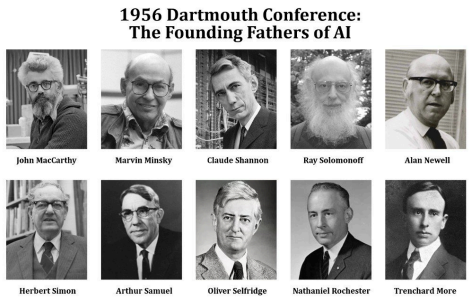
1. J. Qiu, Q. Chen, Y. Dong, J. Zhang, H. Yang, M. Ding, K. Wang, and J. Tang. GCC: Graph Contrastive Coding for Graph Neural Network Pre-Training. *KDD'20*.

人工智能未来...

SEARCH

第三代人工智能

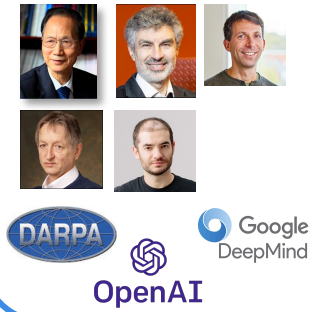
符号AI



符号模型/规则模型/感知机

第一代

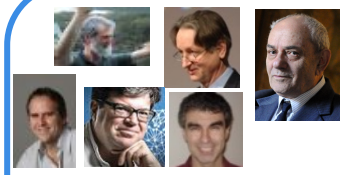
认知智能



张钹院士2016年提出第三代人工智能雏形，DARPA 2018年发布AI Next计划。核心思路是推进统计与知识推理融合的计算；与脑认知的融合

第三代

第二代

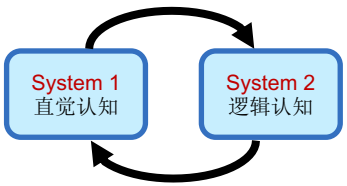


大数据驱动的统计学习方法初步实现了针对文本、图像、语音等的感知与识别

感知智能

目前急需的是高质量超大规模知识图谱（AI的基础设施）以及对知识的理解能力（面向认知的深度学习算法）

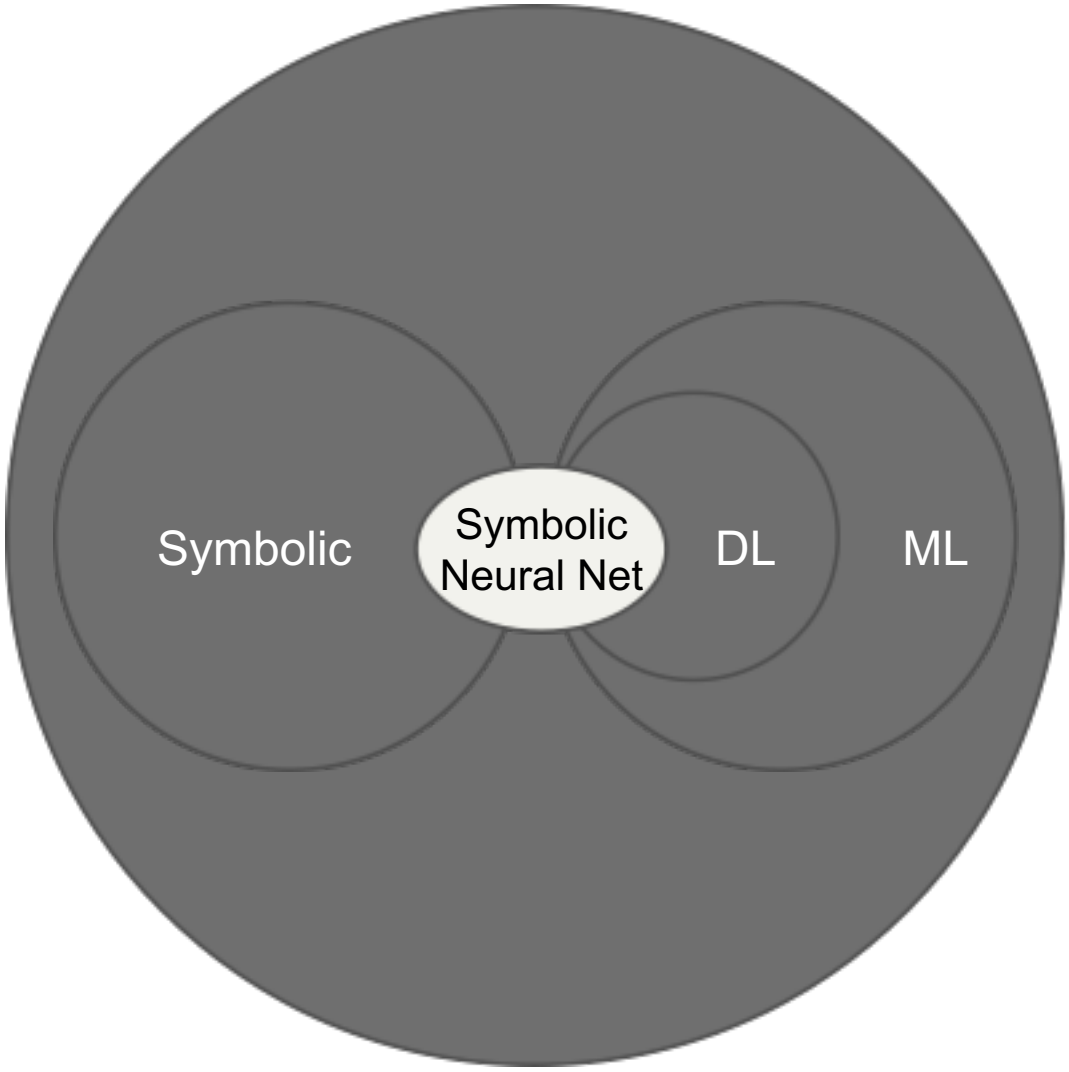
脑认知



第三代人工智能的理论体系

- 早在2015年，张钹老师就提出第三代人工智能体系的雏形；2017年DARPA发起XAI项目，从可解释的机器学习系统、人机交互技术以及可解释的心理学理论三个方面，全面开展可解释性AI系统的研究
- 2018年底，正式公开提出**第三代人工智能的理论框架体系**
  - 建立**可解释、鲁棒性**的人工智能理论和方法
  - 发展**安全、可靠、可信及可扩展**的人工智能技术
  - 推动人工智能**创新**应用
- 具体实施路线图
  - 与**脑科学**融合，发展脑启发的人工智能理论
  - **数据与知识融合**的人工智能理论与方法
- 第三代人工智能的理念在国内外获得广泛影响力

超越深度学习



算法是核心，计算、数据是基础

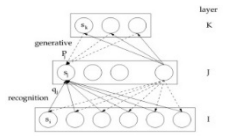
人工智能的核心：算法

算法

实现核心：机器通过**算法**实现人工智能

突破方向：**认知智能**是下一个突破方向

突破途径：人工智能突破主要通过**算法性能的提升**，主要有工程学法和模拟法



方法一：工程学方法

（Engineering Approach）

采用传统的编程技术，利用大量数据处理经验改进提升算法性能。

方法二：模拟法

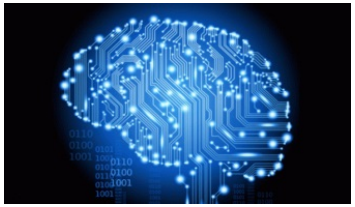
（Modeling Approach）

模仿人类或其他生物所用的方法或机理，提升算法性能，例如遗传算法和神经网络。

基础条件：计算和知识

计算能力

- **现状：**使用GPU并行计算神经网络
- **作用：**提升运算速度，降低计算成本
- **未来：**量子计算、速度更快的芯片



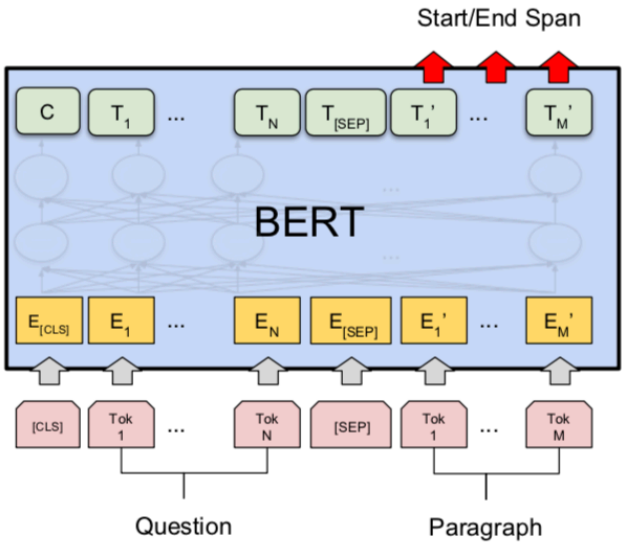
知识

- **现状：**互联网发展积累了片段化知识
- **作用：**训练机器，提升算法性能
- **未来：**面向全世界的常识知识图谱

场景驱动、算法为核心、知识为基础

端到端模型的困境：以BERT为例

- 端到端的黑盒模型缺乏解释性和鲁棒性
  - 不知道推理的逻辑链条
  - 易受无关信息的“对抗攻击”

