面向自然语言处理的机器学习
近一年主要进展

孙　栩

北京大学

xusun@pku.edu.cn

主要内容

- 回顾

- 主要进展
  - 全局学习 (Global Learning)
  - 动态记忆网络 (Dynamic Memory Networks)
  - 深度增强学习 (Deep Reinforcement Learning)

- 展望

回顾

- **近几年，机器学习领域的重要突破集中于深度学习领域**
  - 序列到序列模型(Sequence to Sequence Model)
  - 注意机制(Attention Mechanism)
  - 深度残差学习(Deep Residual Learning)
  - 神经图灵机(Neural Turing Machine)
  - 丢弃算法(DropOut)
  - 深度增强学习(Deep Reinforcement Learning)

回顾

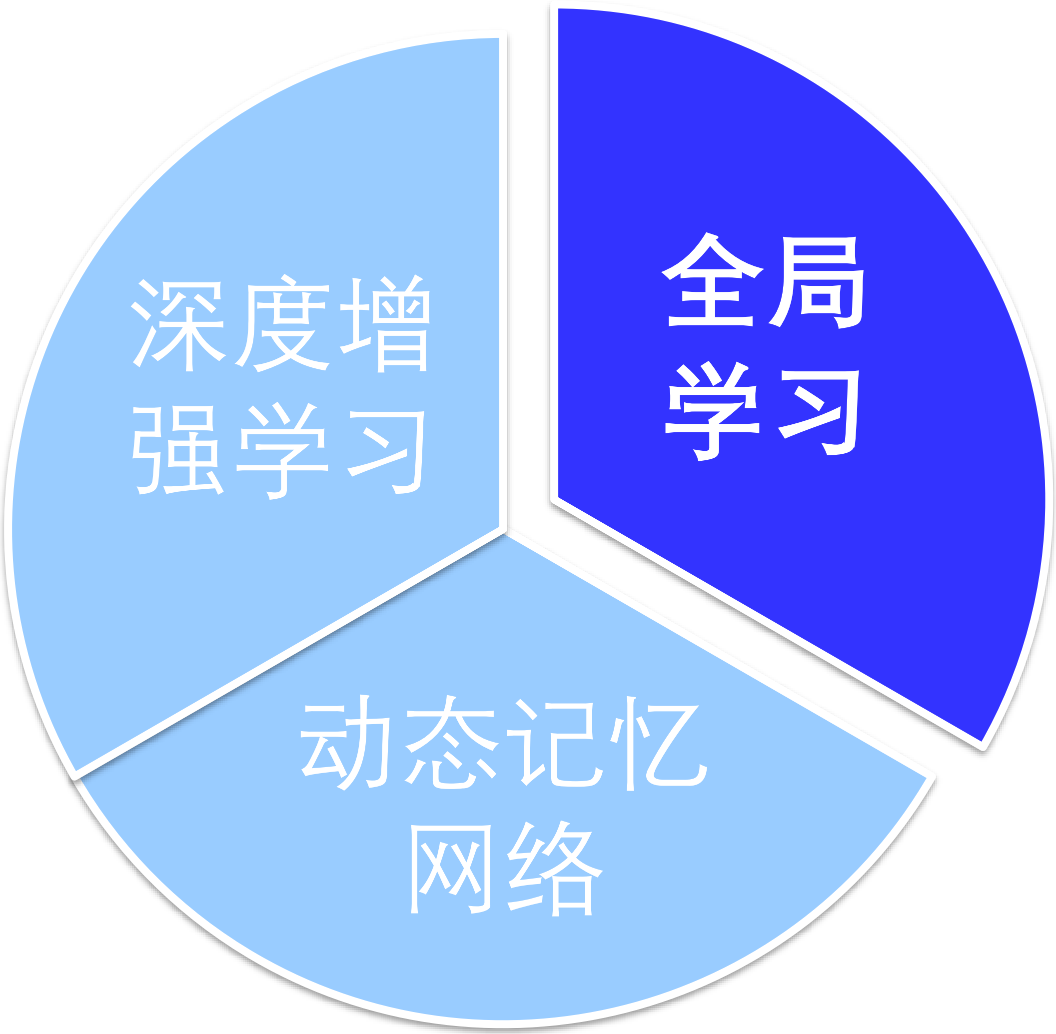
- 近一年来，在主流会议中有以下工作发表
  - ICML 2016 Best Papers
    - Dueling Network Architectures for Deep Reinforcement Learning (Wang et al.)
    - Pixel Recurrent Neural Networks (van den Oord et al.)
    - Ensuring Rapid Mixing and Low Bias for Asynchronous Gibbs Sampling (Sa et al.)
  - NIPS 2015 Best Papers
    - Competitive Distribution Estimation: Why is Good-Turing Good (Orlitsky & Suresh)
    - Fast Convergence of Regularized Learning in Games (Syrgkanis et al.)

