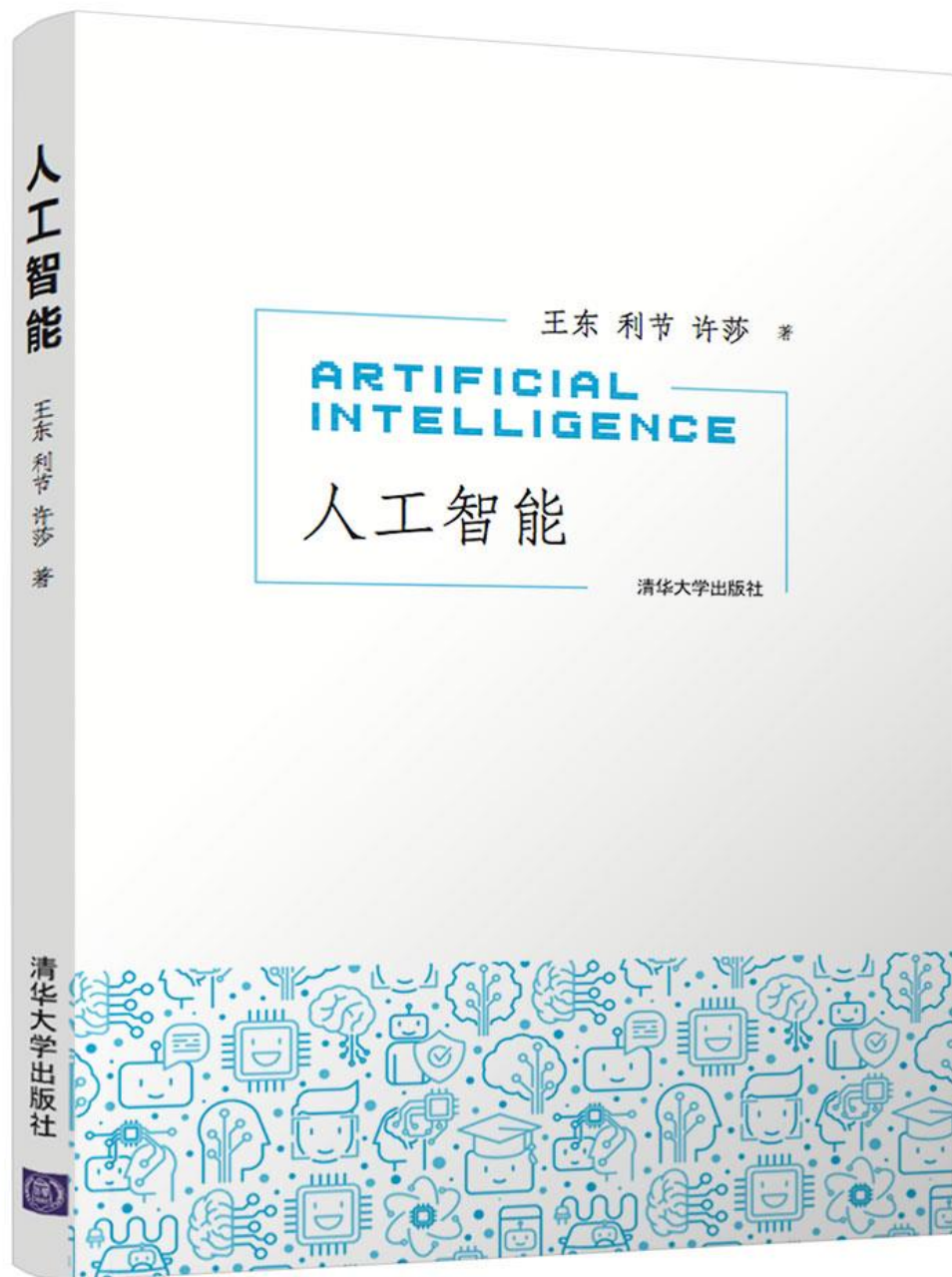


人工智能

王东 & 利节& 许莎



- 教材名称：人工智能
- 出版社：清华大学出版社
- 作者：王东 利节 许莎

本课程应具备的能力：

- Linux下的基本操作
- python语言基础知识

资源下载：

<http://aibook.csit.org/>



学习你的思维

蔡云麒

授课提示

1. 本课件50页，建议授课时间135分钟，每页平均约2.5分钟。
2. 若授课时间为180分钟，可对机器作诗、阅读理解两种任务的具体算法进行扩展性解释。
3. 课件内的图片、视频等素材具有较好的展示效果，建议保留。
4. 建议授课过程中，对“机器是否可实现真正的逻辑思维”进行讨论，激发学生思考。



目 录

- 1、什么是思维
- 2、形象思维
- 3、逻辑思维



目 录

- 1、什么是思维
 - 1.1、思维的特点
 - 1.2、两类思维
- 2、形象思维
- 3、逻辑思维

什么是思维？



思维是理性认识的过程，是人脑对客观事物间接和概括的反映——新华字典

思维的特点

➤ 抽象性

- 从一个太阳、两个苹果、三张椅子等抽象出“数”的概念
- 从太阳、苹果、椅子等抽象出“颜色”的概念



数量：1
颜色：橘黄



数量：2
颜色：红色



数量：3
颜色：蓝色

抽象出的特征

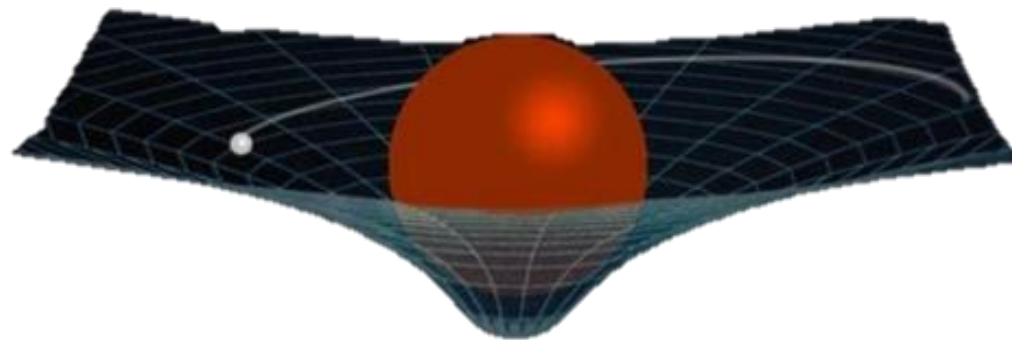
思维的特点

➤ 间接性

- 早晨看见地面潮湿，推知夜里下过雨
- 广义相对论中的时空弯曲



雨后清晨



爱因斯坦相对论



目 录

➤1、什么是思维

1.1、思维的特点

1.2、两类思维

➤2、形象思维

➤3、逻辑思维

两类思维



➤1、形象思维

- 基于直观印象处理问题的思维
- 抽象、想象和联想

➤2、逻辑思维

- 更抽象的思维，需要对事物及其规律做出高度概括
- 概念、判断、推理



目 录

- 1、什么是思维
 - 1.1、思维的特点
 - 1.2、两类思维
- 2、形象思维
- 3、逻辑思维

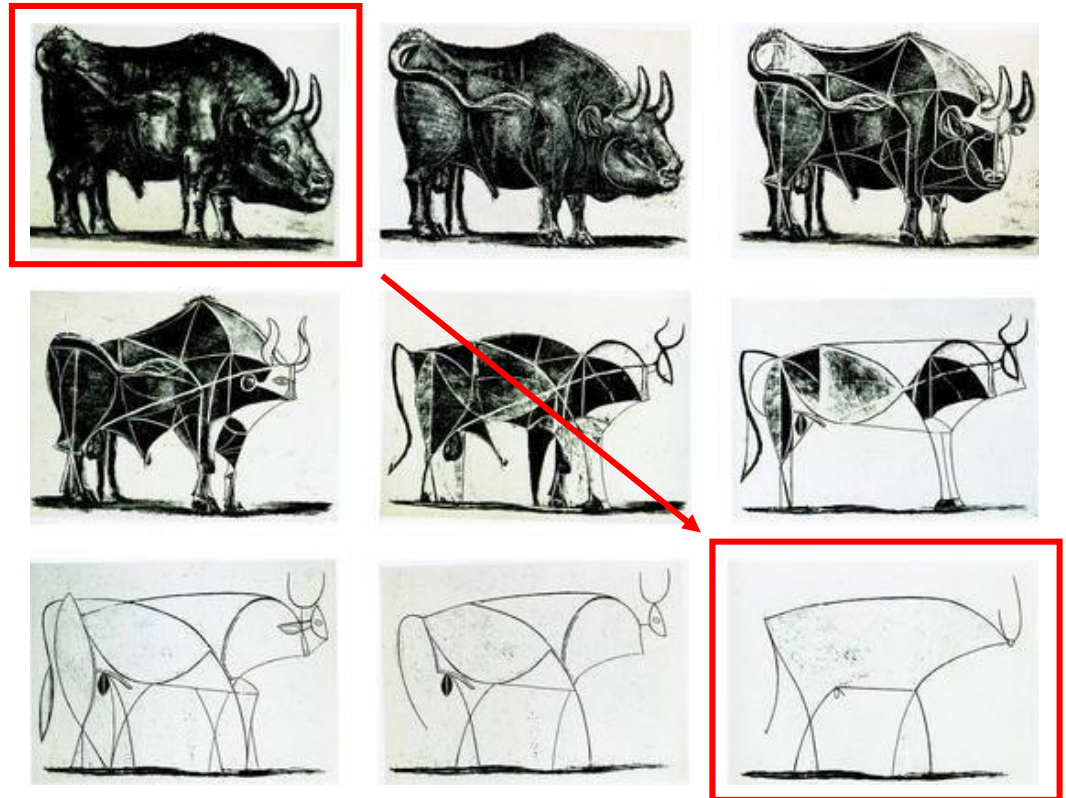


古今中外大师的观点

《书鄢陵王主簿所画折枝二首》

苏轼

论画以形似，见与儿童邻。
赋诗必此诗，定非知诗人。
诗画本一律，天工与清新。
边鸾雀写生，赵昌花传神。
何如此两幅，疏澹含精匀。
谁占一点红，解寄无边春。



特征提取



- **形象思维是指：**人们在认识世界的过程中，依靠**直观印象**解决问题的思维方法。形象思维是在对客观形象感受、体验的基础上，结合主观认识和自身情感**进行认知**，通过一定的形式、手段和工具（如文字语言、绘画色彩、音乐旋律等）对形象进行**再创造**（包括艺术形象和科学形象）的一种思维形式。