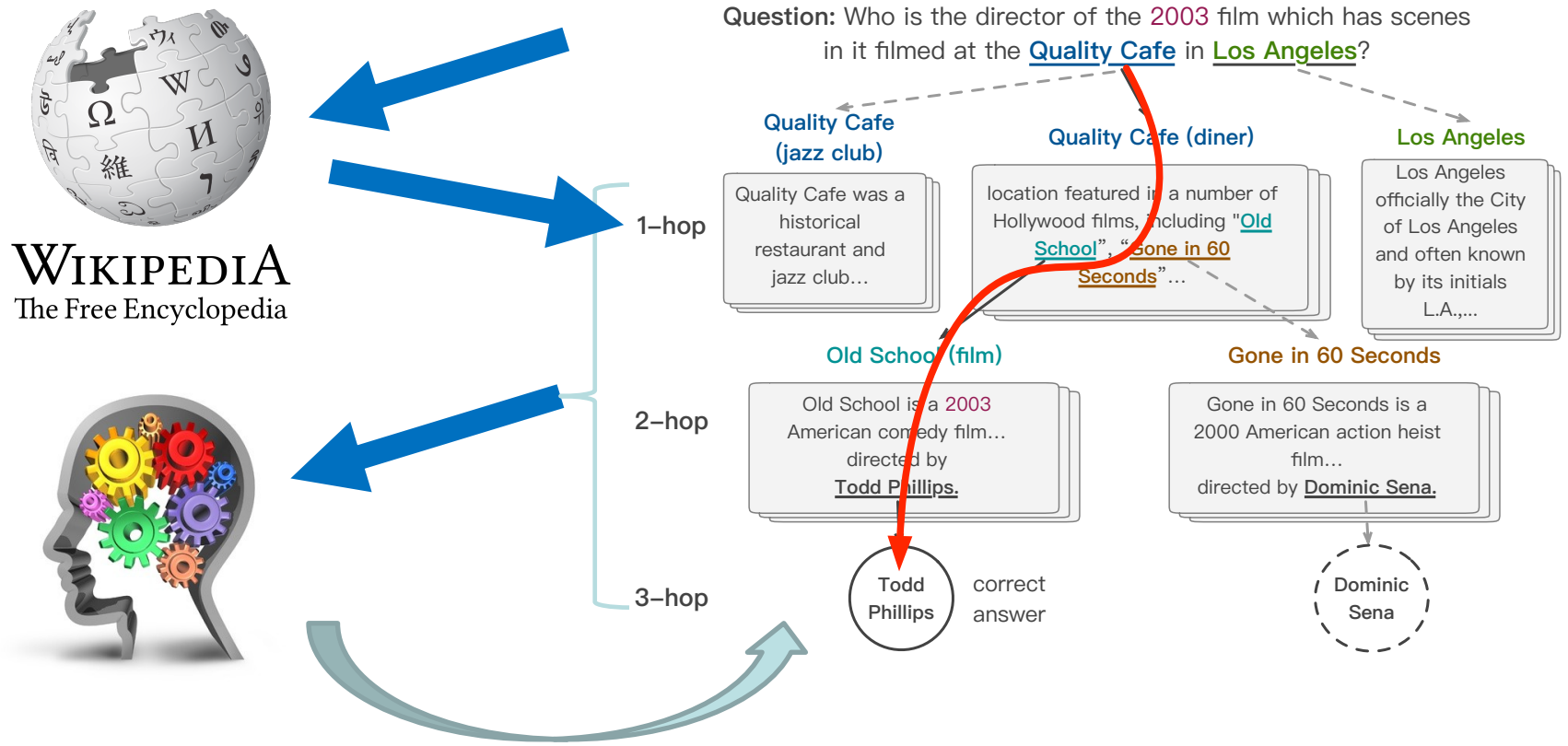


- 无法平衡信息处理精度和规模
  - BERT对于长文本时间空间消耗巨大
  - 直接检索无法考虑多步推理后的相关性

- 为什么人类可以依照逻辑推理，且能利用海量记忆？

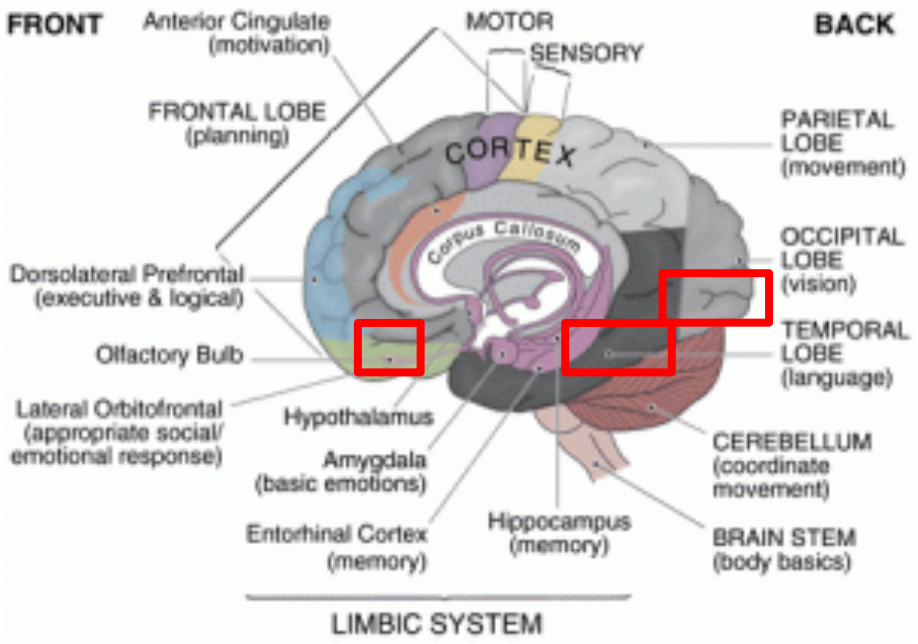
1. Devlin Jacob et. al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL. 2019.

例：复杂阅读理解问答


WIKIPEDIA
The Free Encyclopedia
Question: Who is the director of the 2003 film which has scenes in it filmed at the Quality Cafe in Los Angeles?
1–hop
2–hop
3–hop
Quality Cafe (jazz club)
Quality Cafe was a historical restaurant and jazz club…
Quality Cafe (diner)
location featured in a number of Hollywood films, including "Old School", "Gone in 60 Seconds"…
Los Angeles
Los Angeles officially the City of Los Angeles and often known by its initials L.A.,...
Old School (film)
Old School is a 2003 American comedy film… directed by Todd Phillips.
Gone in 60 Seconds
Gone in 60 Seconds is a 2000 American action heist film… directed by Dominic Sena.
Todd Phillips
correct answer
Dominic Sena

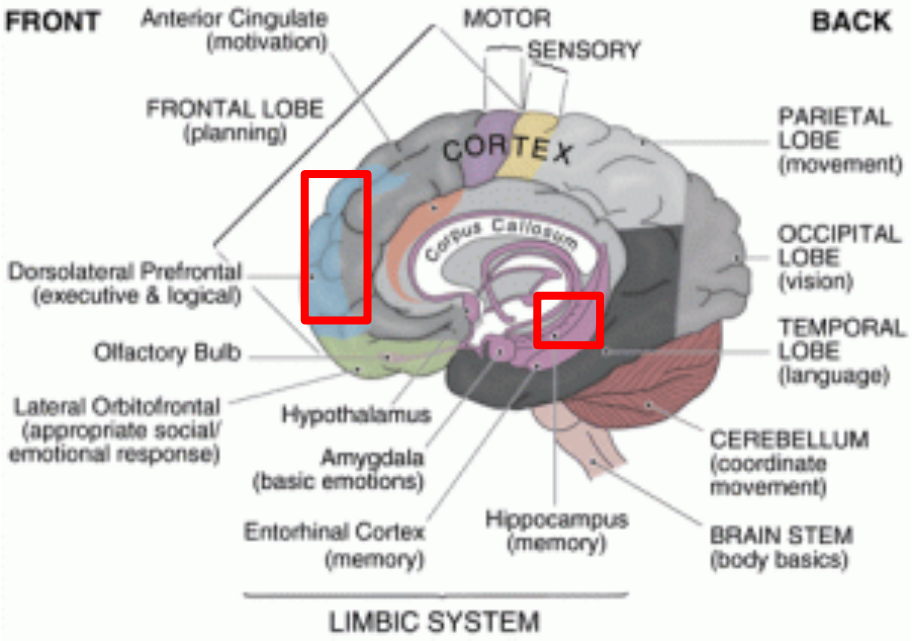
脑科学与感知智能

- 视觉脑区：
  - 主要位于枕叶（V1），通过腹（what）背（where）两条通路进行信息加工。
- 视觉性语言脑区：
  - 顶下小叶的角回（书写与口语转换）
- 听觉性语言脑区：
  - 颞叶的Wernicke’s area
- 语言脑区：
  - Broca’s area（协调发音与语法结构）
- 嗅觉脑区：
  - 嗅球、部分杏仁核
  - ……



脑科学与认知智能

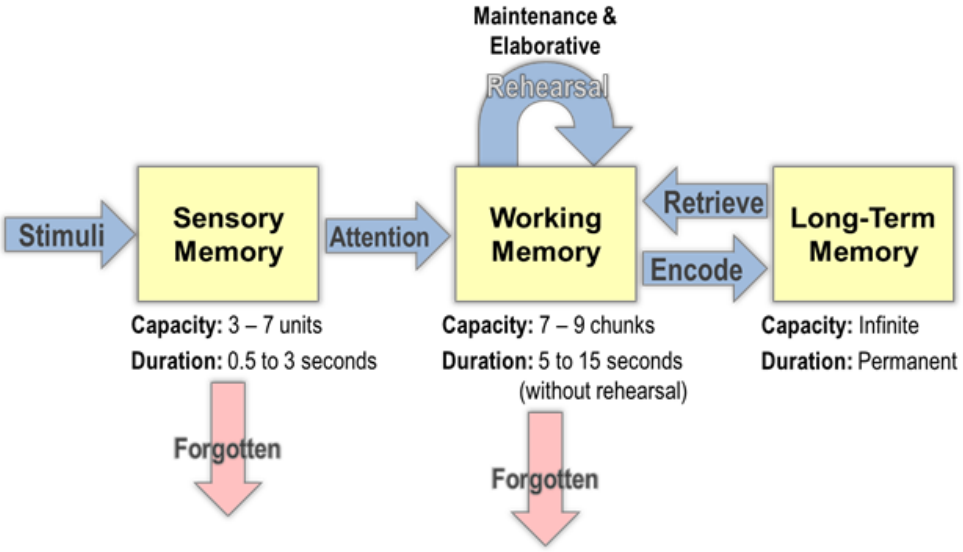
- 记忆
  - 短期记忆在海马体转化为长期（陈述性）记忆
  - 蛋白质磷酸化修饰（中期）
  - 长时程增强、新蛋白产生（长期记忆）
- 推理
  - 主要在前额叶中进行推理
  - 语言脑区参与
  - 对应工作记忆



（睡眠时前额叶兴奋可以做*能推断出自己在做梦*的“清醒梦”）

记忆：工作记忆理论

- 尽管对于认知的微观机理尚未研究清楚，我们仍可以探究宏观框架
- 巴德利的工作记忆（Working Memory）机制是里程碑式的工作，探究工作记忆调用多模态信息与长短期记忆转化（科万的分层注意力理论）



