

Q-LDA: Uncovering Latent Patterns in Text-based Sequential Decision Processes

本文来源

本文发表于 NIPS2017，作者是微软的 Jianshu Chen et al.。

推荐理由

它首次将传统主题模型 LDA 与强化学习 Q 学习算法结合起来，应用于基于文本序列决策的过程。增加了决策过程的可解释性。

主要内容

基于文本序列决策过程是这样的：一种动作空间和状态空间都是以自然语言呈现的文本游戏。在每个时间步骤，决策者（agent）观察一个描述当前游戏环境的文本文档（observation text），以及表征不同可能动作的若干文本文档（action texts）。基于上述这些，agent 选择动作之一并更新状态，直到 agent 达到最终状态结束。

本文提出的模型 q-lda 有两个主题模型实例，一个用于 observation text，另一个用于 action texts。

训练这两个主题模型的过程是这样的：

1. 对 observation text 进行采样
2. 对所有的 action texts 进行采样
3. 通过公式得到 action
4. 通过公式得到 reward
5. 计算更新第 1 和 2 步的采样参数

并通过修改 Q 学习的 Q 公式，将主题模型的参数 θ 加入到公式中，使 Q 学习与主题模型 LDA 融合。

实验结果

最后分别在两个游戏 "Saving John" 和 "machine of Death" 上进行了实验，Q-LDA 比 DRRN(1-layer), DRRN(2-layer), MA-DQN(2-layer) 效果好。