抽象 → 再创造

判断标准✓

机器是否能够达到这个标准





目 录

➤1、什么是思维

1.1、思维的特点

1.2、两类思维

➤2、形象思维

1.1、机器作诗

1.2、机器“盗梦”

➤3、逻辑思维

机器“作诗”

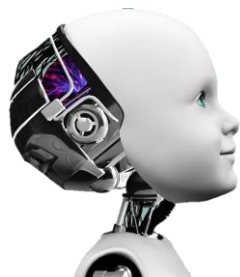
简介



➤ 机器“作诗”：机器是否能够像人类一样创作诗歌？

主题词

输入



输出

七言绝句
五言绝句
宋词
.....

例：“春”

输入



输出

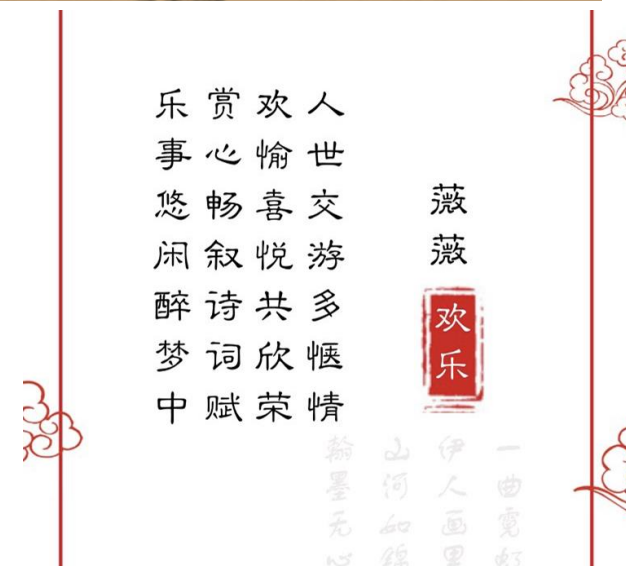
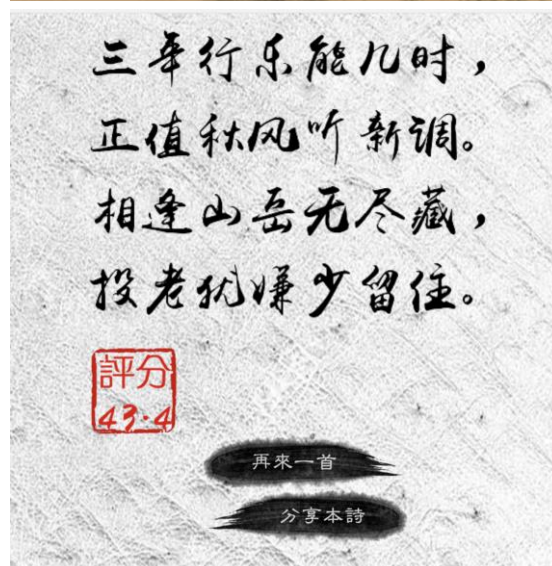
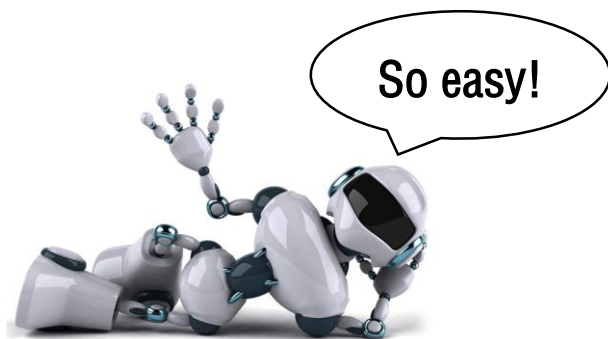
春眠不觉晓
处处闻啼鸟
夜来风雨声
花落知多少

机器“作诗”

简介

➤ 现有的机器“作诗”系统：

- “九哥”作诗机
- 微软律诗绝句
- “编诗姬”
- “薇薇”作诗机
-



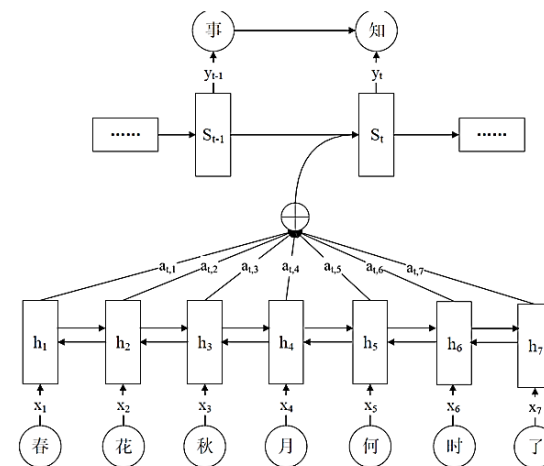
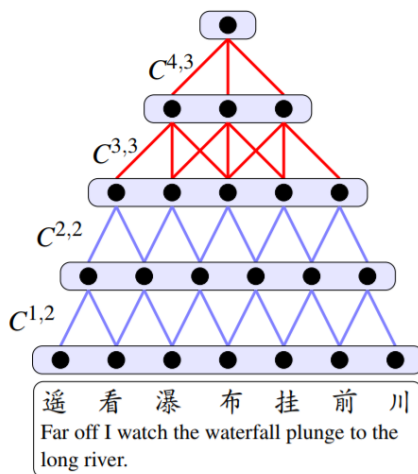
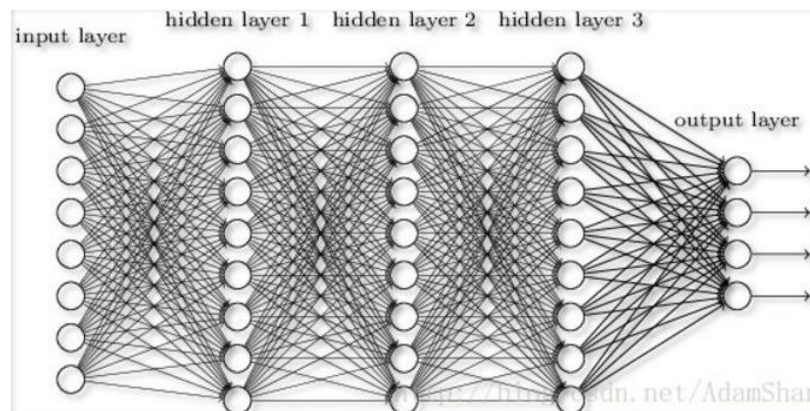
机器“作诗”

简介



➤ 机器“作诗”的方法：

- 基于规则和模板
- 基于遗传算法
- 基于统计机器翻译
- 基于深度学习



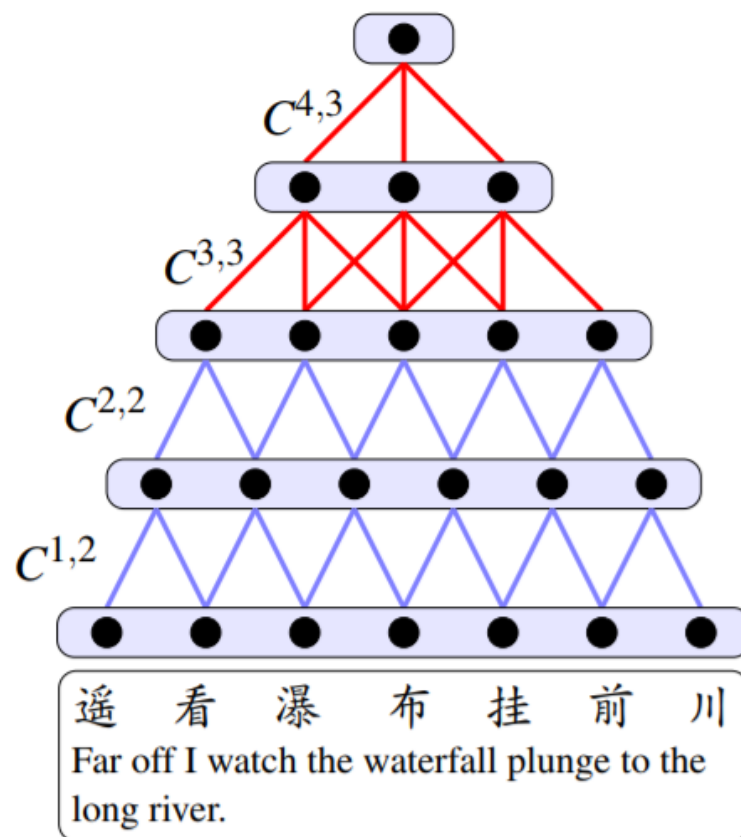
机器“作诗”