- 全局工作空间理论（Global Workspace Theory）是巴斯等人对工作记忆模型的发展，认为“意识”是不同进程争夺全局空间传播信息的结果

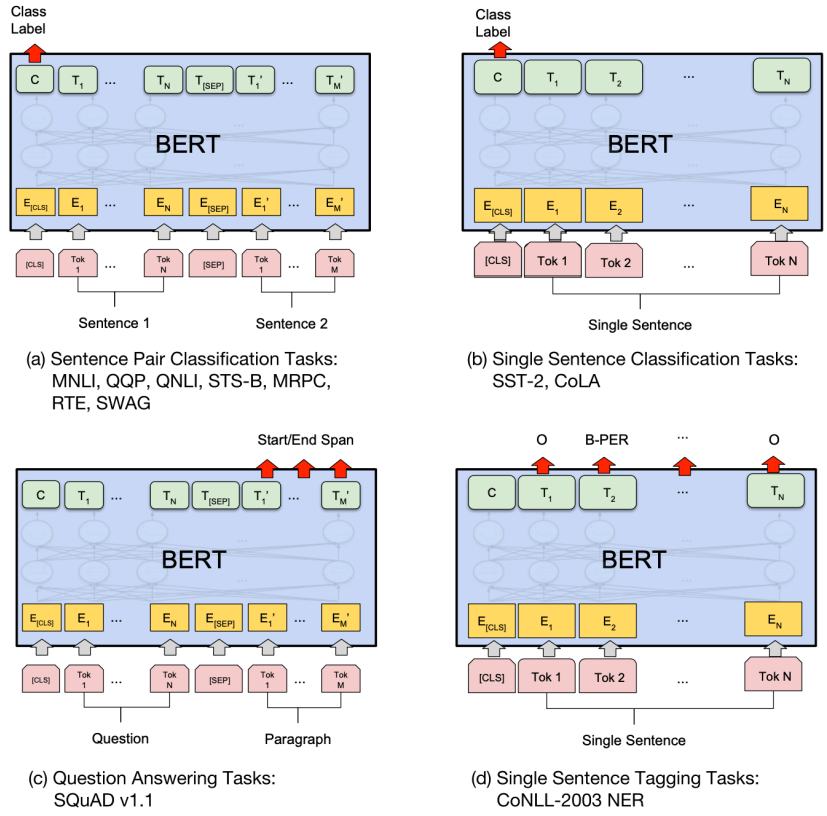
基于记忆机制的相关工作

- 神经网络经典工作LSTM通过门控机制，模拟长期记忆的遗忘与存储，与经典的RNN相比能记住更多步之前的信息。
- Memory Network的一系列工作，提出使用离散的记忆槽位（Memory slots）记录之前计算的隐向量，再通过向量相似度去索引。
- End-to-End Memory Network提出如何端到端学习如何使用记忆槽位记录和检索，并以问答作为例子。
- Dynamic Memory Networks连续通过句子的隐表示来更新目前的“工作记忆向量”。

1. Weston, Jason, Sumit Chopra, and Antoine Bordes. "Memory networks." *arXiv preprint arXiv:1410.3916* (2014).
2. Sukhbaatar, Sainbayar, Jason Weston, and Rob Fergus. "End-to-end memory networks." NIPS. 2015.
3. Kumar, Ankit, et al. "Ask me anything: Dynamic memory networks for natural language processing." ICML. 2016.

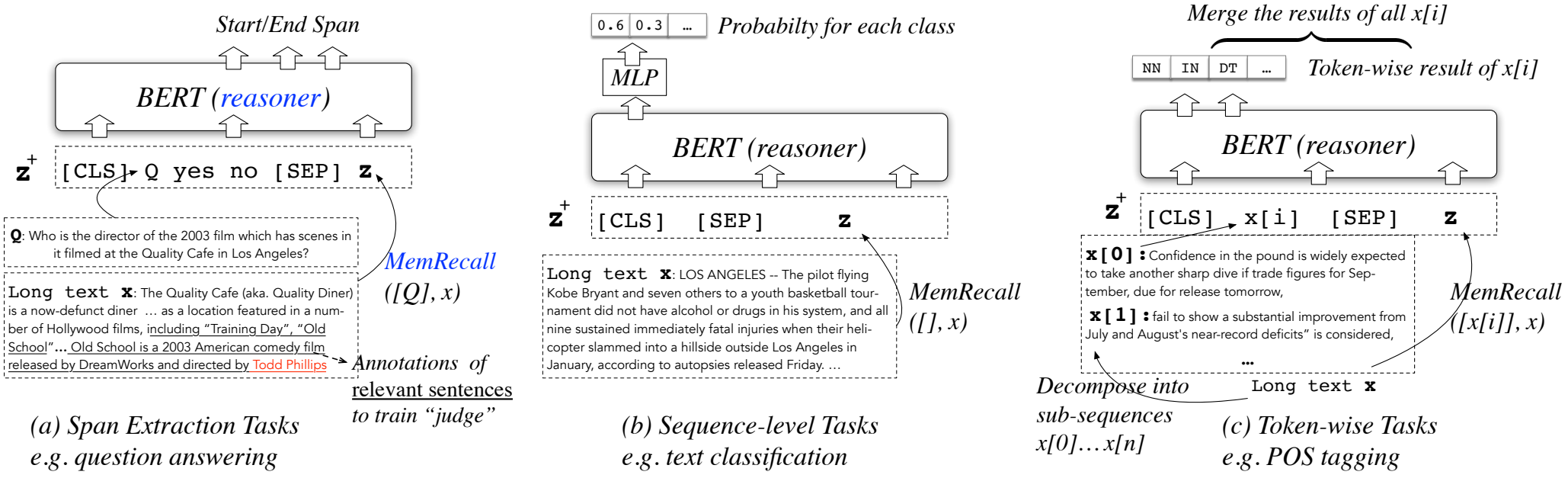
BERT时代新的挑战

- BERT是预训练的Transformer（多层自注意力模型），目前NLP领域最流行的模型之一。
- 训练时内存消耗巨大，并且关于文本长度呈$O(L^2)$增长。
- 简单的分段计算则难以实现长距离的注意力。
- 而人类的工作记忆内存有限，却可以理解长文本，为什么？



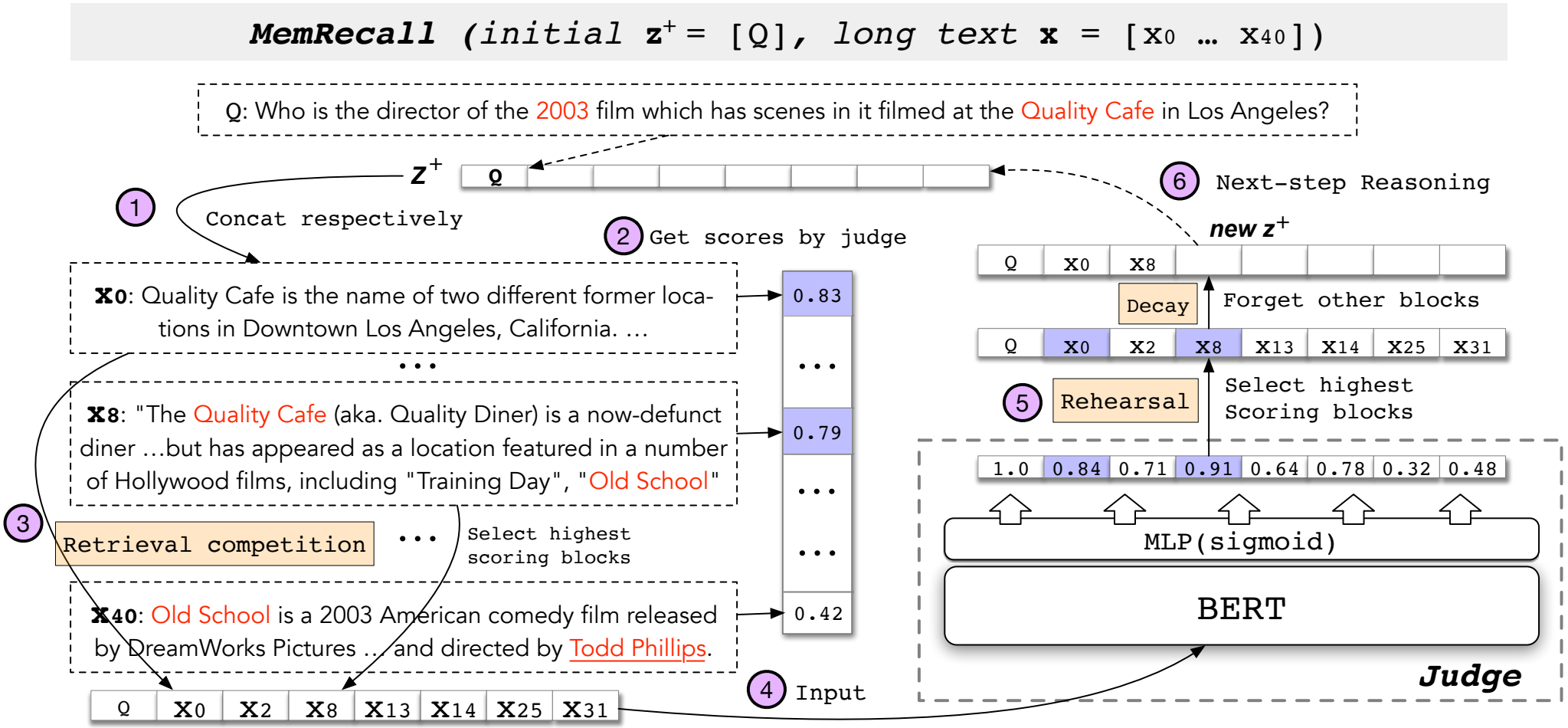
CogLTX: Applying BERT to long texts

- 通过长期记忆到工作记忆的转换，降低BERT处理文本量



1. Ming Ding, Chang Zhou, Hongxia Yang, and Jie Tang. CogLTX: Applying BERT to Long Texts. NeurIPS'20.

MemRecall操作



实验效果

Table 1: NewsQA results (%).

Model	EM	F_1
Match-LSTM [44]	34.9	50.0
BiDAF [37]	37.1	52.3
FastQAExt [47]	42.8	56.1
AMANDA [20]	48.4	63.7
MINIMAL [24]	50.1	63.2
DECAPROP [39]	53.1	66.3
RoBERTa-large [22] (sliding window)	49.6	66.3
CogLTX	**55.2**	**70.1**

Table 2: Results on HotpotQA distractor (dev). (+hyperlink) means usage of extra hyperlink data in Wikipedia. Models beginning with "–" are ablation studies without the corresponding design.