- 这些工作大多不与自然语言处理直接相关
  - 基础理论
  - 游戏对抗
  - 图像生成

回顾：以自然语言处理的视角

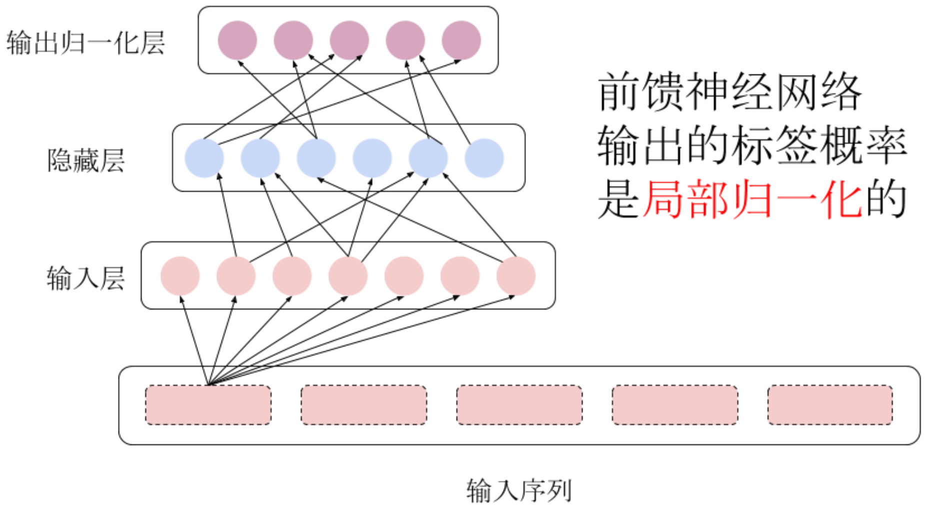
- 自然语言处理相关领域，近一年有一些值得注意的进展
  - 全局学习
  - 动态记忆网络
  - 深度增强学习
- 这些进展主要解决以下挑战
  - 单点分类难以捕捉输出的全局结构化信息
  - 模型与任务紧密耦合，缺少通用化深度学习框架