■ 基于深度学习的诗词生成



➤ 利用神经网络生成绝句诗：

- 首先用《诗学撷英》对用户输入的关键词进行扩展。
- 通过这一扩展，可以得到一个和用户意图相关的目标词集合。
- 基于该集合，利用拼凑法生成第一句诗。利用卷积神经网络（CNN）将第一句诗表示成一个句子向量



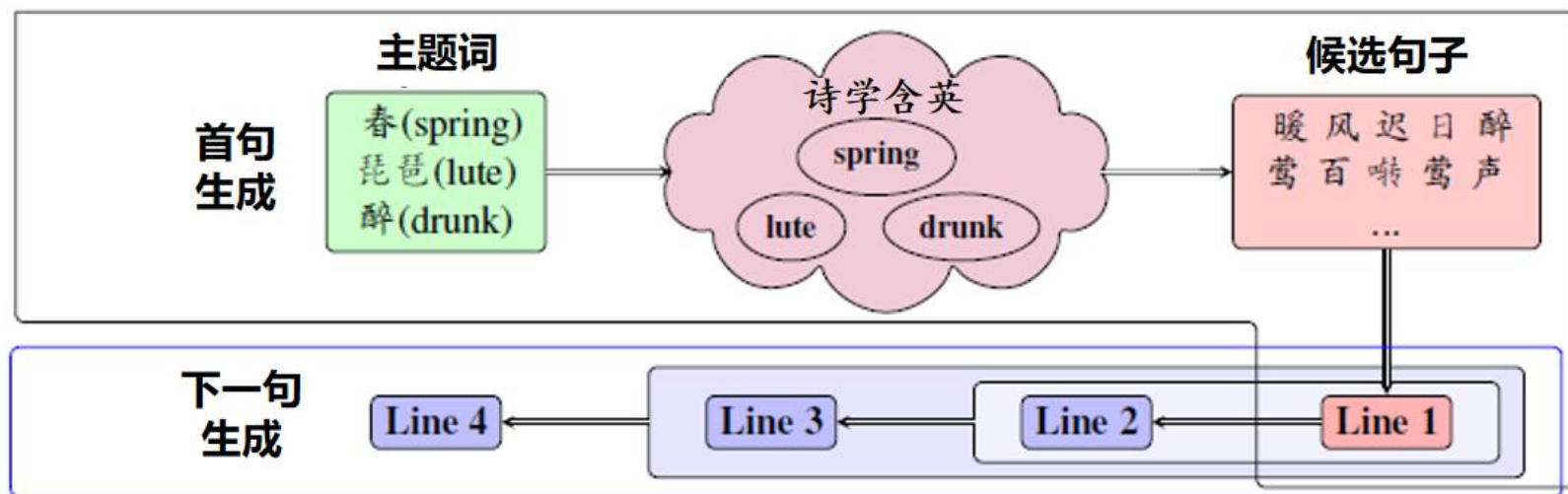
机器“作诗”

■ 基于深度学习的诗词生成



➤ 利用循环神经网络生成绝句诗：

- 循环生成诗句：利用递归神经网络(RNN)将第一句诗的句子向量作为输入得到第二句诗，然后将第一句和第二句诗的句子向量作为输入得到第三句诗，生成第四句时用前三句的句子向量作为条件输入。以此类推



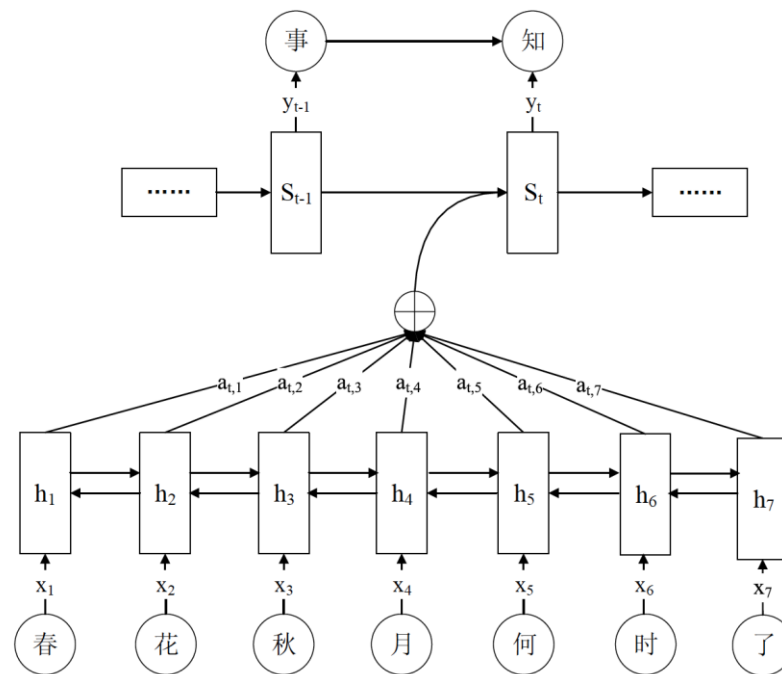
机器“作诗”

基于深度学习的诗词生成



➤ 利用基于注意力机制的机器翻译模型生成宋词：

- 例如：用户输入关键词“春花秋月何时了”。
- 一个双向 RNN 的编码器将关键词编码成一个**隐藏向量序列**（图下方的矩形序列），该向量序列包含了用户的生成意图。
- 在生成过程中，一个单向 RNN 网络递归运行，逐字生成整首诗。在生成每一个字的时候，注意力机制对编码器给出的隐藏向量进行查看，找到与当前生成状态最相关的用户意图，利用这一信息指导下一个字的生成。
- 在生成过程中，需要强制加入断句、押韵、平仄等诗词规则。通过这一生成方式，可以保证生成的字串既能最大程度地符合诗词规则，又能使生成围绕用户的意图展开。同时，这一模型简单灵活，可以用来生成各种体例的诗或词。



机器“作诗”

■ 基于深度学习的诗词生成



➤ 基于图像的作诗方法：

- 不同于前面的给出关键词，现在我们让机器看一幅图，然后作出一首应景的诗。
- 基于这幅输入图像，深度神经网络可以得到“视觉”和“关键词”两种信息
- 利用同样的基于注意力机制的序列对序列网络



扁舟一曲水平堤，
I sing a fishing song on a boat in the lake overflowing its bank,
一棹渔舟日向西。
rowing oars with the sun setting in the west.
长忆西湖水中月，
I often miss the moon reflected in the West Lake,
东风吹过武陵溪。
and the east breeze blowing across the WuLing River.



春风庭院养花姿，
Breeze blows beautiful flowers in the courtyard,
春入帘栊叶满枝。
Spring comes into my window, with leaves covering the branches.
堪笑门前青草树，
Glad to see green grass and trees in front of my door,
谁家芳节几多时。
However spring will not last very long.

机器“作诗”

基于深度学习的诗词生成

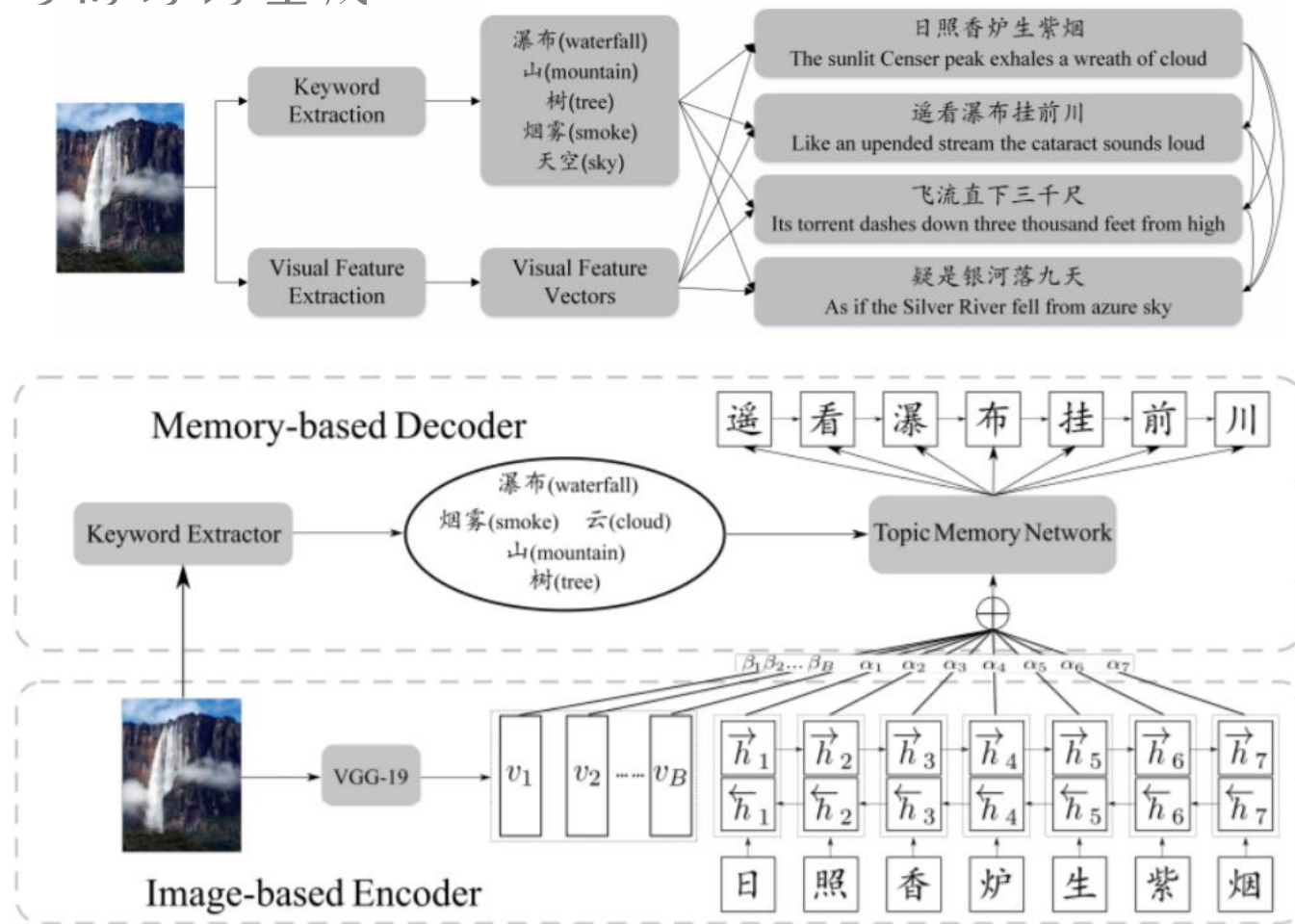


Fig. 6.6 基于图像作诗的神经网络模型 [164]。上图是系统框架。首先由图像得到关键词信息和视觉信息，以这两种信息作为输入，逐句生成整首诗。下图是系统所用的神经网络模型。生成每句诗时，将已经生成的所有句子和视觉信息作为输入，同时参考图像关键词信息。