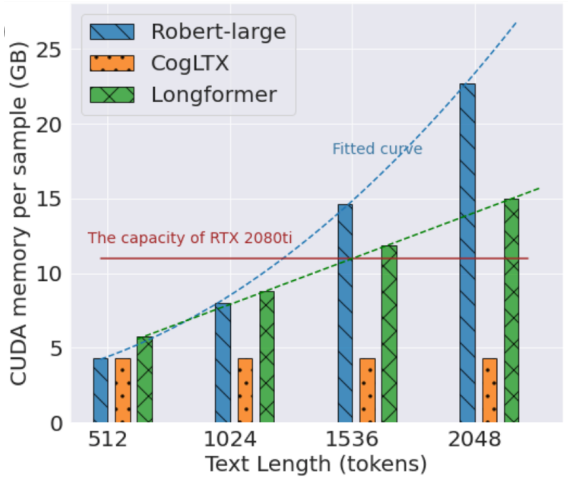
Model	Ans EM	Ans F_1	Sup EM	Sup F_1	Joint EM	Joint F_1
Baseline [49]	45.60	59.02	20.32	64.49	10.83	40.16
DecompRC [25]	55.20	69.63	N/A	N/A	N/A	N/A
QFE [26]	53.86	68.06	57.75	84.49	34.63	59.61
DFGN [32]	56.31	69.69	51.50	81.62	33.62	59.82
SAE [41]	60.36	73.58	56.93	84.63	38.81	64.96
SAE-large	66.92	79.62	61.53	86.86	45.36	**71.45**
HGN [13] (+hyperlink)	66.07	79.36	60.33	87.33	43.57	71.03
HGN-large (+hyperlink)	69.22	82.19	62.76	88.47	47.11	**74.21**
BERT (sliding window) variants						
BERT Plus	55.84	69.76	42.88	80.74	27.13	58.23
LQR-net + BERT	57.20	70.66	50.20	82.42	31.18	59.99
GRN + BERT	55.12	68.98	52.55	84.06	32.88	60.31
EPS + BERT	60.13	73.31	52.55	83.20	35.40	63.41
LQR-net 2 + BERT	60.20	73.78	56.21	84.09	36.56	63.68
P-BERT	61.18	74.16	51.38	82.76	35.42	63.79
EPS + BERT(large)	63.29	76.36	58.25	85.60	41.39	67.92
CogLTX	65.09	78.72	56.15	85.78	39.12	69.21
– multi-step reasoning	62.00	75.39	51.74	83.10	35.85	65.35
– rehearsal & decay	61.44	74.99	7.74	47.37	5.36	37.74
– train-test matching	63.20	77.21	52.57	84.21	36.11	66.90

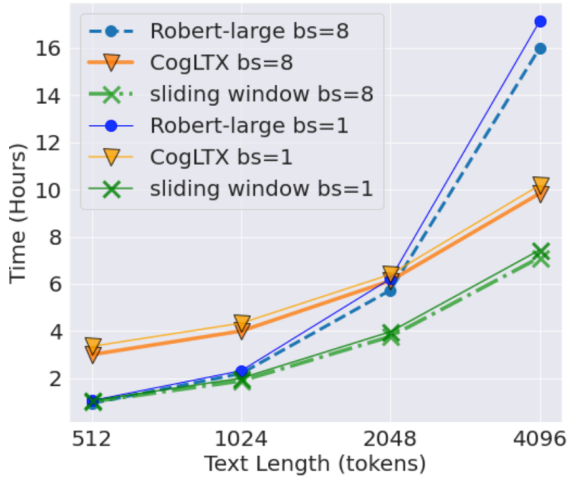
Table 3: 20NewsGroups results (%).

Model	Accuracy
BoW + SVM	63.0
Bi-LSTM	73.2
fastText [15]	79.4
MS-CNN [28]	86.1
Text GCN [50]	86.3
MLP over BERT [29]	85.5
LSTM over BERT [29]	84.7
CogLTX (Glove init)	**87.0**
only long texts	87.4
– intervention (Glove init)	84.8
Bm25 init	86.1

Table 4: A+ result (%).

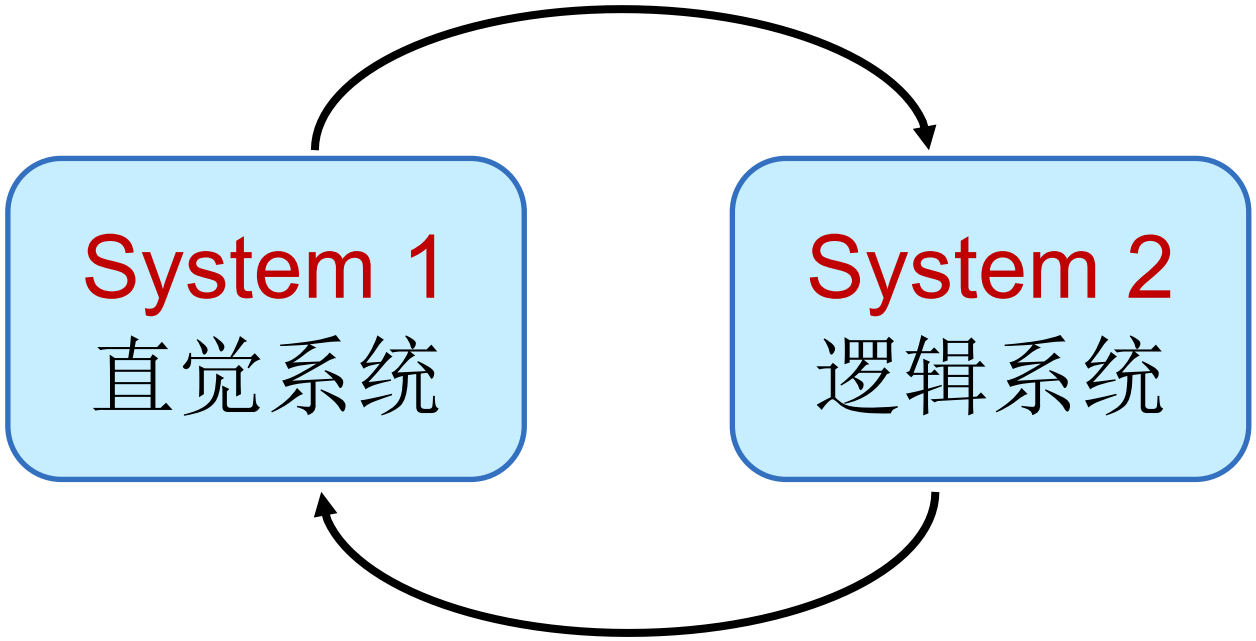
Model	Accuracy	Micro-F_1	Macro-F_1
BoW+SVM	89.9	85.8	55.3
Bi-LSTM	70.7	62.1	48.2
TextCNN	95.3	94.1	91.3
sliding window	94.5	92.7	89.9
CogLTX(tiny)	95.5	94.4	92.4
CogLTX(large)	**98.2**	**97.8**	**97.2**





1. Ming Ding, Chang Zhou, Hongxia Yang, and Jie Tang. CogLTX: Applying BERT to Long Texts. NeurIPS'20.

和认知科学的结合



Dual Process Theory (Cognitive Science)

脑认知的双系统

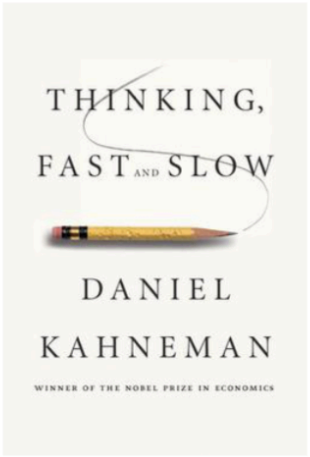
SYSTEM 1 VS. SYSTEM 2 COGNITION

2 systems (and categories of cognitive tasks):

Manipulates high-level / semantic concepts, which can be recombined combinatorially

System 1

- Intuitive, fast, **UNCONSCIOUS**, non-linguistic, habitual
- Current DL



System 2

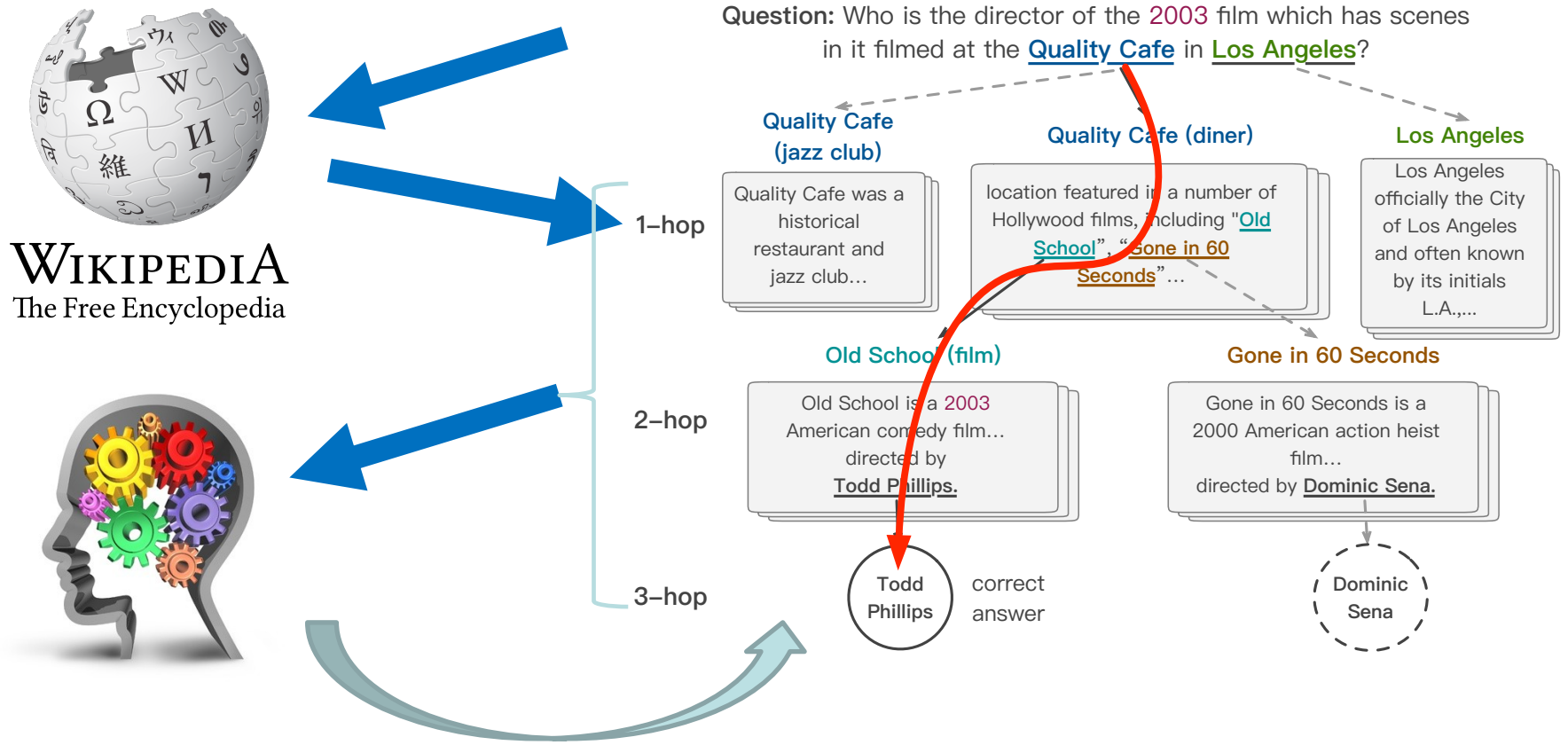
- Slow, logical, sequential, **CONSCIOUS**, linguistic, algorithmic, planning, reasoning
- Future DL

