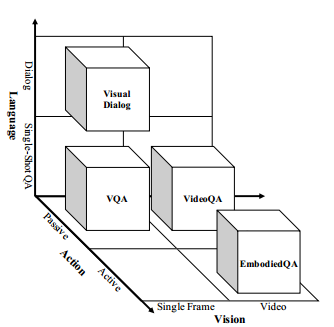
第6周汇报

——张溢炉

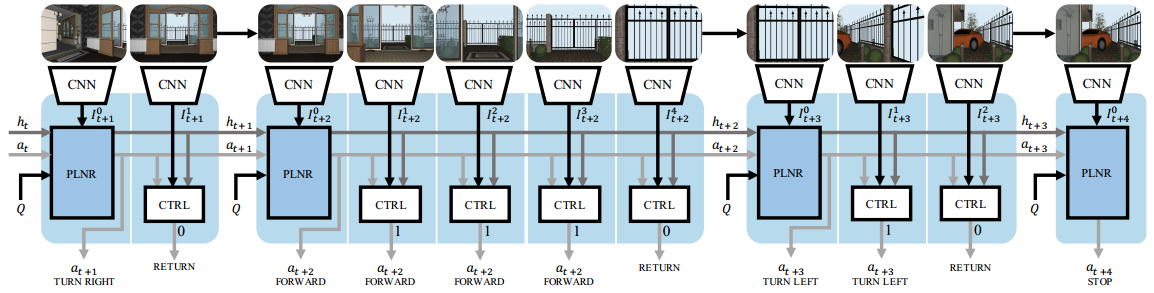
1. 《Embodied Question Answering》总结
2. 内容摘要

文章提出的EQA在视觉、NLP的基础上加了action，相对VQA而言，EQA任务还需通过探索在房间中找到目标物体，然后回答主人的问题。如果再加上拿放物体等动作，可以应用到家庭服务机器人。

1. 总结及理解

* EQA融合了几乎所有AI技术，如CV、NLP、IL、RL、知识图谱等，目的是使agent具有像人一样的感知、理解、决策的能力，最终应该会与传统的定位、机械手等技术相结合构成人形机器人，但这似乎还很遥远，涉及了太多技术，这里充满了机遇和挑战，需要团队合作进行整合各方面技术。
* 这里面涉及很多技术，一个人做的话工作量太大，需要团队合作；另外，3D环境、CV、NLP、RL、IL等整合的计算量很大，目前VQA需要的运算时间已经挺长了（在1080上训练简单模型遍历一遍VQA数据集需要2小时左右，复杂的需要5、6个小时），那么，如果训练集很大，EQA的训练会很长
* 对于解决实际问题，有些场景可能使用传统的slam等技术能更好的解决问题，如机器人可能在某个特定环境下，room不会改变，只会改变一些摆设，room navigator时可以尝试使用传统技术来导航，导航到对应房间之后再采用AI技术。但是科研就是需要不断尝试新的方法，即使它目前效果并不是那么好。
* 下图直观的展示了visual dialog、VQA、videoQA、EQA之间的关系，videoQA、EQA正上方是空白的，可看成是Video Dialog、Embodied Dialog，有一定的挖掘价值。
* 可以通过多任务的预训练来共享权重，使一个模型具有读取多种隐性特征的能力，如文中CNN的预训练
* 模型如下图，其中PLNR模块产生动作，动作集是{前进，左转（9度），右转（9度），停止}，CTRL只决定是否继续执行当前动作。另外question经过编码后输如，answer模块通过最后5帧图像和问题产生答案。

它的训练过程如下：先通过模仿学习预训练PLAR、CTRL和answer模块，再使用强化学习的方法微调PLAR、CTRL模块。



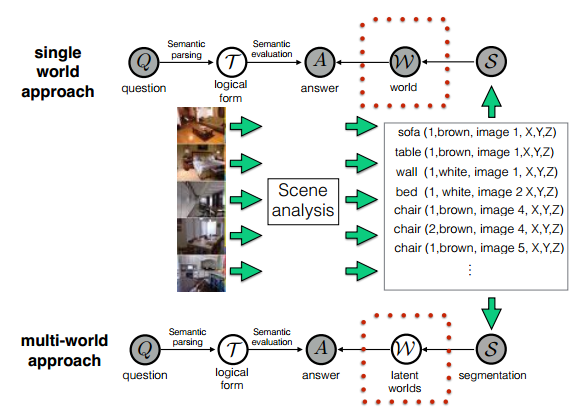
1. 《Interactions in Multiagent Systems: Fairness, Social Optimality and Individual Rationality》总结
2. 内容摘要