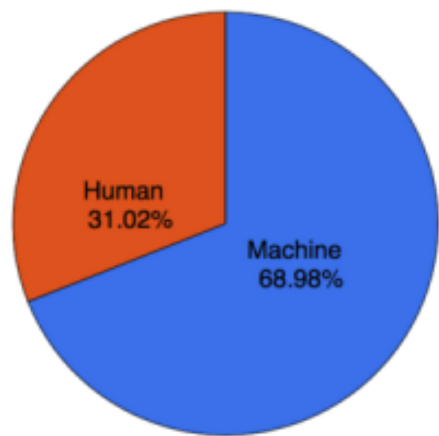
机器“作诗”

■ 薇薇通过图灵测试



➤ 图灵测试：

- 2016 年 3 月，清华大学语音语言技术中心做了一组实验，实验中该中心研发的作诗机器人“薇薇”和一些网络诗人就同一主题进行创作，并邀请北大、社科院等单位的诗词专家进行品评。



(a) 图灵测试中薇薇的作品中有 31% 被专家认为是人

海棠花

红霞淡艳媚妆水，
万朵千峰映碧垂。
一夜东风吹雨过，
满城春色在天辉。

Fig. 6.5 作诗机器人“薇薇”的一首作品 [153]。这首作品是 2016 年 3 月的图灵测试中，在所有人、机作品中得分最高的一首。

机器“作诗”

总结

➤ 机器“作诗”：

- 基于深度学习的模型结构已经能达到不错的效果；
- 仍存在思想空洞等问题；
- 如何让机器像人一样“**寓情于诗**”，写出有意境且有思想的诗词，还需要自然语言处理的学者们不断地尝试和努力。



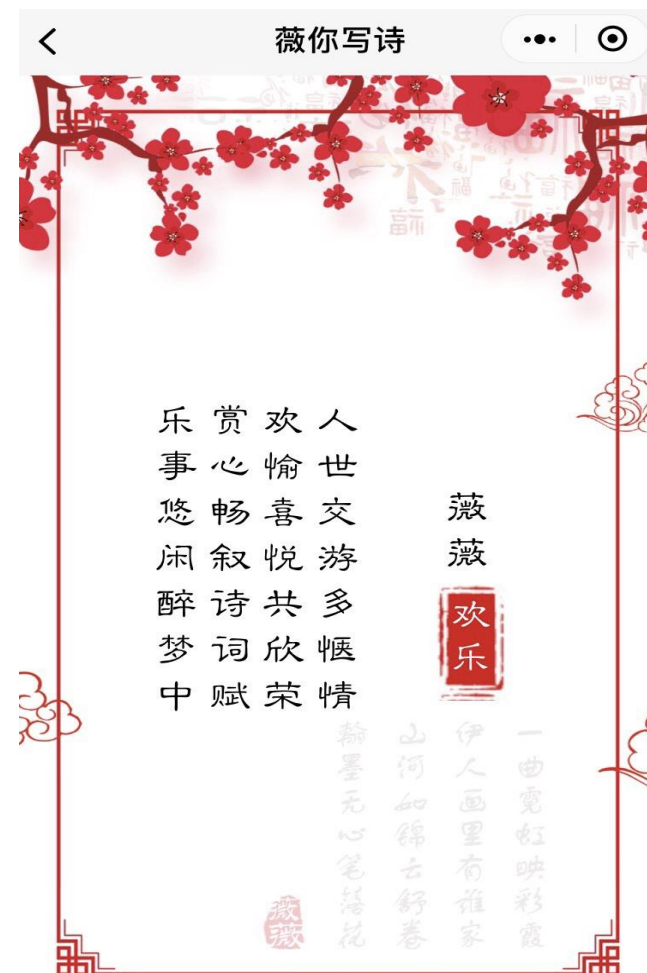
机器“作诗”

体验“薇你写诗”



微信公众号：**清语赋**

微信小程序：**清语赋**





目 录

➤1、形象思维

1.1、机器作诗

1.1.1、作诗系统介绍

1.1.2、机器作诗的方法介绍

1.2、机器“盗梦”

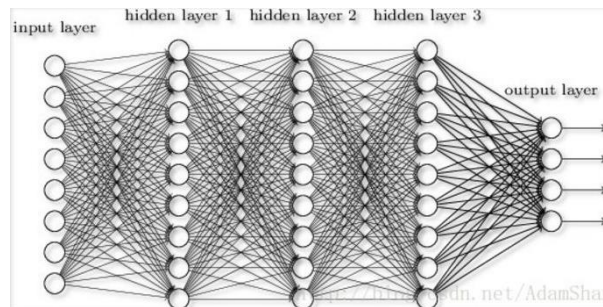
➤2、逻辑思维

2.1、精确逻辑推理

2.2、非精确逻辑推荐

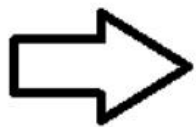
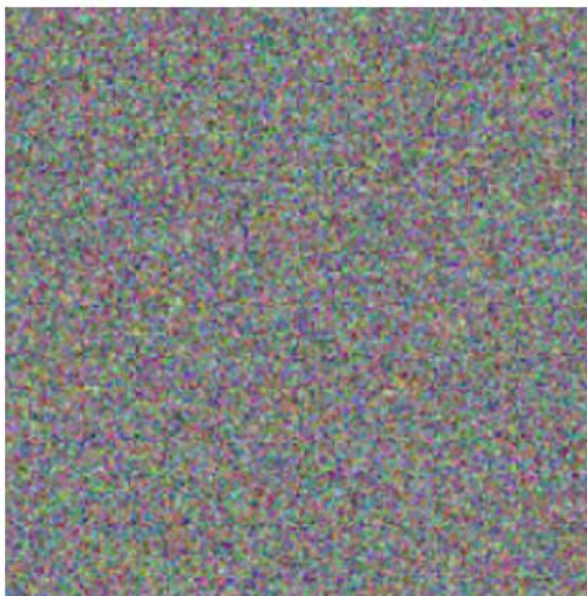
机器“盗梦”

■ 基于深度学习的“Deep Dream”



➤ 神经网络可视化：

- 谷歌的研究员 Alexander Mordvintsev 及其两位同事发现，对于一个用于图片分类的 DNN，可以通过一种称为“Creation by Refinement”的方法来观察该网络对某一概念的学习情况。



我们设想，有一个人工智能机器人通过学习认识了很多的动物，其中就包括海星。这个时候我们给他看一张如图左边所示的噪声图像，他会说：“这不是海星”然后我们不断的调整噪声图像，直到他说“这是个海星！”。于是当前的图像即为这个机器人脑海里海星的模样。

Fig. 6.8 对一张噪声图片进行调整，使得以此图片作为输入，神经网络在代表“海星”的输出节点上激发值最大，即可得到一张海星图片。来自 Google AI 的 Blog³。

机器“盗梦”

■ 基于深度学习的“Deep Dream”



➤ 底层节点强化：

- 如图所示是对一幅山羊图片进行底层节点强化后的结果，即对 DNN 的前几层节点做强化。由于这几层主要代表一些局部特征（线条、色彩等），强化这些层的节点会在图片上加入一些彩色条纹。