Mila

3

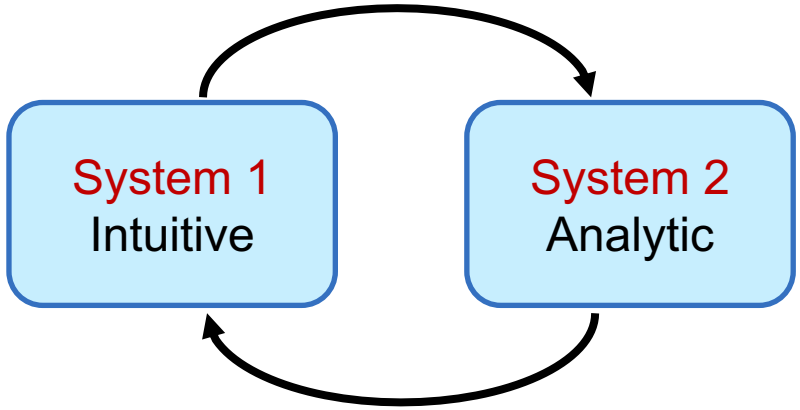
1. From Bengio's NIPS'2019 Keynote

认知图谱：知识表示，推理和决策



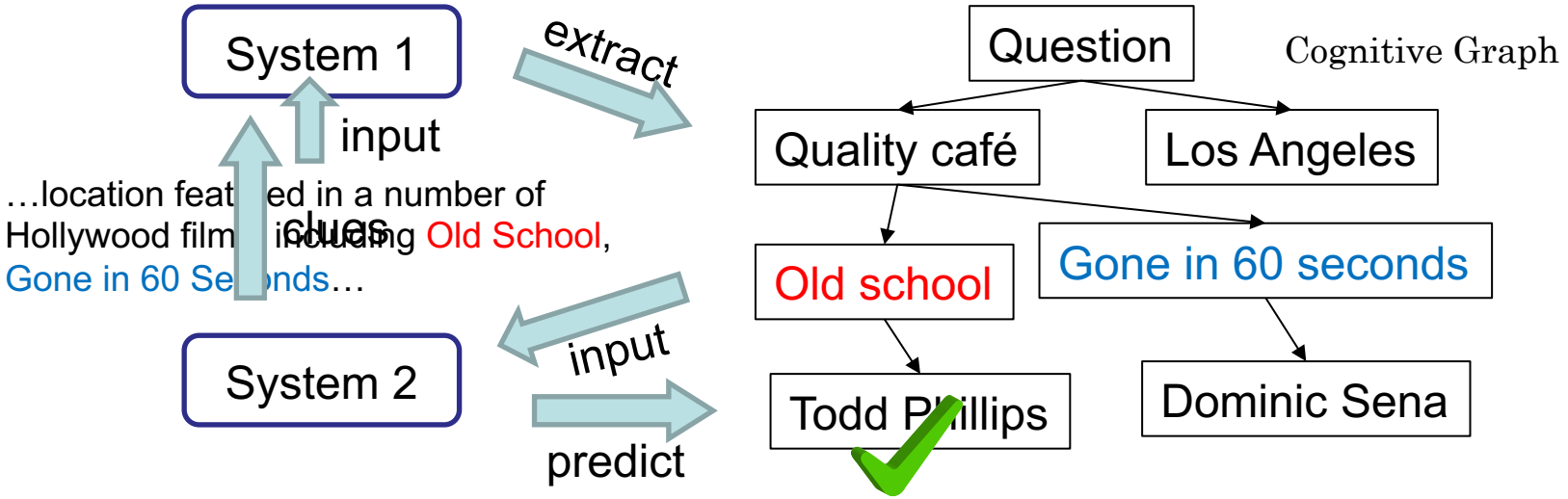
认知图谱的推理模型

- System 1:
  - Knowledge expansion by association in text when reading
- System 2:
  - Decision making w/ all the information

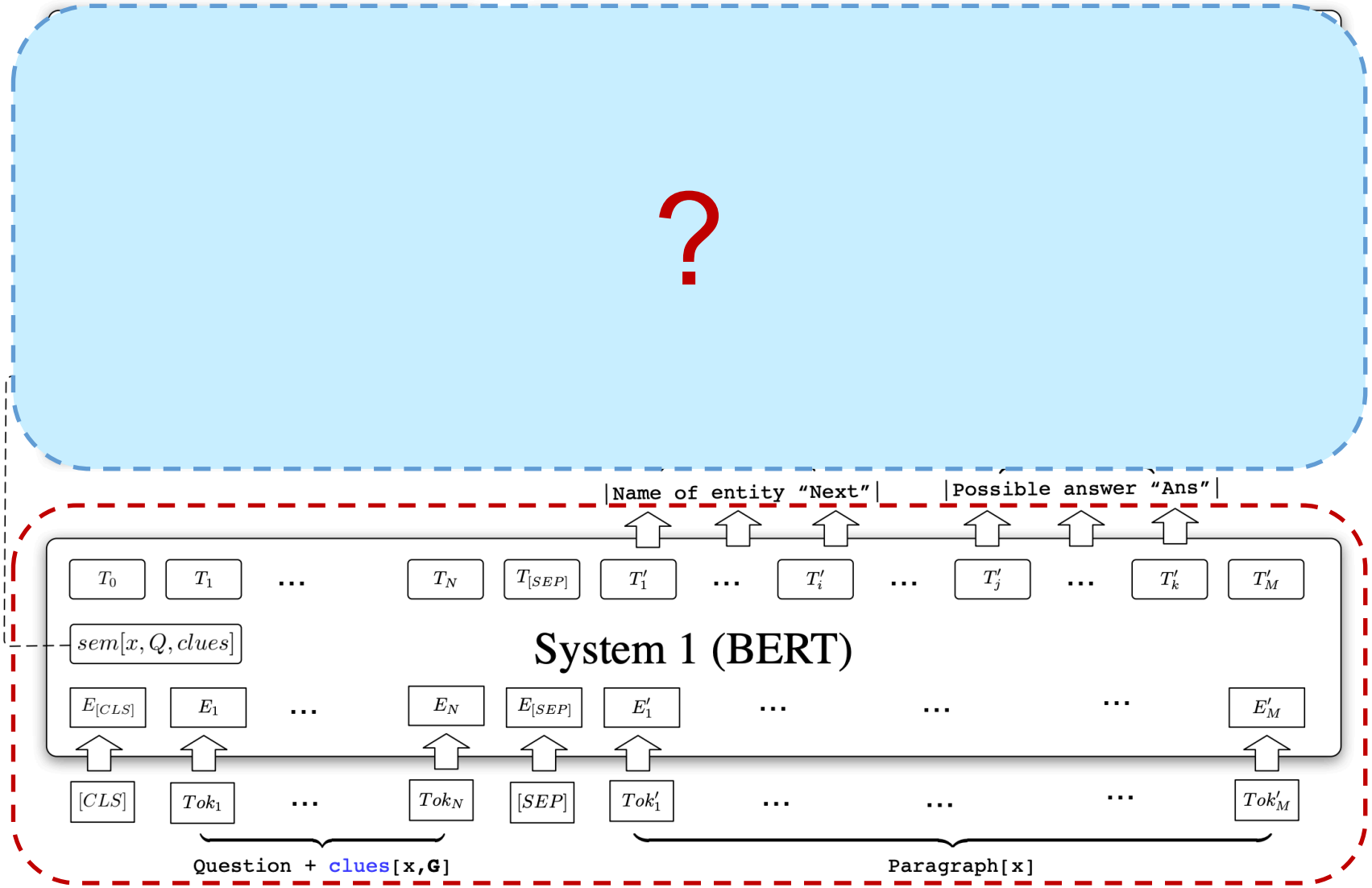


认知图谱的推理模型

- An iterative framework corresponding to dual process theory
- System 1
  - extract entities to build the cognitive graph
  - generate semantic vectors for each node
- System 2
  - Do reasoning based on semantic vectors and graph
  - Feed clues to System 1 to extract next-hop entities

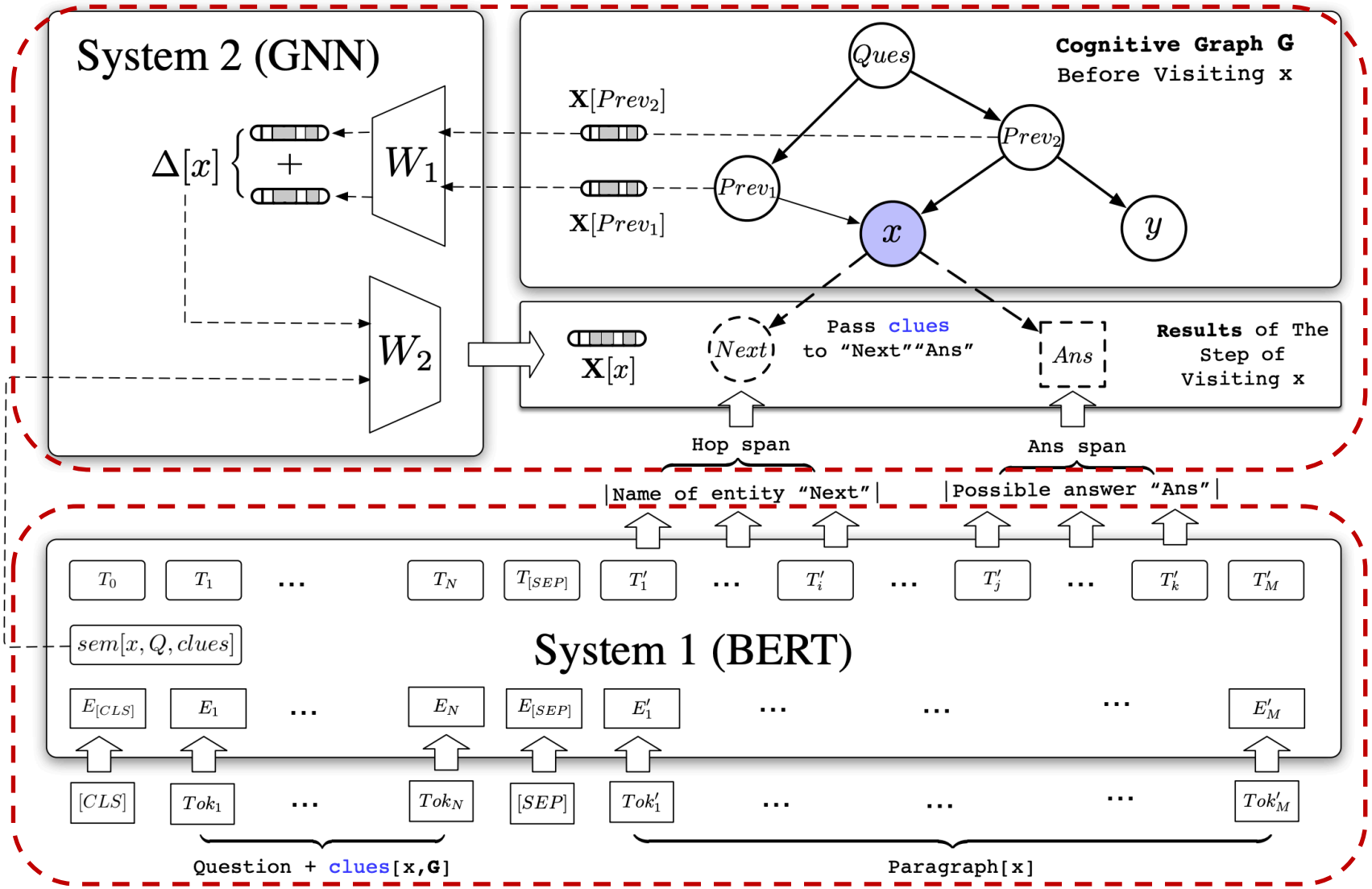


认知图谱的推理模型



1. M. Ding, C. Zhou, Q. Chen, H. Yang, and J. Tang. Cognitive Graph for Multi-Hop Reading Comprehension at Scale. ACL'19.