主要进展



全局学习

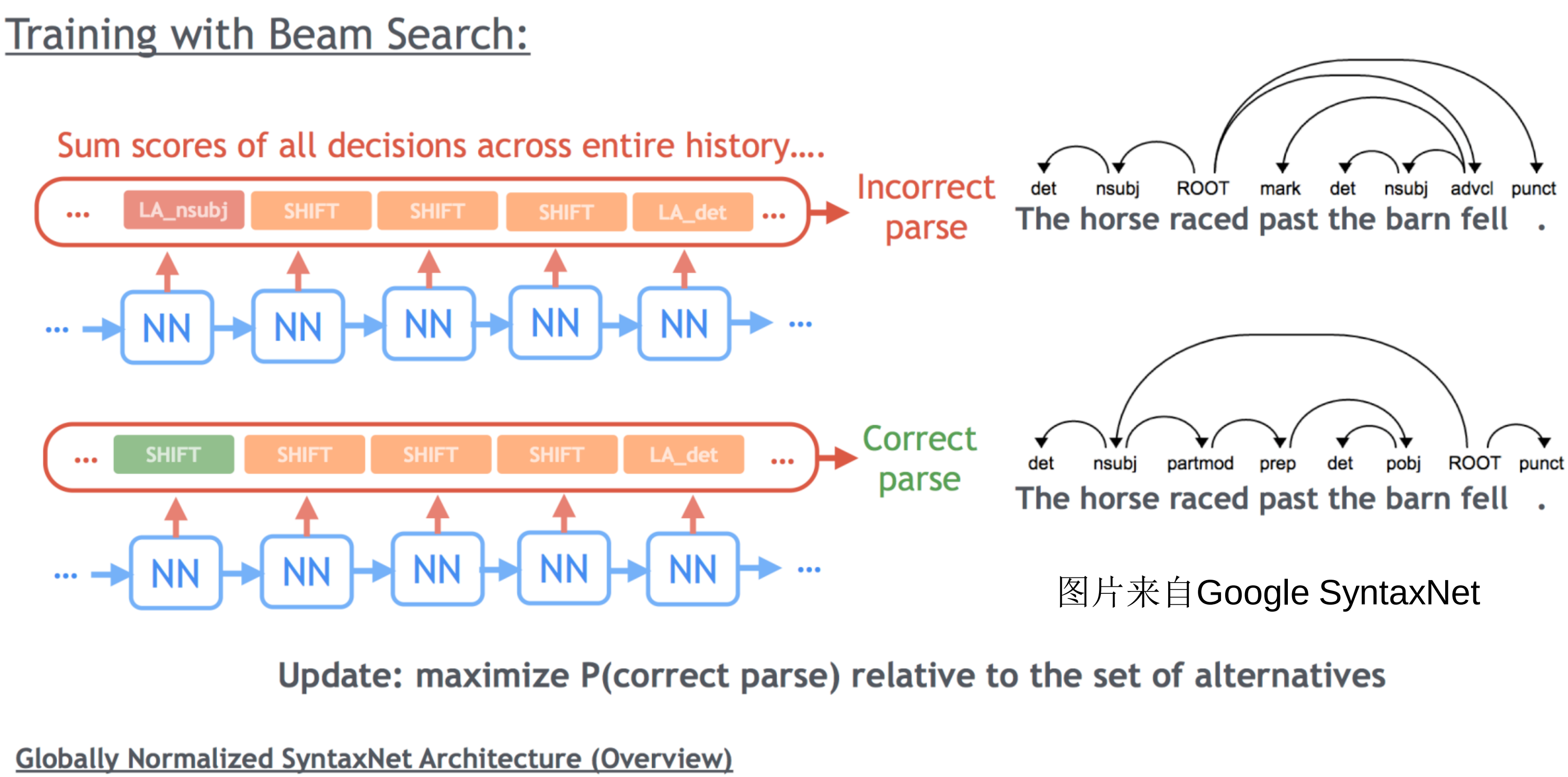
- 借鉴CRF全局归一化的思想
  - 最大化输出序列在所有可能输出序列中的概率
  - 而不是最大化每一步的输出在该步所有可能输出中的概率
- 不严格的类比
  - 感知器 -> 结构化感知器
  - 局部学习 -> 全局学习
- 应用于基于转移的神经网络
  - Zhou et al. (ACL 2015)
    - 依存句法分析
  - Andor et al. (ACL 2016)
    - 依存句法分析、词性标注、语句压缩



前馈神经网络输出的标签概率是局部归一化的

全局学习

◻ 示例：基于转移的依存句法分析



图片来自Google SyntaxNet

2016年5月12日开源公布的Google SyntaxNet是该算法实现

全局学习

- **优势**
  - 速度较快
    - 前馈神经网络较循环神经网络有较大的计算速度优势
    - beam search限定了可能输出序列的空间
  - 效果较好
    - 解决了局部学习的标签偏置问题

Method	WSJ UAS	WSJ LAS	Union-News UAS	Union-News LAS	Union-Web UAS	Union-Web LAS	Union-QTB UAS	Union-QTB LAS
Martins et al. (2013)	92.89	90.55	93.10	91.13	88.23	85.04	94.21	91.54
Zhang and McDonald (2014)	93.22	91.02	93.32	91.48	88.65	85.59	93.37	90.69
Weiss et al. (2015)	93.99	92.05	93.91	92.25	89.29	86.44	94.17	92.06
Alberti et al. (2015)	94.23	92.36	94.10	92.55	89.55	86.85	94.74	93.04
Our Local (B=1)	93.17	91.18	93.11	91.46	88.42	85.58	92.49	90.38
Our Local (B=32)	93.58	91.66	93.65	92.03	88.96	86.17	93.22	91.17
Our Global (B=32)	**94.41**	**92.55**	**94.44**	**92.93**	**90.17**	**87.54**	**95.40**	**93.64**

Table 2: Final English dependency parsing test set results (without tri-training for any method).