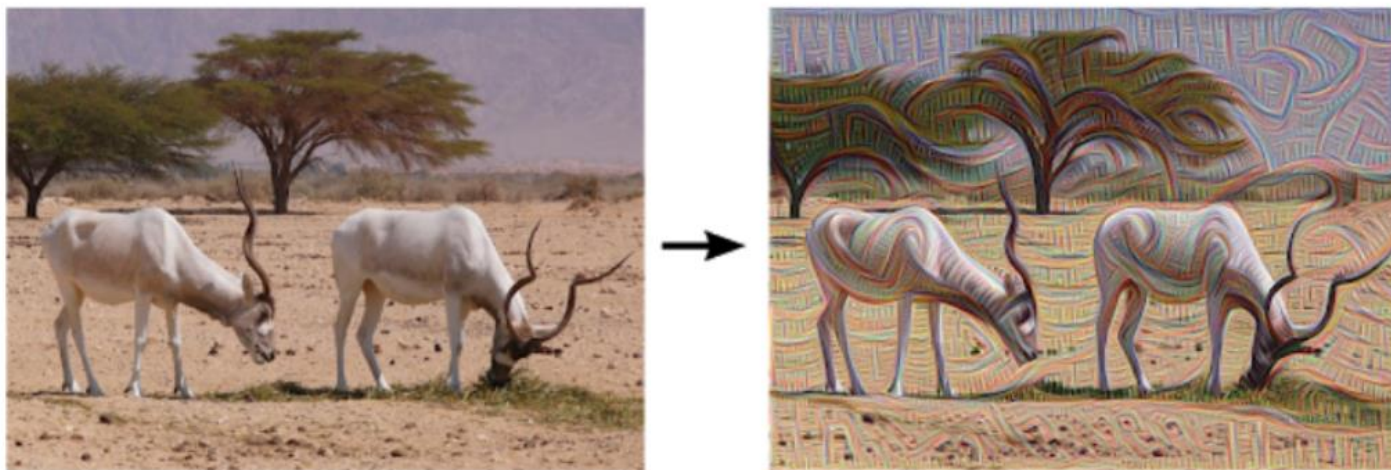


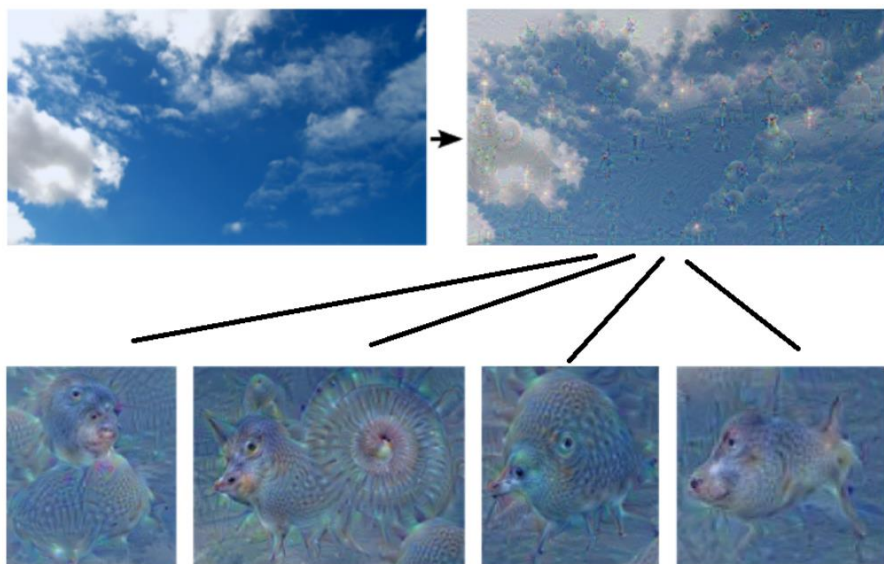
Fig. 6.9 对一幅图片强化底层节点时，会在图片中加入一些类似斑马纹的彩色条纹。来自 Google AI 的 blog。

机器“盗梦”

■ 基于深度学习的“Deep Dream”

➤ 高层节点强化：

- 高层节点代表了抽象的、复杂的特征，如某种动物的轮廓、形态等，对这些特征进行强化将产生梦幻般的效果。
- 输入一副蓝天图片，机器会在某些节点上产生较大的激发。因为该神经网络是用来区分动物的，这些节点事实上对应着某些动物形象。当我们对这些特征进行强化时，就会在图片中显现出这些动物的轮廓来。



这就类似于我们小时候盯着云朵看，偶尔会发现一些类似动物的图案，然后越看越像，越看越迷乱，即产生了自我暗示。Google 将这一强化系统称为 Deep Dream。



Fig. 6.10 对一幅图片强化高层节点时，会在图片中加入一些复杂的形象。由于该神经网络是识别动物的，因此强化出来的形象都是动物形象。来自 Google AI 的 blog。

机器“盗梦”

基于深度学习的“Deep Dream”



Fig. 6.11 Deep dream 生成的几幅图像。上面一行是原图，下面一行是由 Deep Dream 生成的图像。来自 Google AI 的 blog。



Fig. 6.12 Deep dream 合成的复杂图像，其中局部特征包含一定真实性，整体照片则是合成出来的非现实场景。图片来源于 WIRED。⁵

机器“盗梦”

■ 基于深度学习的“Deep Dream”

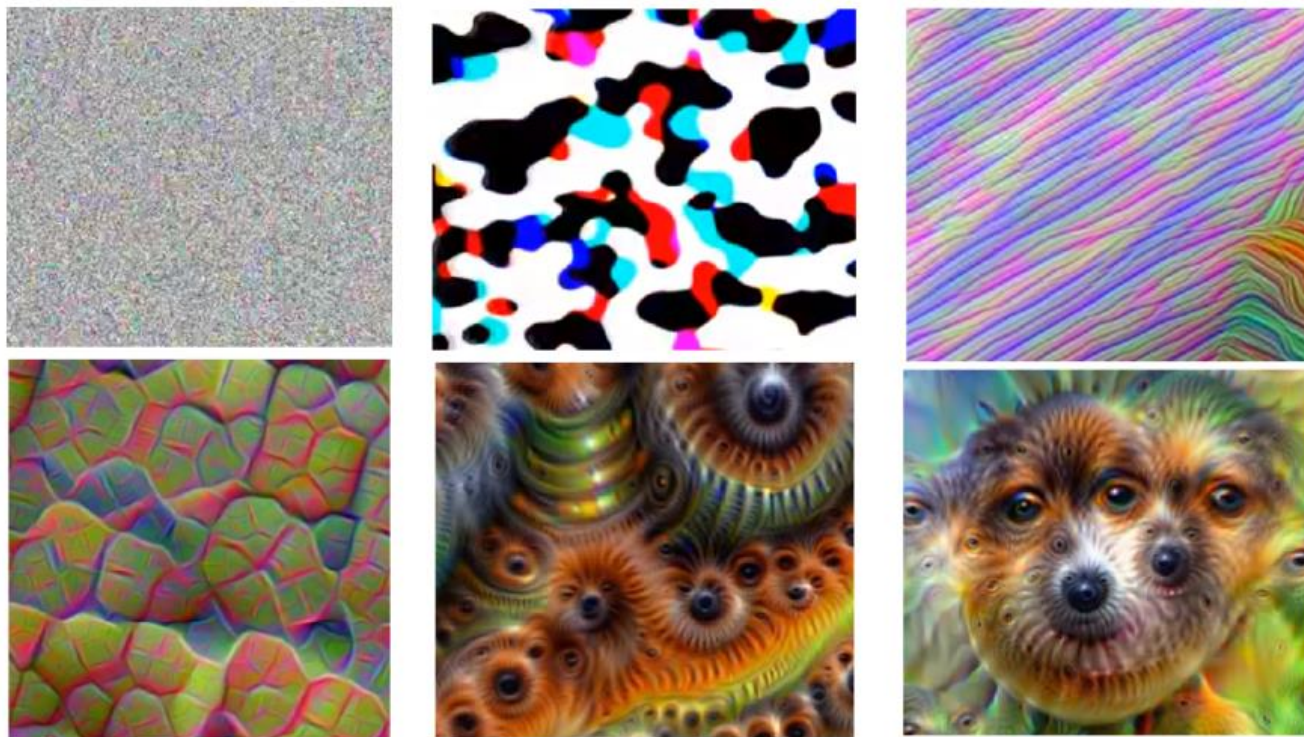
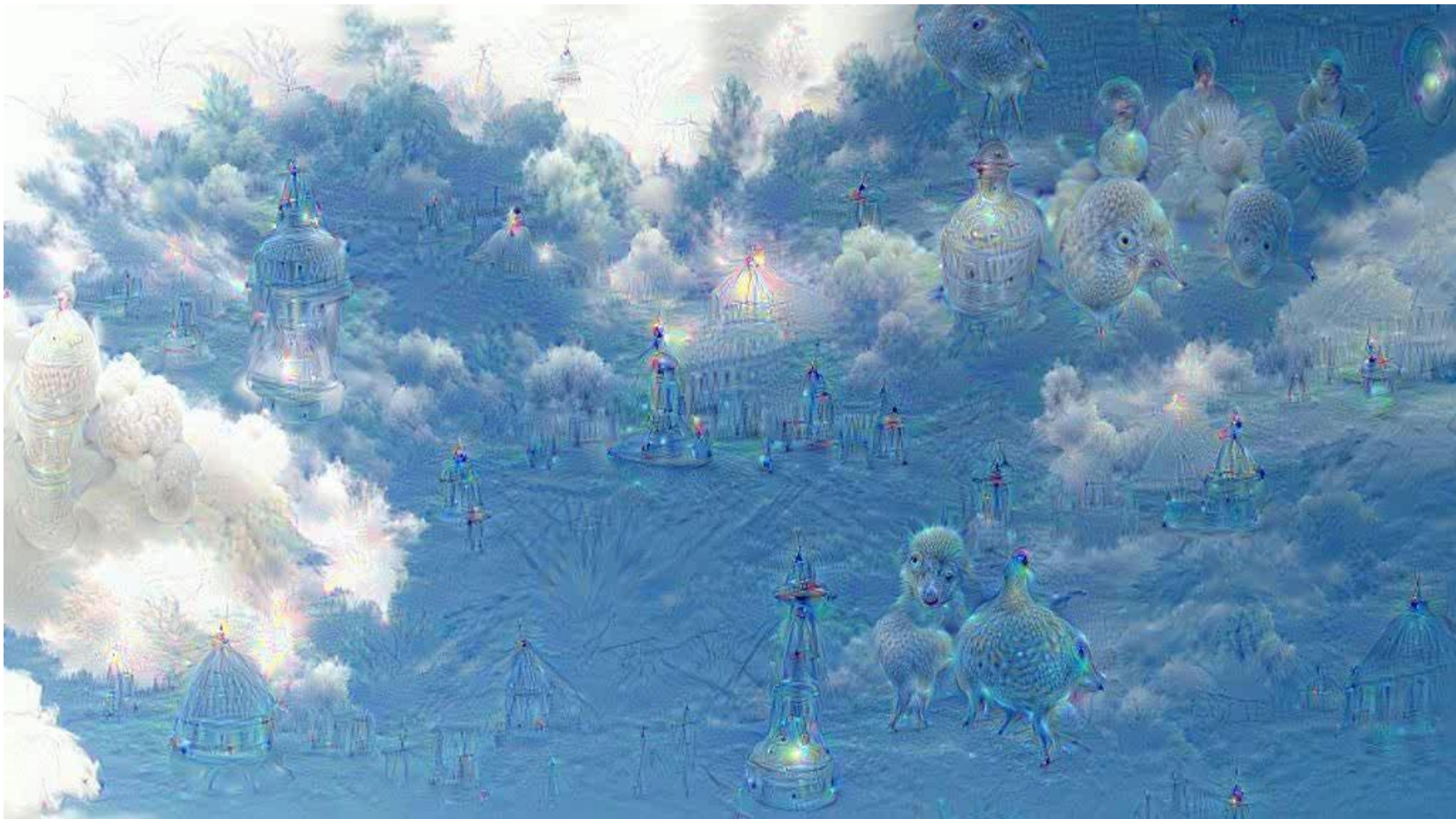


