

# 我的科研经历和未来研究、学习计划

“ 罗敏中

我的研究经历主要分为三个阶段:

- **1.应用于物理场景或受物理模型启发的算法**:用于物理应用如能谱识别,辐照实验图片生成的算法,或基于物理模型的统计套利,彩票套利和分类优化算法(持续至今,仍在探索);
- **2.智能符号计算**:自动公式推导和定理证明(有阶段性的成果);
- **3.从概念学习到逻辑推理**:human-level的智能算法(当前和未来的主要方向,最终可以反过来解决第二个符号计算的问题);

其中,和物理相关工作是受本科专业和硕士所在研究机构的氛围影响展开的.首先我引入了一些计算机和AI的方法做物理场景的应用研究,比如构建Kafka为流处理系统的加速器数据分析应用、能谱识别、辐照实验图片生成的算法等,反过来,受物理模型的启发,也将物理模型的原理用于改进其他场景的算法,如最大熵分类方法、基于吉布斯自由能的市场趋势判别、基于巨正则系综的体育彩票套利模型等等;

符号计算是尝试赋予计算机数理逻辑的推导能力,是硕士期间自己选的一个感兴趣的研究方向,起源于在做数学证明题时想到能否机械地设计通用的数学证明方法,然后调研时发现这个分支叫符号计算,是一个方兴未艾的研究方向.正是在符号计算方法创新上的思考和努力让我走向了对人工智能的研究;

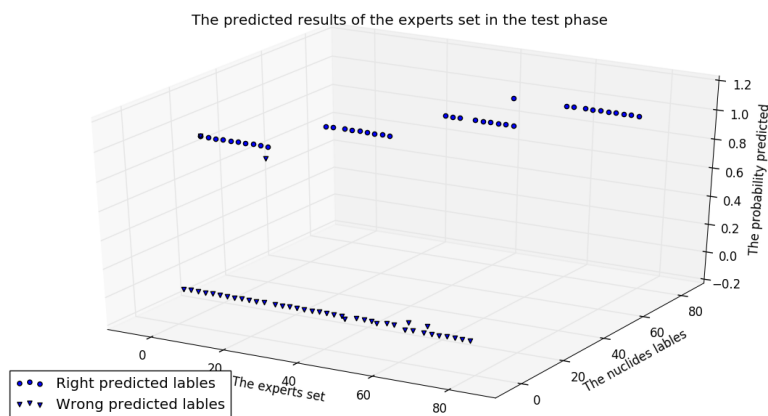
从概念学习到逻辑推理的智能方面,我尤其是想结合认知心理学和神经科学的实验启发,引入相关数学理论工具,基于虚拟的3D环境Fuzzy World,来构建具备human-level的智能算法;

## 1.应用于物理场景或受物理模型启发的算法研究

这些工作主要是我的早期工作,这个方向的工作内容也许博士以后不会涉及太多,但是其中一些思维模式仍然有价值.虽然我的学科背景是偏工程物理方向的,但是这并非一个做人工智能研究的劣势,因为很多研究思维和方式是可迁移的.