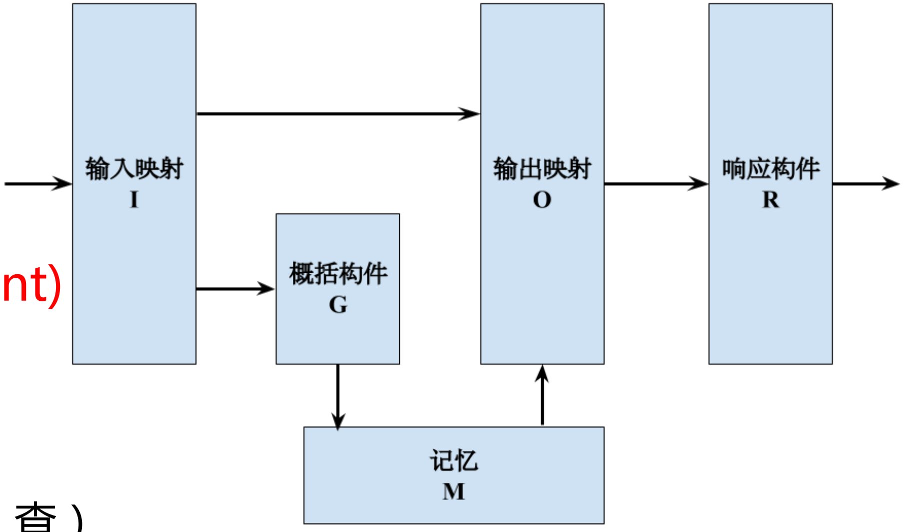
- **为深度学习在结构化预测中的应用提供了新思路**

主要进展

记忆网络

- **记忆网络的核心思想：模拟计算机存储**
- **一般构成**
  - 记忆 (Memory)
    - 存储器：可以看成由对象构成的数组
  - 输入映射 (Input Map)
    - 将输入转换为某种表示
  - 概括构件 (Generalization Component)
    - 根据输入更新记忆（增、删、改）
  - 输出映射 (Output Map)
    - 根据输入选择记忆中有用的内容输出（查）
  - 响应构件 (Response Component)
    - 根据输出做出响应

Memory Networks
(Weston et al., ICLR 2015)