Fig. 6.13 由随机图片开始，迭代生成梦境的过程。每次迭代时，首先对当前图片进行强化，然后对强化后的图片做局部放大，最后将放大后的图片送入下一轮迭代，准备再次强化。⁷

晚上十点钟，小明准时的上床睡觉了。当他闭上眼睛后，脑海里出现了像电视雪花屏一样的白噪画面，还没来得及过多的观察，画面突然变得五颜六色起来，小明试着观察其中的某处，突然他看到色块里面还有着很多细致的纹理。纹理逐渐的在小明的脑海中放大，直到出现了很多类似于鳞片一样的多边形。最后画面变得越来越具象，变的越来越像以前小明家养的一只棕色小狗的脸。

机器“盗梦”

■ 基于深度学习的“Deep Dream”





目 录

➤1、形 象 思 维

1.1、机 器 作 诗

1.2、机 器 “盗 梦”

➤2、逻 辑 思 维

2.1、精 确 逻 辑 推 理

2.2.1、定 理 证 明

2.2、非 精 确 逻 辑 推 理

精确逻辑推理

推理+启发搜索



➤1、推理+启发搜索

- 1954, Martin Davis, JOHNNIAC, 证明两个偶数相加还是偶数
- 1956, Newell, Shaw 和 Simon, “逻辑机器” (Logic Theory Machine, LTM)
- 1959, 王浩, 9分钟证明《数学原理》中所有定理

➤2、机器几何定理证明

- 1978年, 吴教授提出几何定理证明新方法, 该方法将几何定理证明问题转化为方程式计算问题, 从而极大提高了证明效率。吴教授和他的学生们把该方法扩展到扩展到微分几何, 智能 CAD、机器人、计算机视觉等各个方面, 做出了重要贡献。

精确逻辑推理

几何定理证明

机器几何定理证明

基于吴方法的自动几何定理证明大致步骤如下：

1. 将几何问题代数化，将给定的几何条件翻译成多项式方程，（hypothesis polynomials） $f_1, f_2, \dots, f_r$ ，将几何结论也翻译成多项式 $g_1, g_2, \dots, g_s$ （conclusion polynomials）
2. 用伪除方法（pseudodivision）将条件多项式变换成三角列形式，

$$\begin{cases} f_1 = f_1(u_1, \dots, u_d, x_1) \\ f_2 = f_2(u_1, \dots, u_d, x_1, x_2) \\ \vdots \\ f_r = f_r(u_1, \dots, u_d, x_1, \dots, x_r) \end{cases}$$

代数簇 $V(f_1, \dots, f_r)$ 包含条件多项式生成的代数簇的不可约分支（irreducible components）

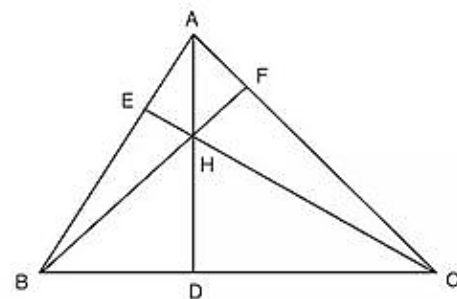


图2. 三角形垂心定理。

例如我们来证明三角形垂心定理：如图2所示，三角形的顶点坐标是

$$A = (u_2, u_3), B = (0, 0), C = (u_1, 0),$$

三个垂足为

$$D = (u_2, 0), E = (x_1, x_2), F = (x_3, x_4),$$

由垂直条件得到多项式方程

$$\begin{aligned} x_2 u_2 - x_1 u_3 &= 0 \\ x_4(u_2 - u_1) - u_3(x_3 - u_1) &= 0 \\ x_2 u_3 + u_2(x_1 - u_1) &= 0 \\ x_4 u_3 + x_3(u_2 - u_1) &= 0, \end{aligned}$$

我们假设AD和CE交于G点，AD和BF交于H点， $G = (u_2, x_5)$ ， $H = (u_2, x_6)$ 。由G、E、C三点共线，H、B、F三点共线，增加方程

$$\begin{aligned} (x_2 - x_5)(x_1 - u_1) - x_2(x_1 - u_2) &= 0 \\ x_6 x_3 - x_4 u_2 &= 0 \end{aligned}$$

我们需要证明G和H重合，即 $x_5 - x_6 = 0$ 。



精确逻辑推理

几何定理证明

机器人路径规划

吴方法可以用来求解多项式方程组。将一般的多项式方程组化解为三角列形式，非常类似于线性方程组的高斯消元法

(Gauss elimination)。我们通过数值方法求解单变元多项式 $f_1(u_1, \dots, u_d, x_1) = 0$ ，求得 x_1 ；然后将 x_1 代入第二个方程 $f_2(u_1, \dots, u_d, x_1, x_2) = 0$ ，求得 x_2 ；再将 x_1, x_2 代入第三个方程

$f_3(u_1, \dots, u_d, x_1, x_2, x_3) = 0$ ，求得 x_3 ；以此类推，逐步求得所有的未知变量 $(x_1, x_2, \dots, x_r)$ 。



图3. 机械臂逆向运动学。

在机器人 (robotics) 领域，机械臂路径规划是一个经典问题。一条机械臂有多个关节，每个关节有旋转自由度或者伸缩自由度，我们将这些自由度由变元 $(\theta_1, \theta_2, \dots, \theta_r, d_1, d_2, \dots, d_s)$ 表示。机械手的位置和朝向由这些变元控制，同时它们可以表示成多项式函数：

$$f(\cos \theta_1, \sin \theta_1, \cos \theta_2, \sin \theta_2, \dots, \cos \theta_r, \sin \theta_r, d_1, d_2, \dots, d_s)$$

同时我们有限制 $\cos^2 \theta_i + \sin^2 \theta_i = 1$ 。在机器人应用中，机器人通过三维扫描获得物体的三维几何位置信息，从而得到最终机械手的位置和朝向，通过反解各个关节的旋转角度，和机械臂的伸缩，使得机械手达到目标位置，从而可以实现抓取。这被称为是逆向运动学问题 (inverse kinematics)，需要求解多项式方程组，而吴方法正是解多项式方程组的有力武器。



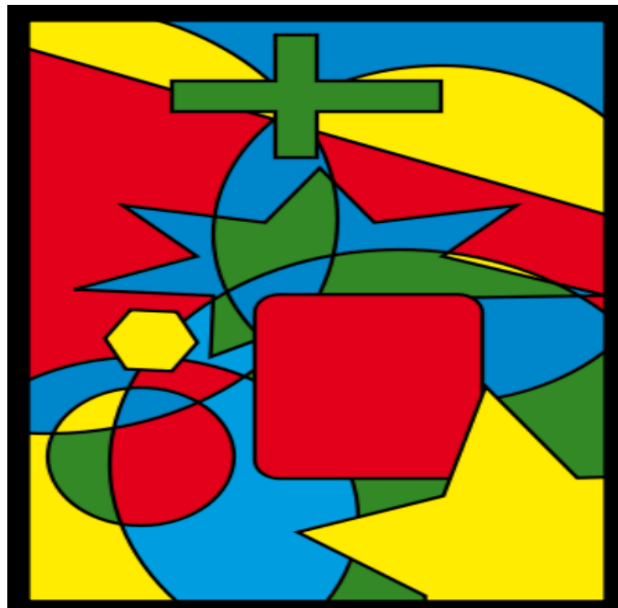
精确逻辑推理

几何定理证明



➤ **四色定理：** 任意一幅地图，不管多么复杂，都可以用四种颜色完成染色，并保证相邻国家不会同色。

- 该问题由英国制图员Francis Guthrie在1852年提出，一直困扰了数学家100多年，直到1976年，数学家Kenneth Appel和Wolfgang Haken借助计算机首次得到一个完整的证明，四色问题也终于成为四色定理。一些科学家认为，与这一定理本身相比，“计算机证明”这一点对数学界的意义更加重大，它表明计算机有可能成为数学家不可或缺的工具，就象显微镜对生物学家一样。





➤ 精确推理总结：

- 值得说明的是，上述定理证明方法本质上是在某一**确定逻辑系统**所定义的命题空间中的搜索过程。在这一过程中，所有公理和推理原则都是既定的、假设正确的，这将保证所发现的推理过程是无可质疑的，因此是一种精确的逻辑推理。机器定理证明本质上是人类智能在机器上的复现和加速。一方面，它确实可以发现一些未知的结论，而且可能是很重要的结论，如四色定理；另一方面，这种方法所发现的所有结论都蕴含在前提假设中，始终只能在一个**封闭域**中推导知识。



目 录

➤1、形 象 思 维

1.1、机 器 作 诗

1.2、机 器 “盗 梦”

➤2、逻 辑 思 维

2.1、精 确 逻 辑 推 理

2.2、非 精 确 逻 辑 推 荐：

2.2.1、机 器 问 答

2.2.3、阅 读 理 解

非精确推理

机器问答

ELIZA

Person: Men are all alike.