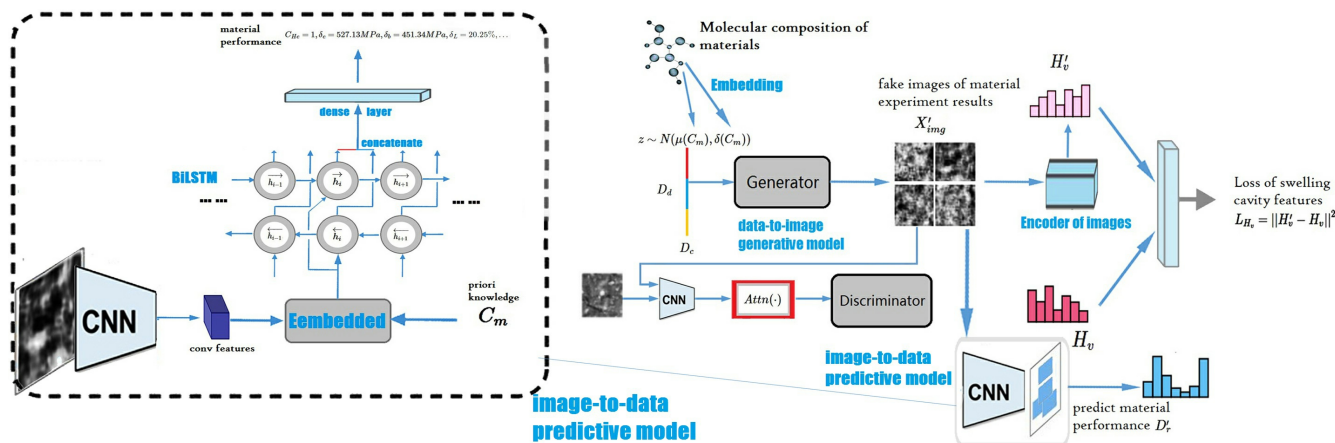
### 复杂能谱识别的在线学习算法

**Core Idea:**能谱由于本底干扰和成像的随机性是一个复杂分布的模式识别问题.在K-Means方法作为基本解决方法的基础上,给出了多个样本集上训练的弱分类器集合  $\{f^{(i)}(x, \theta_i)\}$  的加权学习算法及其在线学习形式,并证明了集成分类器损失  $\sum_{t=1}^T w^{(t)} \odot v_t$  的上界;



## 辐照材料肿胀实验图片的生成模型

**Core Idea:**材料受辐照而肿胀,在电子显微镜下的成像  $X_i$  蕴含了其性能参数的改变信息  $D_r$ ,而性能参数的改变信息  $D$  也决定了成像  $X_i$  的形式;本研究分别构建了data-to-image的生成模型  $\mathbb{P}(X_i|D, C_m)$  和image-to-data的预测模型  $\mathbb{P}(D_r|X_i, C_m)$ ,其中  $C_m$  是由材料分子决定的特征向量;



## 基于最大熵原理的分类方法

**Core Idea:**熵本质上是不确定性的度量,条件概率  $\mathbb{P}(Y|X)$  的对应熵  $H(P) = - \sum_{x,y} P(y, x) \log P(y|x)$ ,而最大熵的设定相当于使得不确定性最大,这在  $\mathbb{P}(Y|X)$  未知时是最保守有效的,这结合神经网络用可改进分类问题的效果(详细地,  $H(P)$  的上界可证得为  $-\sum_i p_i \{-\lambda_0 - \sum_r \lambda_r f_r(x_i)\}$ ,等价地对应层的损失可设为  $\lambda \odot (p \otimes f(x))$ ).

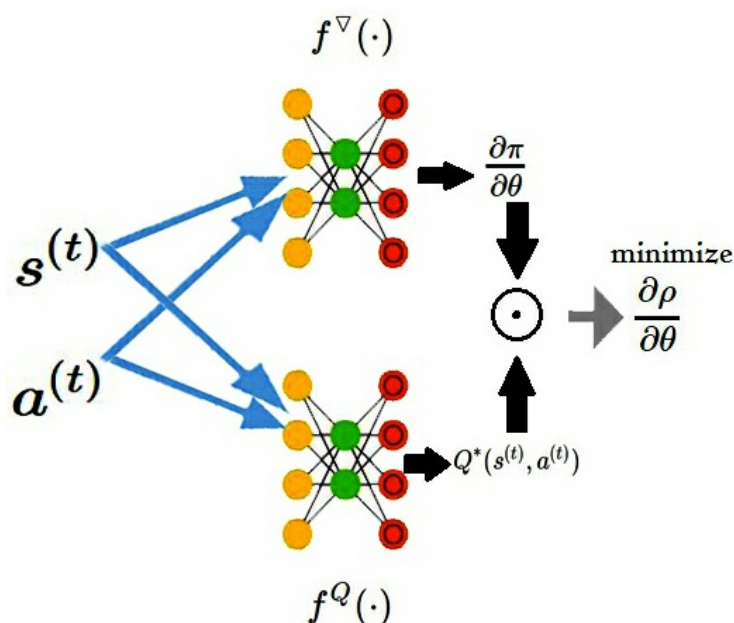
## 强化学习的正交策略梯度

**Core Idea:**对于强化学习问题,可以证明代价 $\rho$ 相对于策略网络的参数 $\theta$ 的梯度为

$$\frac{\partial \rho}{\partial \theta} = \sum_a d^\pi(s) \sum_a \frac{\partial \pi(a|s)}{\partial \theta} Q(a, s) \text{ 也就是 } \frac{\partial \pi(a|s)}{\partial \theta} \odot Q(a, s), \text{ 若 } \frac{\partial \rho}{\partial \theta} = 0, \text{ 那么可}$$