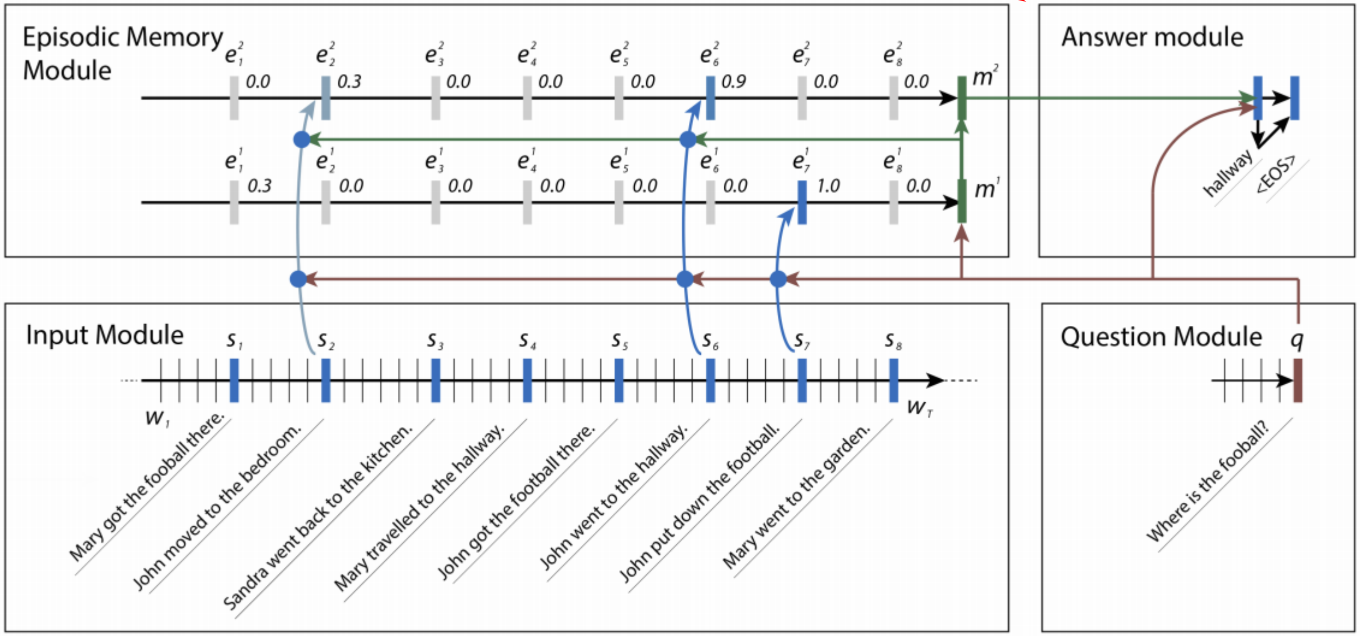


- **相关工作：神经图灵机 (Neural Turing Machine)**

动态记忆网络（Dynamic Memory Networks）

- 主要改进：片段记忆模块 (Episodic Memory Module)
  - 大致是记忆、概括构件和输出映射的整合
    - 转换/整合输入为记忆片段，多个记忆片段转换/整合为输出
  - 注意机制(attention)
    - 通过注意机制优化记忆片段

(Kumar et al., ICML 2016)



动态记忆网络

特点

- 整合了现阶段的几乎所有深度学习成果
  - 端到端模型 (Seq-to-seq Model)
  - 注意机制 (Attention Mechanism)
  - 门控循环单元 (Gated Recurrent Unit)
- 模块化
  - 每个模块有特定的目标
  - 模块耦合度低、每个模块内部可以单独设计
- 通用化
  - 模型适应几乎所有输出为序列的自然语言处理任务

未来发展

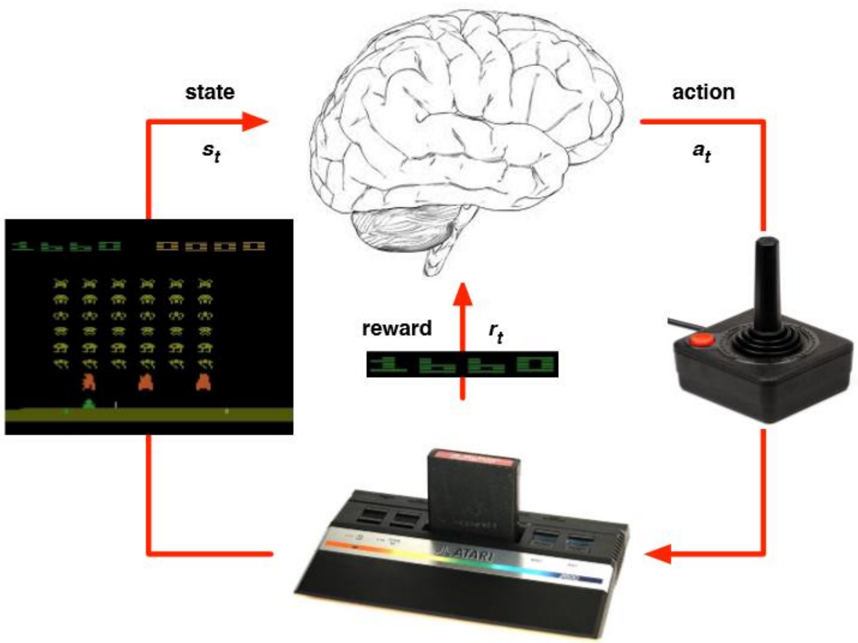
- 是否有模型能够近一步涵盖输出为树或图结构的任务？

主要进展

增强学习

何为增强学习？

- 决策学习
- 给定如下模型



- Agent是学习的对象
- 目标是最大化reward

图片来自David Silver

深度增强学习

- **深度增强学习是用神经网络表示**
  - 基于值的学习
    - 如Deep Q-Networks (DQN)
  - 基于策略的学习
    - 如Deep Policy Networks
  - 基于模型的学习
    - 如AlphaGo
- **深度增强学习首先在游戏对抗中取得优异表现**
  - Mnih et al. (2013), Mnih et al. (Nature 2015)
- **目前该技术开始被用于自然语言处理中的特定任务**
  - 基于文本的游戏
  - 文本生成、对话机器人