认知图谱的推理模型



1. M. Ding, C. Zhou, Q. Chen, H. Yang, and J. Tang. Cognitive Graph for Multi-Hop Reading Comprehension at Scale. ACL'19.

认知图谱的推理效果

- HotpotQA is a dataset with leaderboard similar to SQuAD
- CogQA ranked 1st from 21, Feb to 15, May (nearly 3 month)

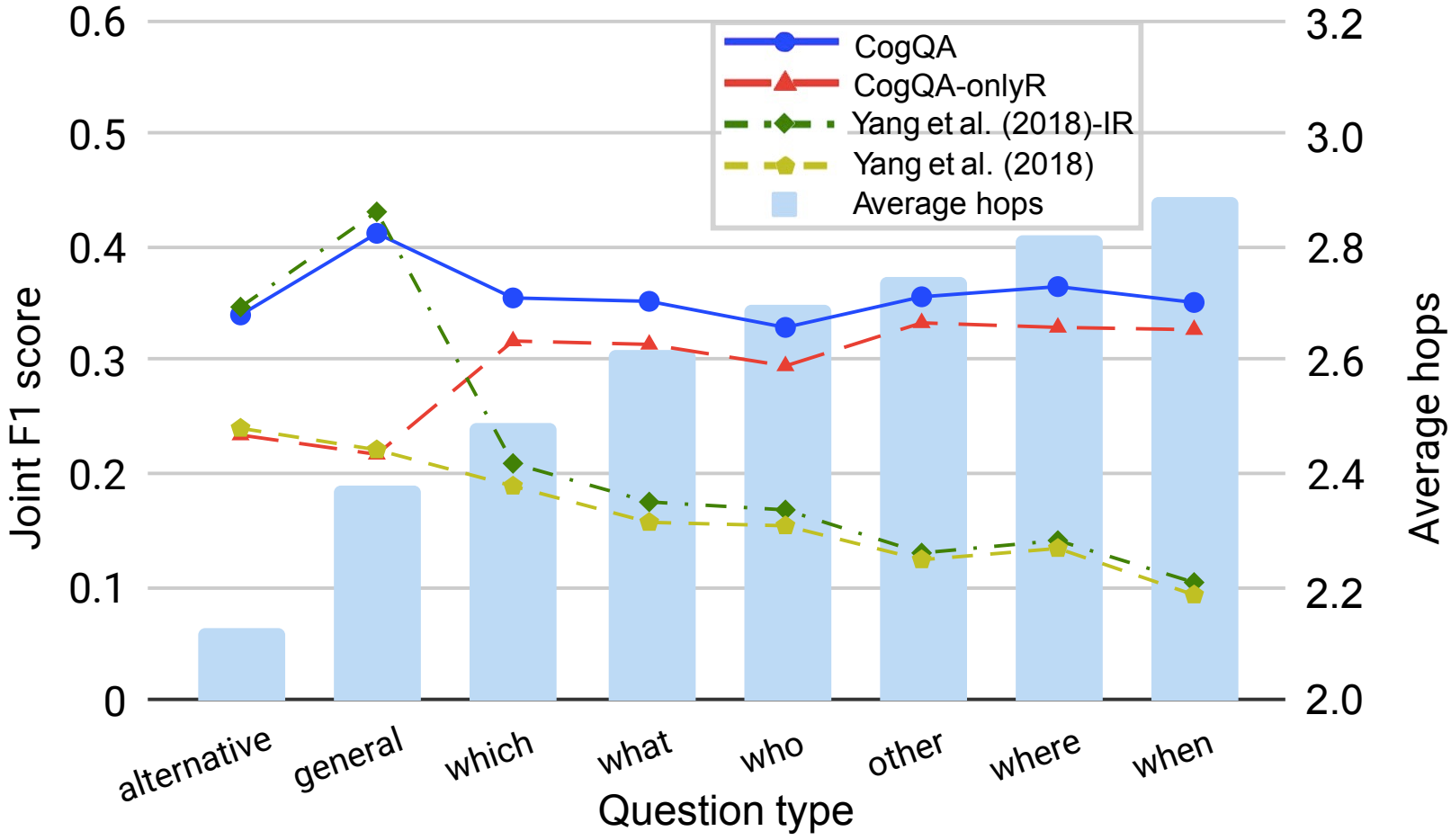
	Model	Ans				Sup				Joint			
		EM	F_1	Prec	Recall	EM	F_1	Prec	Recall	EM	F_1	Prec	Recall
Dev	Yang et al. (2018)	23.9	32.9	34.9	33.9	5.1	40.9	47.2	40.8	2.5	17.2	20.4	17.8
	Yang et al. (2018)-IR	24.6	34.0	35.7	34.8	10.9	49.3	52.5	52.1	5.2	21.1	22.7	23.2
	BERT	22.7	31.6	33.4	31.9	6.5	42.4	54.6	38.7	3.1	17.8	24.3	16.2
	CogQA-sys1	33.6	45.0	47.6	45.4	**23.7**	58.3	**67.3**	56.2	12.3	32.5	39.0	31.8
	CogQA-onlyR	34.6	46.2	48.8	46.7	14.7	48.2	56.4	47.7	8.3	29.9	36.2	30.1
	CogQA-onlyQ	30.7	40.4	42.9	40.7	23.4	49.9	56.5	48.5	**12.4**	30.1	35.2	29.9
	CogQA	**37.6**	**49.4**	**52.2**	**49.9**	23.1	**58.5**	64.3	**59.7**	12.2	**35.3**	**40.3**	**36.5**
Test	Yang et al. (2018)	24.0	32.9	-	-	3.86	37.7	-	-	1.9	16.2	-	-
	QFE	28.7	38.1	-	-	14.2	44.4	-	-	8.7	23.1	-	-
	DecompRC	30.0	40.7	-	-	N/A	N/A	-	-	N/A	N/A	-	-
	MultiQA	30.7	40.2	-	-	N/A	N/A	-	-	N/A	N/A	-	-
	GRN	27.3	36.5	-	-	12.2	48.8	-	-	7.4	23.6	-	-
	CogQA	**37.1**	**48.9**	-	-	**22.8**	**57.7**	-	-	**12.4**	**34.9**	-	-

Table 1: Results on HotpotQA (fullwiki setting). The test set is not public. The maintainer of HotpotQA only offers EM and F_1 for every submission. N/A means the model cannot find supporting facts.

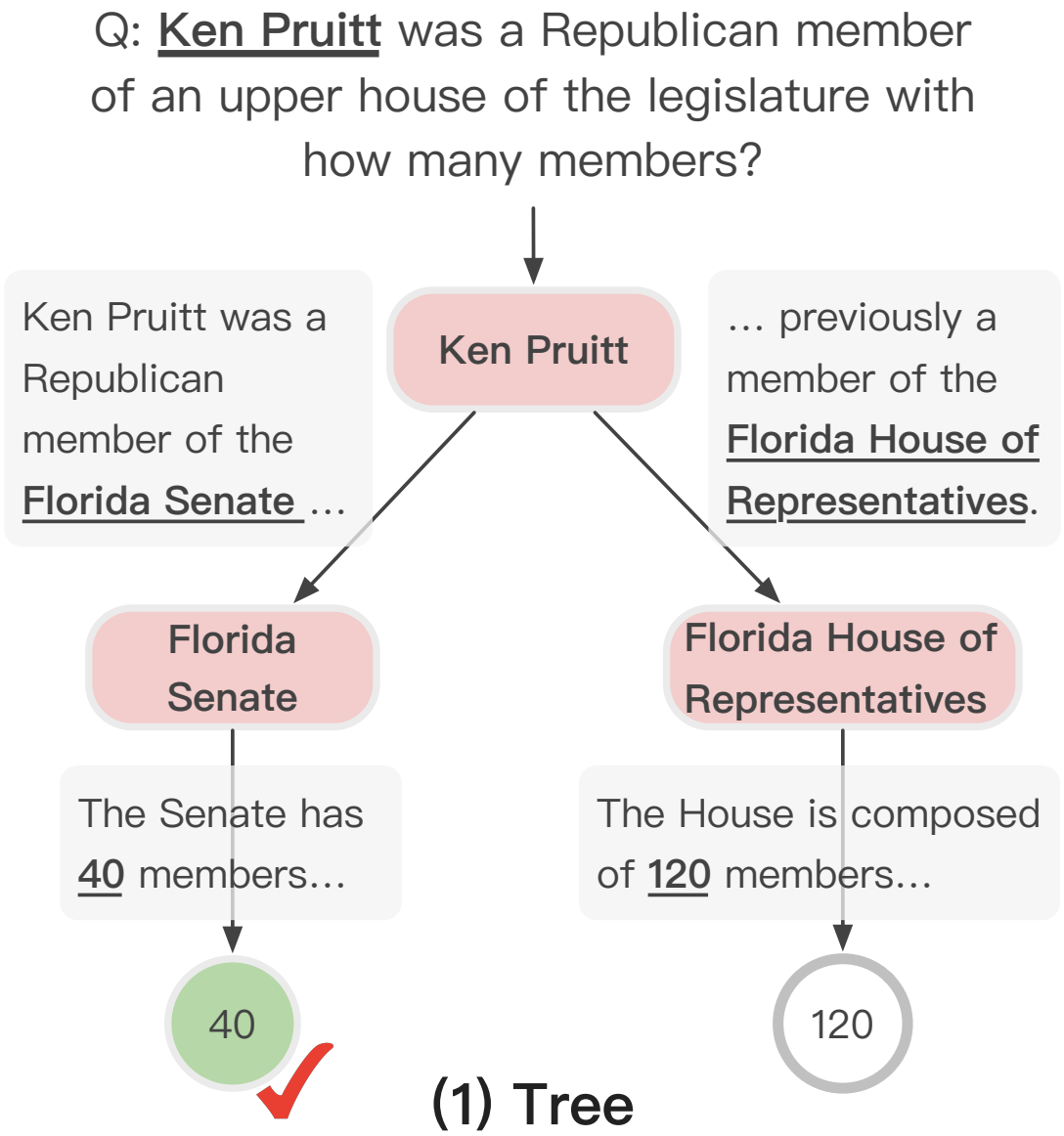
** Code available at https://github.com/THUDM/CogQA

认知图谱的推理效果

CogQA Performs much better on question with more hops !