ELIZA: In what way?

Person: They're always bugging us about something or other.

ELIZA: Can you think of a specific example?

Person: Well, my boyfriend made me come here.

ELIZA: Your boyfriend made you come here.

Person: He says I'm depressed much of the time.

judge: What is the meaning of life?

ALICE: The meaning of life is part of God's mysterious plans.

judge: Make me a sandwich.

ALICE: What are the ingredients?

judge: Turkey and mustard.

ALICE: I don't have any turkey and mustard?

judge: What do you have?



非精确推理

问答机器人介绍

- ELIZA(1966),Start(1993),Alice(1995)
- 微软小冰
- 图灵机器人
- 爱客服



www.yqdown.com [友情下载] 提供



非精确推理

基于模板的问答系统

➤ 模板替换

ELIZA

Person: Men are all alike.

ELIZA: In what way?

Person: They're always bugging us about something or other.

ELIZA: Can you think of a specific example?

Person: Well, my boyfriend made me come here.

ELIZA: Your boyfriend made you come here.

Person: He says I'm depressed much of the time.

ELIZA: I am sorry to hear you are depressed.

Person: It's true. I am unhappy.

ELIZA: Do you think coming here will help you not to be unhappy?

Person: I need some help, that much seems certain.

ELIZA: What would it mean to you if you got some help?

Person: Perhaps I could learn to get along with my mother.

ELIZA: Tell me more about your family.

➤ 模板通式样例

(X me Y)	→ (X you Y)
(I remember Y)	→ (Why do remember X just now?)
(My {family-member} is Y)	→ (Who else in your family is Y?)
(X {family-member} Y)	→ (Tell me more about your family)



非精确推理

■ 基于语义分析的机器问答

- 自然语言提问，进行语义映射，基于知识库给出问题答案。



Google

姚明老婆的国籍

All Images News Videos Maps More Settings Tools

About 126,000 results (0.54 seconds)

Ye Li / Nationality

Chinese

People also search for

Yao Ming Chinese

Sui Feifei Chinese

Sun Mingming Chinese

Feedback

非精确推理

基于深度学习的机器问答

➤ 利用维基百科回答开放式问题：



Q: How many of Warsaw's inhabitants spoke Polish in 1933?

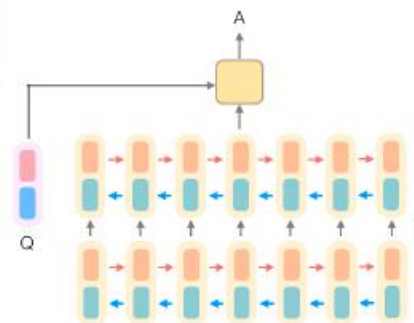


Document
Retriever



Document
Reader

833,500



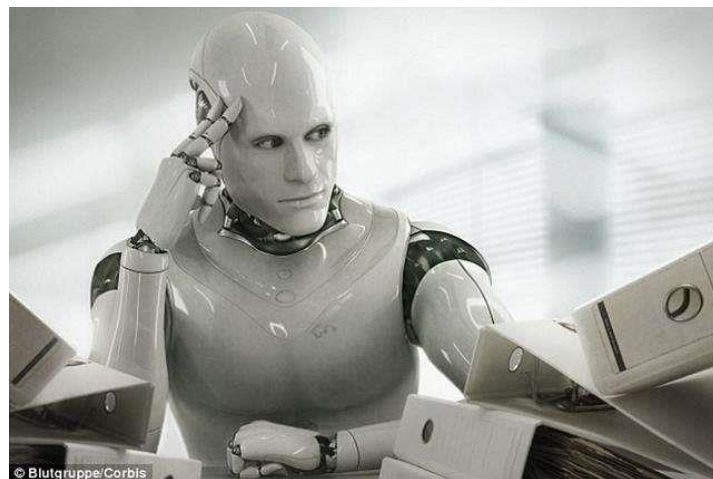
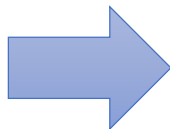
Chen D, Fisch A, Weston J, et al. Reading Wikipedia to Answer Open-Domain Questions[C]// Meeting of the Association for Computational Linguistics. 2017:1870-1879.

非精确推理

机器问答总结

➤ 机器问答总结：

- 基于语义分析的问答比较精确，适用于特别领域。基于深度学习的模型建模更加简洁，在一通用问题上效果更优。
- 基于结构化知识库的问答较基于非结构化的知识库更简单，但基于非结构化知识库的问答用途更广，更类似于人类的问答方式。
- 如何利用“终身学习（life-long learning）”训练机器，使其能够像人一样拥有较长的记忆和灵活的问答方式将是未来的努力方向。





目 录

➤1、形 象 思 维

1.1、机 器 作 诗

1.2、机 器 “盗 梦”

➤2、逻 辑 思 维

2.1、精 确 逻 辑 推 理

2.2、非 精 确 逻 辑 推 荐：

2.2.1、机 器 问 答

2.2.3、阅 读 理 解

非精确推理

机器阅读理解简介

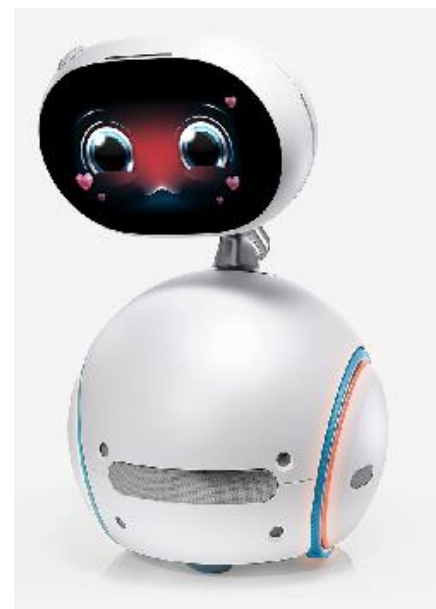


- **机器阅读理解：**是一个典型的非精确推理任务，要解答这类题型，需要对叙事和问题的语义有充分理解，再通过一系列推理过程发现答案。

故事：有一天白雪公主的外婆来了。外婆说：你看，外面的树长的多高啊，你照顾的真好。白雪公主说：外婆，是侍女小明照顾的好。

问题：谁在照顾树？

答案：小明



非精确推理

机器阅读理解



- SQuAD数据集演变为阅读理解一个比赛：
<https://rajpurkar.github.io/SQuAD-explorer/>

Rank	Model	EM	F1
	Human Performance Stanford University (Rajpurkar et al. '16)	82.304	91.221
1 Jan 22, 2018	Hybrid AoA Reader (ensemble) Joint Laboratory of HIT and iFLYTEK Research	82.482	89.281
1 Mar 06, 2018	QANet (ensemble) Google Brain & CMU	82.744	89.045
1 Feb 19, 2018	Reinforced Mnemonic Reader + A2D (ensemble model) Microsoft Research Asia & NUDT	82.849	88.764
2 Jan 05, 2018	SLQA+ (ensemble) Alibaba iDST NLP	82.440	88.607
3 Feb 02, 2018	Reinforced Mnemonic Reader (ensemble model) NUDT and Fudan University https://arxiv.org/abs/1705.02798	82.283	88.533
3 Feb 27, 2018	QANet (single model) Google Brain & CMU	82.209	88.608
3 Jan 03, 2018	r-net+ (ensemble) Microsoft Research Asia	82.650	88.493

SQuAD 阅读理解是由斯坦福大学收集的一批用于评测阅读理解任务的数据库。该数据库包括从 536 篇维基百科文章中提取出的 2.3 万个段落，将这些段落记为叙事 (C)；同时，又用人工方法生成了 10.8 万个问题 (Q)。问题类型包括事件、日期、人名、地点等。研究者首先做了一个实验，让人来回答这些问题，发现人类的准确率可达 82.3%。

(进入链接) 可以看到很多表现出色的深度学习系统都已经超过了号称有精确推理能力的人类所能达到的水平。这说明在很多任务中并不需要精确推理，而人也并不是每时每刻都在使用精确推理。非精确推理应该被认为是人工智能的重要组成部分，而不仅仅是某种权宜之计。

非精确推理

其他任务

- **寻找凸边界**: 发现一个由群中某点作为顶点的凸包, 可以包围群中所有点。
- **Delaunay 三角化**: 将群中所有的点联结成三角形, 并保证所有三角形的外接圆内部不包含群中其它点。
- **最短旅行**: 求解从一个点出发访问所有点并回到起始点的最短路径。