图片来自David Silver, Google DeepMind

深度增强学习

- **学会玩基于文本的游戏** (Narasimhan et al., EMNLP 2015; He et al., ACL 2016)
  - 自然语言理解任务：通过填写部分关键语句，完成故事的编写
    - 不同的填写方式会导致不同的故事结局
  - 目标：填写最少的句子完成一个有“好结局”的故事
  - 深度增强学习的应用(DQN架构)
    - state：每次的描述文本
    - action：固定的选项
    - reward：对每个结局根据情感偏向有得分，对中间步骤给定负值

(a) Parser-based (b) Choiced-based (c) Hypertext-based

Figure 3: Different types of text games

深度增强学习

更多应用

- 消除无用句以改进对话生成 (Li et al., EMNLP 2016)
- 引入外部佐证以改进信息抽取 (Narasimhan et al., EMNLP 2016)
- 通过引用排序以进行指代消解 (Clark & Manning, EMNLP 2016)
- 预测和追踪热点讨论话题 (He et al., EMNLP 2016)

特点

- 可以无限地自动生成训练数据
- 深度学习的表示能力扩展了增强学习的能力

前景广阔

The future of deep learning

Unsupervised learning[91–98] had a catalytic effect in reviving interest in deep learning, but has since been overshadowed by the successes of purely supervised learning. Although we have not focused on it in this Review, we expect unsupervised learning to become far more important in the longer term. Human and animal learning is largely unsupervised: we discover the structure of the world by observing it, not by being told the name of every object.

Human vision is an active process that sequentially samples the optic array in an intelligent, task-specific way using a small, high-resolution fovea with a large, low-resolution surround. We expect much of the future progress in vision to come from systems that are trained end-to-end and combine ConvNets with RNNs that use reinforcement learning to decide where to look. Systems combining deep learning and reinforcement learning are in their infancy, but they already outperform passive vision systems[99] at classification tasks and produce impressive results in learning to play many different video games[100].

Natural language understanding is another area in which deep learning is poised to make a large impact over the next few years. We expect systems that use RNNs to understand sentences or whole documents will become much better when they learn strategies for selectively attending to one part at a time[76,86].

Ultimately, major progress in artificial intelligence will come about through systems that combine representation learning with complex reasoning. Although deep learning and simple reasoning have been used for speech and handwriting recognition for a long time, new paradigms are needed to replace rule-based manipulation of symbolic expressions by operations on large vectors[101]. ■