本文主要讲述了多智能体，内容较多，自己浏览了一遍，还未仔细看文中公式，对多智能体有了初步的了解。

1. 总结及理解

* 类似于人类决策的3个假设：智能体会根据自身经验进行行推理论证，根据不同环境作出有利于自身的决定；智能体也会综合其他智能体的利益来作决定，就像人遵从个人利益、集体利益和社会利益那样；另外，本文还利用了认知心理学的知识，如人们会对对自己友善的人友善。
* 在多智能体系统中会面临利益冲突、信息有限和回报等问题，多智能体环境可分为合作环境和竞争环境。在合作环境下，智能体注重集体利益和社会利益，在竞争环境下智能体注重个体利益。在多智能体系统中需要均衡Fairness, Social Optimality和Individual Rationality。
* 纳什平衡（Nash equilibrium）：为了使利益最大化，应该在对手出不同策略的时候我们的收益都相等（不然在游戏中，对方可以改变不同策略的概率让我们的期望收入减少）。这种方法的一个好处是不需要沟通，可以自己做决策。纯纳什平衡有时并不能公平对待每个agent，如果要获得fairness，得采用混合纳什平衡，给不同策略赋予一定的概率。
* 传统的由superagent作决策的方法有以下几个缺点：
* 因为计算量会随着agent的增加而不断增大，不利于扩展
* 单个agent都必须与superagent通讯，通讯花费大，也可能相互干扰
* 一旦superagent出现故障，会造成整个系统崩溃
* 文中主要以两个agent的游戏来介绍算法，涉及只玩一次和不断玩下去的情况。玩一次的游戏意味着大家只有一次机会，大家都会比较在乎结果，容易引发矛盾；而不断玩下去的游戏，大家对一次游戏的结果并不是那么在乎，可以寻找一个策略使得大家总的回报差不多。
* 第3章加入了很多人类的情绪在模型中，比如厌恶不公平现象、人不犯我我不犯人和人若犯我我必犯人的思想、惩罚违背集体利益的人等，感觉将想法转换成数学表达是一个有趣的过程，自己会在之后的学习中好好体会，比如调整下系数改变对某个方面的关注程度。
* 公平的策略有利于解决各个agent 之间利益冲突的问题，从而增进合作
* 第4章讲述合作环境中的social optimality，我的理解是将共同利益最大化。各个agent有着共同的利益和相同的回报函数，个体回报的增加会促使集体回报的增加，但有时个体回报的增加会导致其他agent回报的损失，从而导致集体回报的减少，所以单个agent在行动时除了考虑自身回报，还得考虑集体回报。
* 在合作环境中，更偏向于合作，主要考虑其中的fairness、social optimality；在竞争环境中，更偏向于竞争，主要考虑其中的individual rationality、social optimality。这两者之间也没有明显的界限，有时也可能相互转变。如当前的中美贸易战，从之前更偏向于合作的合作关系演变成互相伤害的竞争关系。
* 文章主要以两者游戏的方式来讲述这些方法，多智能体系统也可以通过随机配对的方式简化为两个智能体的问题。其中主要运用到强化学习、博弈论、认知心理学等方法，可以通过改变不同策略或action的概率来改变回报。
* 以后继续看需要注意的方面
* 如何设计多智能体系统环境并以数学模型来表示？是否类似于强化学习的隐马尔科夫模型？
* 如何表示及实现Fairness, Social Optimality和Individual Rationality均衡？通过调节公式中不同项的系数？
* 如何学习策略？
* 如何实现智能体之间有效的沟通？减小通讯的信息量？如何实现各智能体独立作出决策？如：未来自动驾驶技术及各种无人驾驶的落地，如果各智能体间需要通讯才能作出决策，那么很可能出现通讯混乱的情况，最好各智能体能独立作出决策，在决策过程中也需考虑Fairness, Social Optimality和Individual Rationality等因素。

1. 《A Multi-World Approach to Question Answering about Real-World Scenes based on Uncertain Input》总结
2. 总结及理解

* 此篇论文是《deep learning-based models have dominated standard VQA benchmarks》中提到的标准VQA基准测试之一，自己粗略读了一下，了解一下以前的方法，有需要再细读
* 模型如下图所示，上半部single-world approach是之前的方法，其在图像分割S只产生一种对图像的理解W，单一的理解势必在回答问题时会有些局限。于是作者提出了下半部的muti-world approach。这两个方法将当前NLP和CV结合起来，是打破客观（CV）与主观（自然语言）的一种尝试。图像分割会提取图像中的object及其属性、位置等信息。根据x、y、z三维坐标可以判断各个object之间的位置关系，如y轴方向为从左到右，那么y值较小的物体在左边。
* VQA的问题和答案具有主观性，不同的人可能有不同的问题、答案及表示方式，而图像具有客观性，但现实中图像可能会有干扰和被遮挡的情况，而且图像一般是二维的。这给VQA任务带来了很多挑战。如
* 同一个物体可能有多个名称，这个问题可以利用词向量来解决，在词向量空间中词义相近的词的坐标相距比较近
* 图像存在被遮挡和截断的情况，这不利于物体识别
* 人类回答某个问题的句子可能比较长，这会增加产生答案的难度
* 物体的一些状态比较难检测，如门是否开着
* 疑问
* 产生问题、图像和答案表示的具体模型？
* 怎么获得图像中物体的三维坐标？