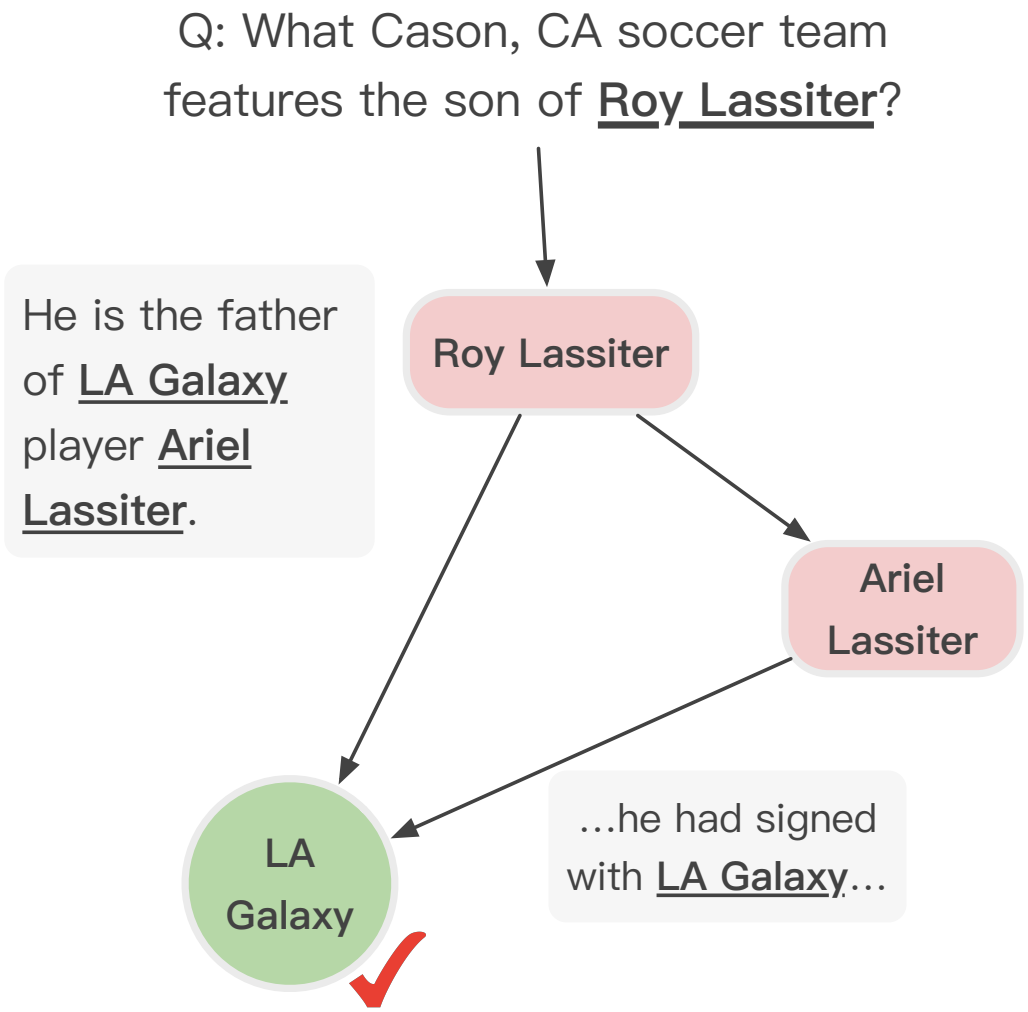


认知图谱的推理效果



- Tree-shape Cognitive Graph
- Users can verify the answer by comparing it with another possible reasoning chain.
- "*Upper House*" in the question is similar to "*Senate*" not "*House of Representative*"

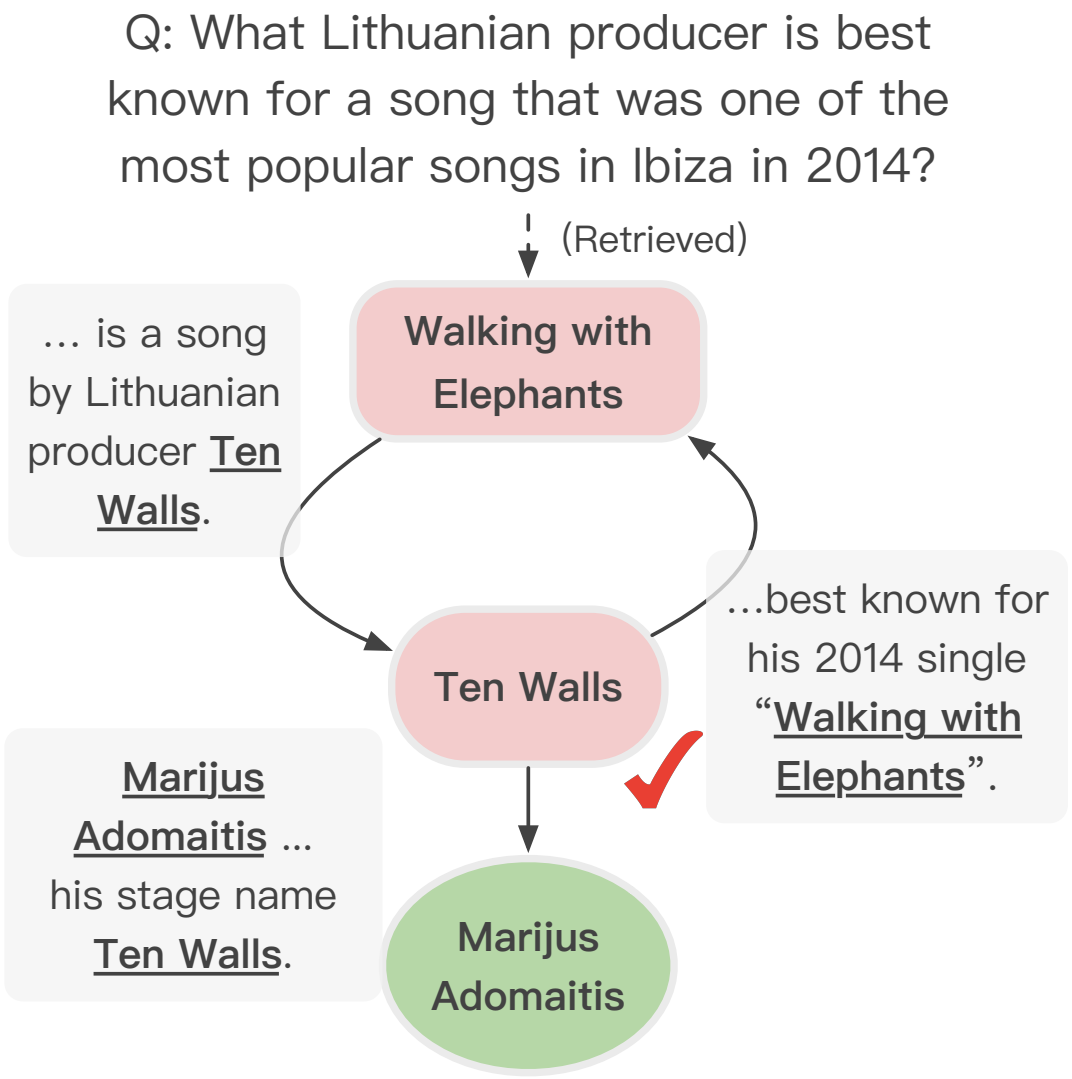
认知图谱的推理效果



(2) DAG

- DAG-shape Cognitive Graph
- Multiple supporting facts provides richer information, increasing the credibility of the answer.

认知图谱的推理效果



(3) Cyclic Graph

- CogQA gives the answer "Marijus Adomaitis" while the ground truth is "Ten Walls".
- By examining, Ten Walls is just the stage name of Marijus Adomaitis!
- Without cognitive graphs, black-box models cannot achieve it.

Next

挑战与未来(Next 10)

Edward Feigenbaum
Turing Award Winner

认知与推理

—Trillion-scale common-sense knowledge graph

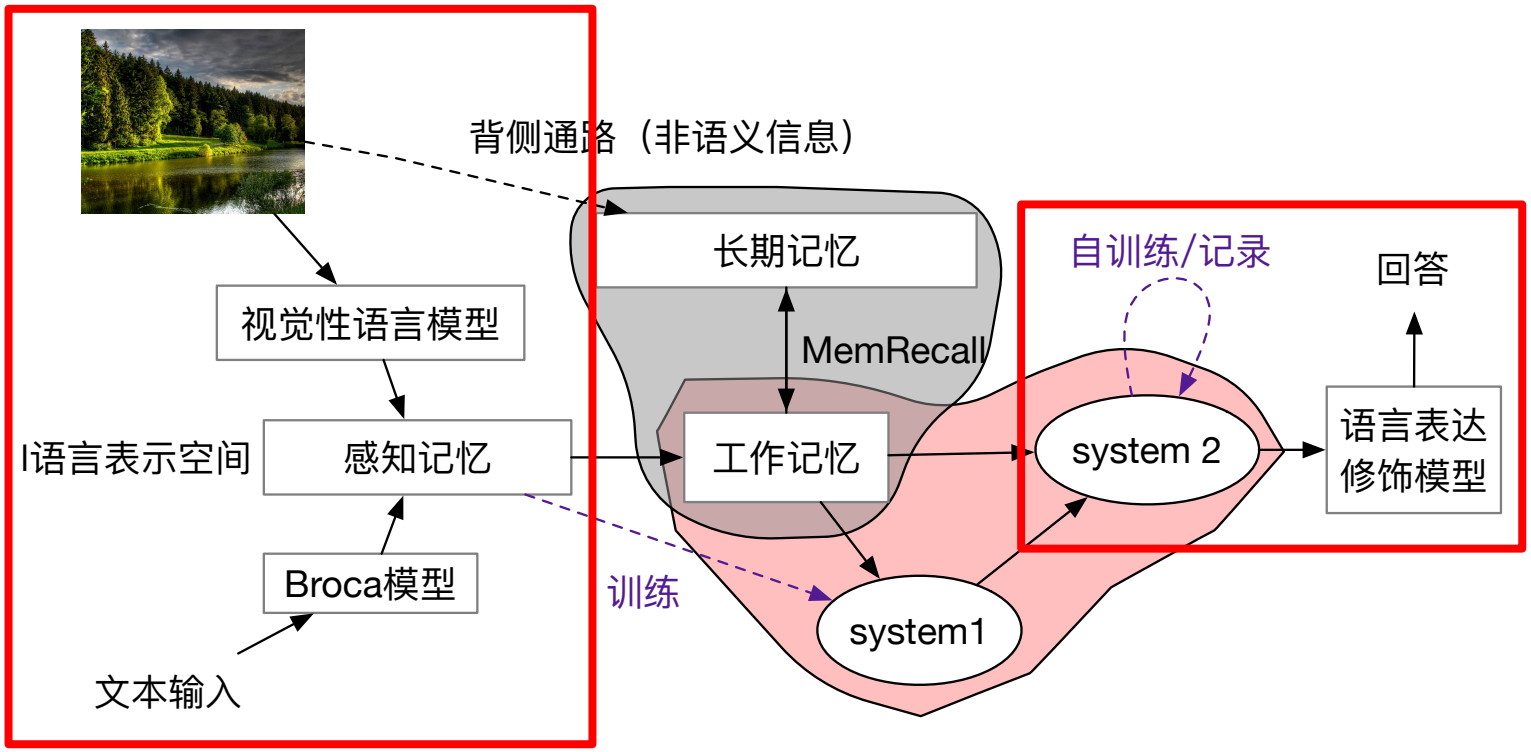
Tim Berners Lee
Turing Award Winner

大数据 ⇨ 知识 ⇨ 智能

* AI = Knowledge + Intelligence

1. J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. ArnetMiner: Extraction and Mining of Academic Social Networks. KDD'08. pp.990-998.