主要观点：增强学习在未来的Deep Learning的发展中会有重要作用

Yann Lecun
Yoshua Bengio
Geoffery Hinton

Nature 2015

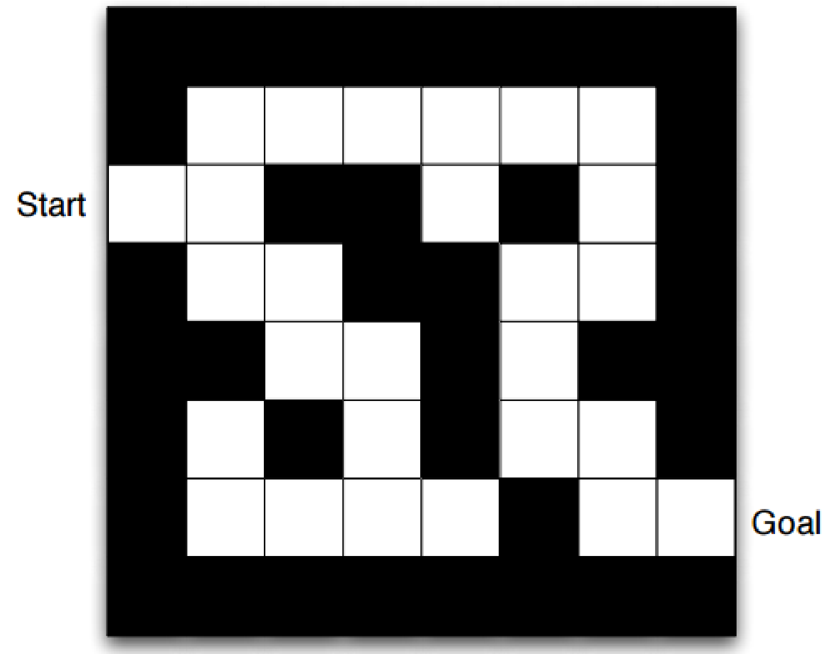
展望

- **总结**
  - 探索更为有效的结构化深度学习方式
  - 探索更为通用的计算模型
  - 探索深度增强学习在自然语言处理领域中的应用
- **困难与挑战**
  - 缺少严格的数学理论基础
  - 梯度计算带来的速度瓶颈
- **发展趋势**
  - 自然语言处理领域中的深度学习将进一步结构化、通用化
  - 更为有效的训练手段将推动深度学习进入一般实用领域
  - 深度增强学习将开辟新的领域

谢　谢！

Thank You

- **本质都是学最佳策略(Policy)，但Agent的实现方式有多种**
  - 基于值的学习 (Value-based Learning)
  - 基于策略的学习(Policy-based Learning)
  - 基于模型的学习(Model-based Learning)
- **之后会用一个简单例子来辅助说明：走迷宫**
  - 模型
    - state：agent的当前位置
    - reward：每步获得-1
    - action：上、下、左、右
      移动一格
  - agent目标
    - 最少步数走出迷宫



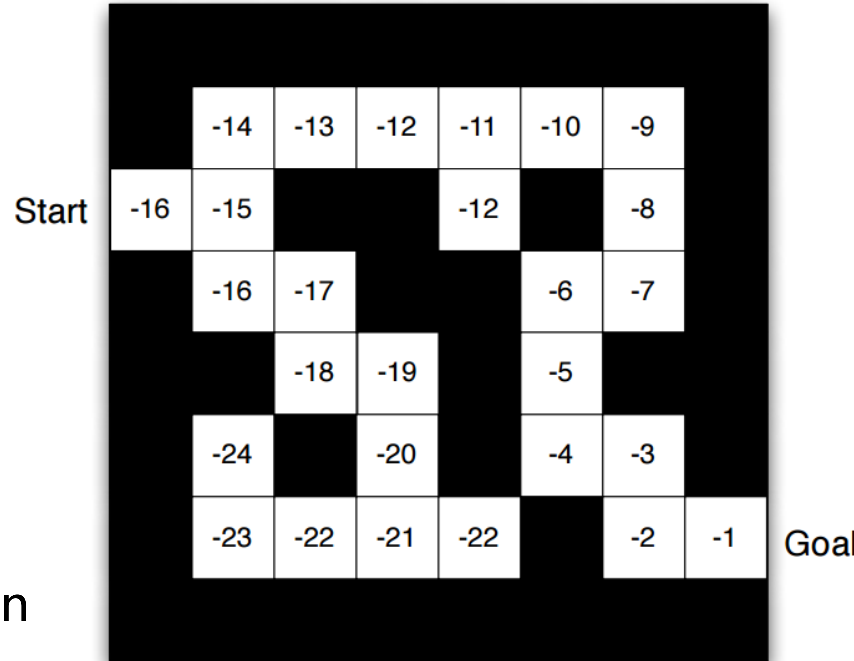
图片来自David Silver

基于值的学习 (Value-based Learning)