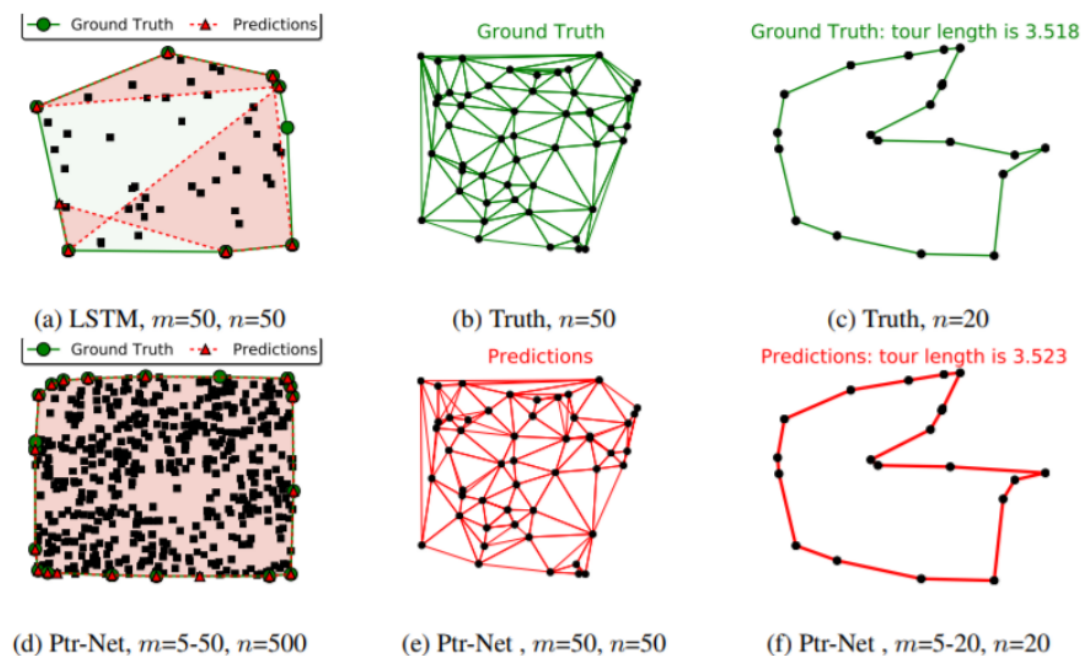


Fig. 6.19 基于指针网络的近似推理。共有三个任务: (左) 寻找凸边界; (中) 三角化; (右) 最短距离旅行。上边三幅图是最优解, 下边三幅图是基于指针网络得到的近似解。图片来自 [151]。

➤ 思考习题：

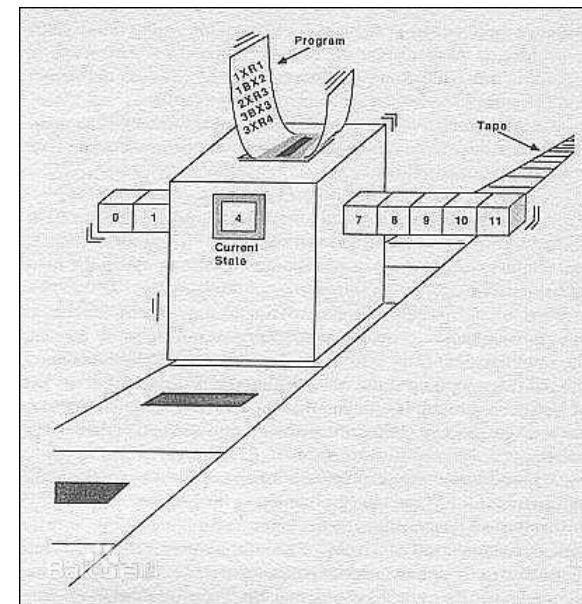
- 1. 形象思维与逻辑思维的区别是什么？Deep Dream 是如何学习形象思维的？
- 2. 精确推理与非精确推理的区别是什么？哪种推理更接近人类？
- 3. 机器智能有可能超过人类吗？

拓展阅读

图灵机

➤ 神经图灵机

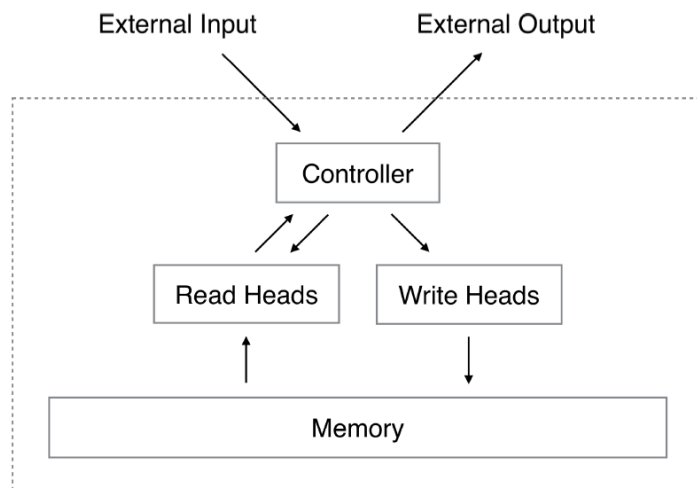
- **图灵机**：现代计算机的雏形，由四部分组成：
 - ✧ 无限长的纸带：纸带上有一个个小方格，每个方格可以存储一个符号，纸带可以向左或向右移动；
 - ✧ 读写头：可以读取和修改方格内的内容；
 - ✧ 状态寄存器：保存机器现在的状态；
 - ✧ 控制规则（程序）：根据机器现在的状态和方格内的内容作出操作，并更改状态寄存器的内容。
- **现代计算机**：计算模块 + 控制模块 + 外部记忆模块
- **认知神经学研究**表明：在人类的大脑中也存在一个类似于“working memory”的过程，能在大脑思考时，短时保存信息。
- 然而，现在的**机器学习模型**一般只有“计算模块”，而缺少了“控制模块”和“外部记忆模块”。



➤ 神经图灵机

- 神经图灵机包括两个基本模块：

- ✧ 基于神经网络的控制器（Controller）：负责接收输入，产生输出，同时读写记忆模块。控制器可以选择任何合适的神经网络模型。
- ✧ 记忆模块（Memory）：模型运算过程中的“working memory”。



➤ 神经图灵机

- 为了保证整个模型的可导性，不采用读取或写入单个记忆单元的方式，而提出了“**模糊读写**”的方式：

- 读操作： $w_t(i)$ 是对第 i 个记忆单元的读取权重，类似于“注意力”；

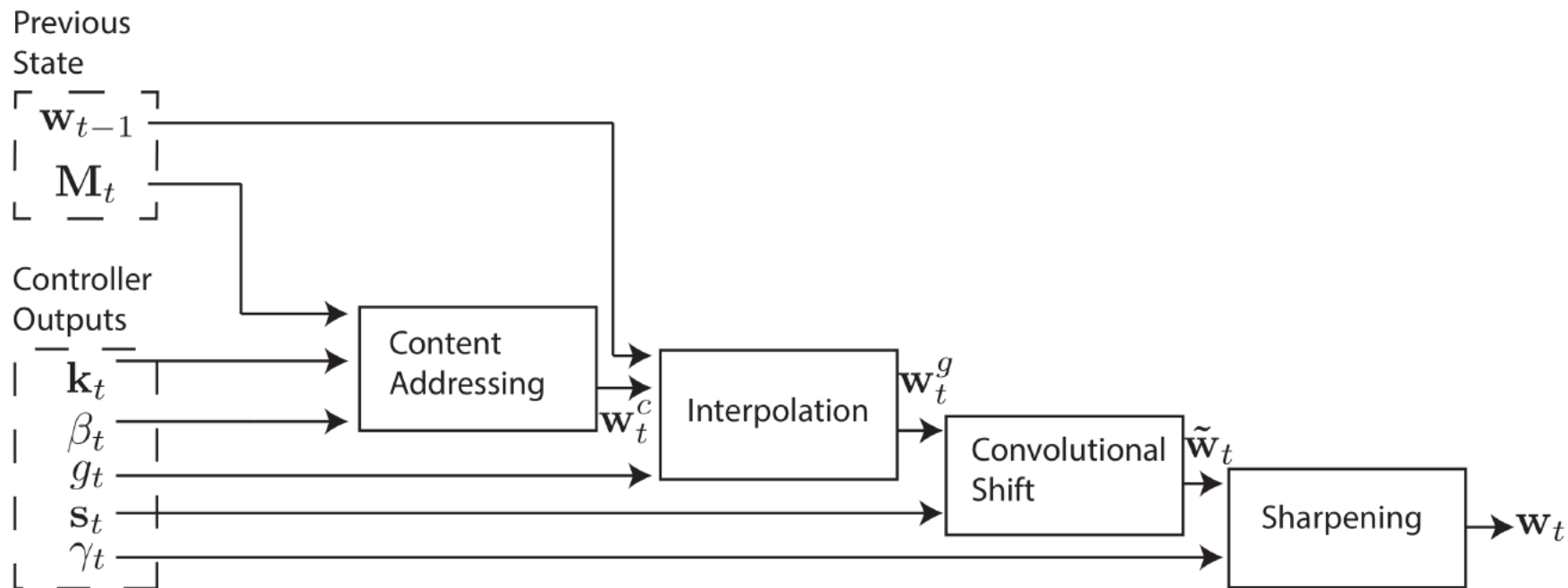
$$\sum_i w_t(i) = 1, \quad 0 \leq w_t(i) \leq 1, \forall i.$$
$$\mathbf{r}_t \leftarrow \sum_i w_t(i) \mathbf{M}_t(i),$$

- 写操作：先擦除，后写入

$$\tilde{\mathbf{M}}_t(i) \leftarrow \mathbf{M}_{t-1}(i) [\mathbf{1} - w_t(i) \mathbf{e}_t],$$
$$\mathbf{M}_t(i) \leftarrow \tilde{\mathbf{M}}_t(i) + w_t(i) \mathbf{a}_t.$$

➤ 神经图灵机

- 为了保证整个模型的可导性，不采用只读取或写入一个记忆单元的方式，提出了“模糊读写”的方式：
 - 寻址：确定“注意力”向量 w_t



The end