面向计算的认知框架


背侧通路（非语义信息）
长期记忆
自训练/记录
回答
视觉性语言模型
MemRecall
l语言表示空间
感知记忆
工作记忆
system 2
语言表达
修饰模型
Broca模型
训练
system1
文本输入

挑战与未来(Next 30)

意识

—让计算机具有自我意识

认知推理 ⇨ 记忆 ⇨ 自我意识

- Next AI = Reasoning + Memory + Consciousness

1. J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. ArnetMiner: Extraction and Mining of Academic Social Networks. KDD'08. pp.990-998.

意识：全局工作理论—GWT

所谓的有意识学习，就是有目标的机器学习，给定数据，训练一个模型，通过模型进行分类（决策），所以有意识学习输出是一个短期记忆模型，但这个短期记忆模型比较简单：数据有限、无背景知识。所以无意识就对应着长期记忆模型，长期记忆模型有点类似半监督或者无监督学习模型，或者当下比较流行的预训练和自监督学习；无意识处理对应多种长期记忆，所以无意识可以考虑多个不同的处理器，这些处理器之间可以有链接，也可以没有，很多时候是并行处理，但针对特定任务，比如有意识思考某个问题的时候，形成特定连接，包括无意识处理器（无监督模型）和有意识处理器（有监督模型）之间的连接，这里可以考虑成fine-tune。当然连接权重可以通过外界反馈强化学习来实现。无意识处理器之间的连接以及和有意识处理器之间的连接可以类比为注意力机制。最后值得注意的是长期记忆的构造和实现，人脑记忆保存的是模型图，而不是概念图。每个长期记忆都可能是一个模型，可以生成样本，具体学习方法，可以想象一下是一个层次聚类。通过这样就可以用有监督、无监督、强化、注意力、fine-tune来实现GWT模型。

Related Publications

For more, check http://keg.cs.tsinghua.edu.cn/jietang

- Wenzheng Feng, Jie Zhang, Yuxiao Dong, Yu Han, Huanbo Luan, Qian Xu, Qiang Yang, Evgeny Kharlamov, and Jie Tang. Graph Random Neural Networks for Semi-Supervised Learning on Graphs. NeurIPS'20.
- Ming Ding, Chang Zhou, Hongxia Yang, and Jie Tang. CogLTX: Applying BERT to Long Texts. NeurIPS'20.
- Jiezhong Qiu, Chi Wang, Ben Liao, Richard Peng, and Jie Tang. Concentration Bounds for Co-occurrence Matrices of Markov Chains. NeurIPS'20.
- Xiao Liu, Fanjin Zhang, Zhenyu Hou, Li Mian, Zhaoyu Wang, Jing Zhang, and Jie Tang. Self-supervised Learning: Generative or Contrastive. https://arxiv.org/pdf/2006.08218.pdf
- Jiezhong Qiu, Qibin Chen, Yuxiao Dong, Jing Zhang, Hongxia Yang, Ming Ding, Kuansan Wang, and Jie Tang. GCC: Graph Contrastive Coding for Structural Graph Representation Pre-Training. KDD'20.
- Zhen Yang, Ming Ding, Chang Zhou, Hongxia Yang, Jingren Zhou, and Jie Tang. Understanding Negative Sampling in Graph Representation Learning. KDD'20.
- Yukuo Cen, Jianwei Zhang, Xu Zou, Chang Zhou, Hongxia Yang, and Jie Tang. Controllable Multi-Interest Framework for Recommendation. KDD'20.
- Yuxiao Dong, Ziniu Hu, Kuansan Wang, Yizhou Sun and Jie Tang. Heterogeneous Network Representation Learning. IJCAI'20.
- Ming Ding, Chang Zhou, Qibin Chen, Hongxia Yang, and Jie Tang. Cognitive Graph for Multi-Hop Reading Comprehension at Scale. ACL'19.
- Jie Zhang, Yuxiao Dong, Yan Wang, Jie Tang, and Ming Ding. ProNE: Fast and Scalable Network Representation Learning. IJCAI'19.
- Yukuo Cen, Xu Zou, Jianwei Zhang, Hongxia Yang, Jingren Zhou and Jie Tang. Representation Learning for Attributed Multiplex Heterogeneous Network. KDD'19.
- Fanjin Zhang, Xiao Liu, Jie Tang, Yuxiao Dong, Peiran Yao, Jie Zhang, Xiaotao Gu, Yan Wang, Bin Shao, Rui Li, and Kuansan Wang. OAG: Toward Linking Large-scale Heterogeneous Entity Graphs. KDD'19.
- Qibin Chen, Junyang Lin, Yichang Zhang, Hongxia Yang, Jingren Zhou and Jie Tang. Towards Knowledge-Based Personalized Product Description Generation in E-commerce. KDD'19.
- Yifeng Zhao, Xiangwei Wang, Hongxia Yang, Le Song, and Jie Tang. Large Scale Evolving Graphs with Burst Detection. IJCAI'19.
- Yu Han, Jie Tang, and Qian Chen. Network Embedding under Partial Monitoring for Evolving Networks. IJCAI'19.
- Jiezhong Qiu, Yuxiao Dong, Hao Ma, Jian Li, Chi Wang, Kuansan Wang, and Jie Tang. NetSMF: Large-Scale Network Embedding as Sparse Matrix Factorization. WWW'19.
- Jiezhong Qiu, Jian Tang, Hao Ma, Yuxiao Dong, Kuansan Wang, and Jie Tang. DeepInf: Modeling Influence Locality in Large Social Networks. KDD'18.
- Jiezhong Qiu, Yuxiao Dong, Hao Ma, Jian Li, Kuansan Wang, and Jie Tang. Network Embedding as Matrix Factorization: Unifying DeepWalk, LINE, PTE, and node2vec. WSDM'18.
- Jie Tang, Jing Zhang, Limin Yao, Juanzi Li, Li Zhang, and Zhong Su. ArnetMiner: Extraction and Mining of Academic Social Networks. KDD'08.

Thank you !

Jie Tang, KEG, Tsinghua U

Download all data & Codes

http://keg.cs.tsinghua.edu.cn/jietang
https://keg.cs.tsinghua.edu.cn/cogdl/
https://github.com/THUDM

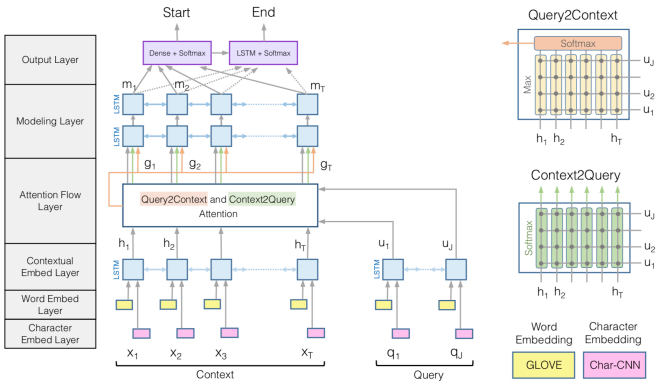
认知图谱：算法与认知的结合

Question: Who is the director of the 2003 film which has scenes in it filmed at the Quality Cafe in Los Angeles?

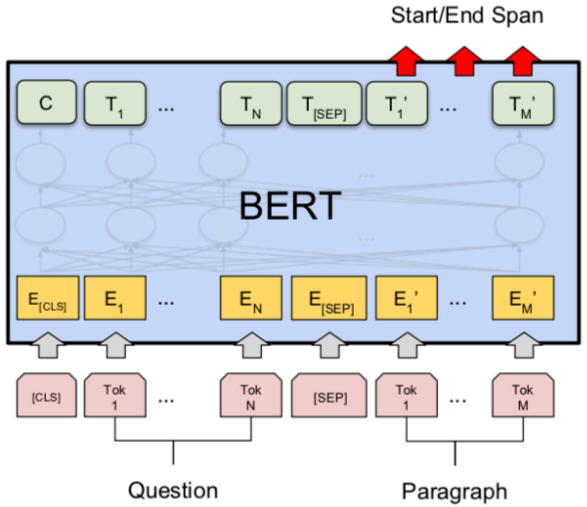
端到端模型的困境：BIDAF, BERT, XLNet

- 目标：理解整个文档，而不仅仅是局部片段
- 但仍然缺乏在知识层面上的推理能力

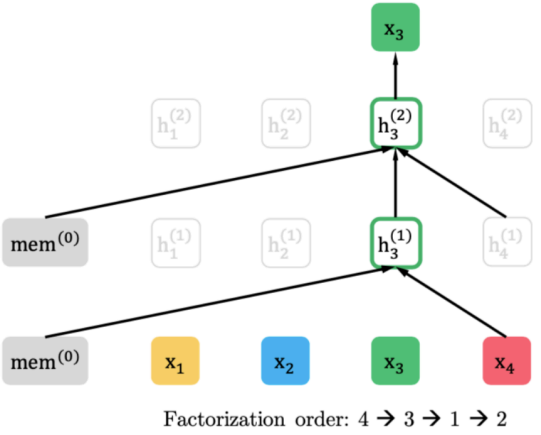
BiDAF



BERT



XLNet



挑战：可解释性

- 大部分阅读理解方法都只能看做**黑盒**：
  - 输入：问题和文档
  - 输出：答案文本块（在文档中的起止位置）

- 如何让用户可以验证答案的对错：
  - 推理路径或者子图
  - 每个推理节点上的支撑事实
  - 用于对比的其他可能答案和推理路径