- 为每个state和/或action赋予一个价值：动作价值函数
- 将问题转化为求解最优价值函数的问题
- 典型实现：Q-learning

走迷宫：数值代表每个state的价值

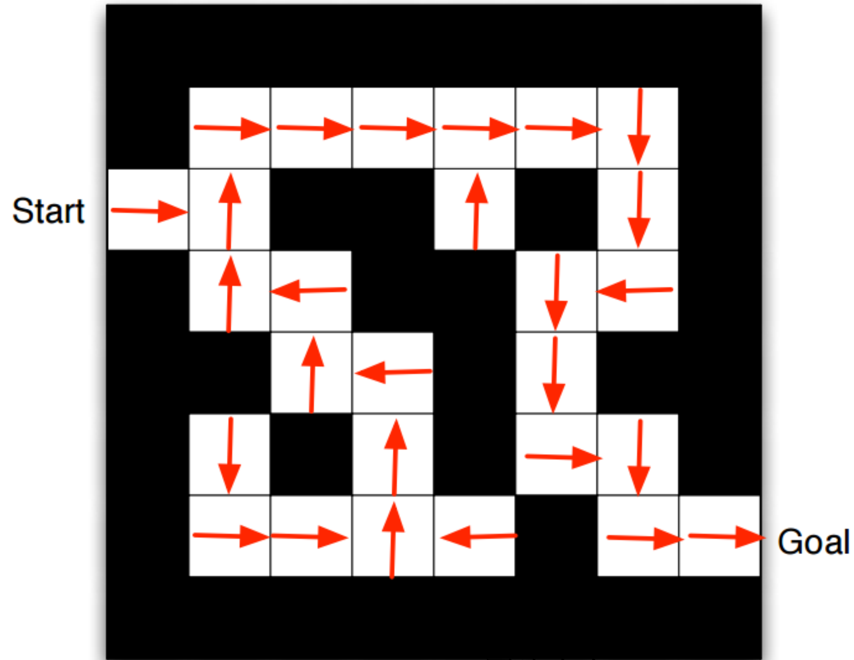
- 不了解模型
  - 不知道有哪些state
  - 不知道某个state可以有哪些action
- 极为简略的价值函数
  - 价值仅与state相关
- 在每个state下
  - 选择可以导致最大价值的action



图片来自David Silver

增强学习

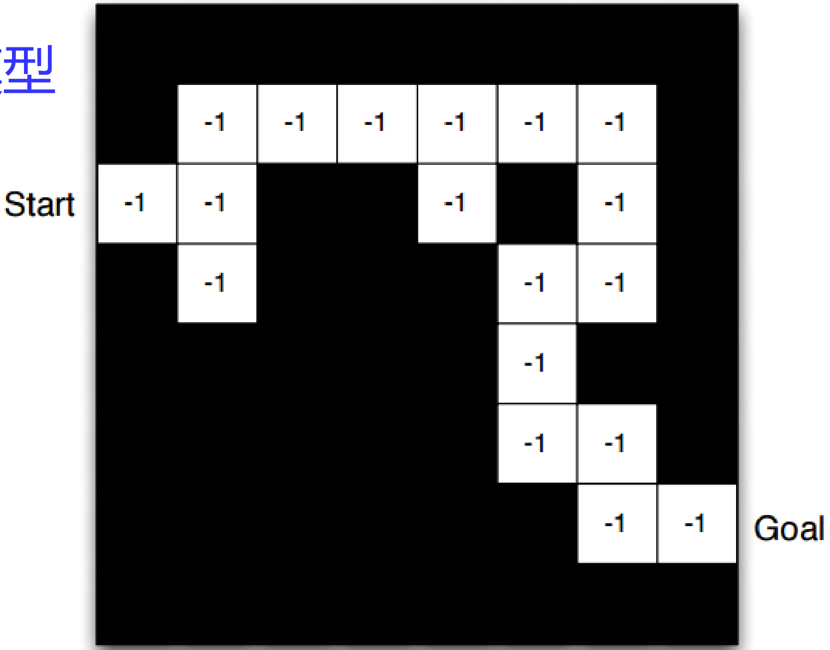
- 基于策略的学习(Policy-based Learning)
  - 直接根据reward或result优化policy
  - 不严格的说类似感知器：如果结果好，那么这个动作就好
  - 典型实现：Policy Gradient
- 走迷宫：方向代表每个state的action
  - 同样不了解模型
  - 每个state下
    - 根据给定的策略选择action



图片来自David Silver

增强学习

- **基于模型的学习(Model-based Learning)**
  - 建立environment的模型
    - 接下来的state是什么，对应的reward是什么
    - 除非环境完全已知，这个模型大多数情况下是不完善的
  - 通过规划的方式得到最优动作
- **走迷宫：建构环境**
  - 建立或估计了environment的模型
    - 有哪些state？
    - reward是怎样的？
  - 在每个state下
    - 通过规划选择最优action



图片来自David Silver