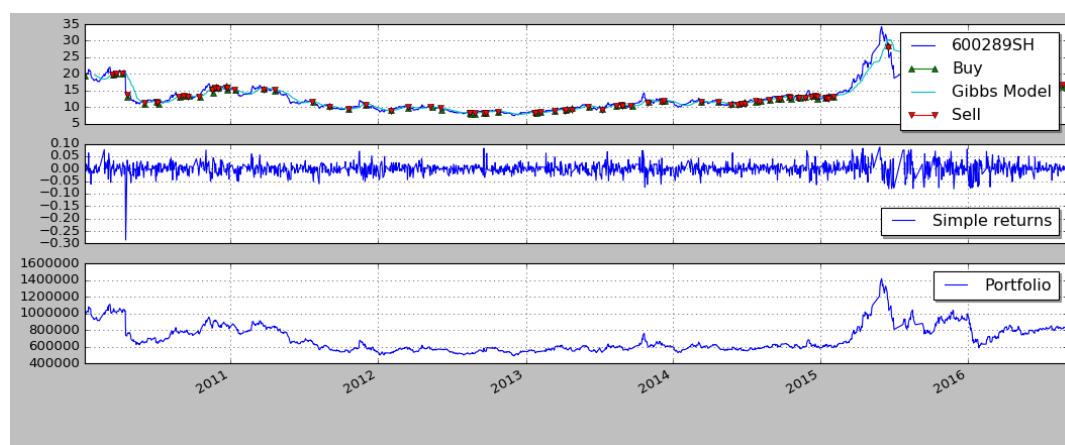
以说此时 $\frac{\partial \pi(a|s)}{\partial \theta}$ 和 $Q(a, s)$ 正交,同时此时代价最优;因此分别构建了可以预测梯度

$\frac{\partial \pi(a|s)}{\partial \theta}$ 的网络 $f^\nabla(\cdot)$ 和可以预测Q-Value的网络 $f^Q(\cdot)$ ,由此通过梯度下降可得最优的参数 $\theta^*$ ,进而可得最优策略 $a^* \sim \pi(a^{(t+1)} | s^{(t)}, \theta^*)$ ;



## Gibbs自由能用于市场趋势判别

**Core Idea:**Gibbs自由能 $\Delta G = \Delta H_{\text{sys}} - T\Delta S_{\text{sys}}$ 决定了反应的方向,那么是否可以使用物理内涵的模型来模拟市场趋势?首先将市场指标拆分为市场Hidden Feature  $X_h$ , Stable Factors  $X_s$ , Volatility Factors  $X_v$ ,再引入神经网络 $\mathbb{P}(\Delta G|X_h)$ ,  $\mathbb{P}(T|X_s)$ ,  $\mathbb{P}(\Delta S|X_v)$ 将因子映射到物理空间,最后利用 $\Delta H_{\text{sys}} = \Delta G + T\Delta S_{\text{sys}}$ 来预测市场方向;



## 巨正则系综的体育彩票套利模型

**Core Idea:**体育赛事的结果是参赛者基本面、博彩公司赔率、必发交易所行情、历史数据等诸多因子的综合作用复杂模型;首先通过爬虫程序收集实时数据,后端通过类巨正则系综的模型判断走向;巨正则系综的系统能量 $E$ 、压强 $P$ 和粒子数 $N$ (类比彩票中各类因子)会在某一平均值附近有一个起伏.体系是一个开放系统,与大热库大粒子源接触(类比投注市场),有能量交换,达到热平衡,温度相等;有粒子交换,达到化学平衡,化学势相等,源足够大(类比必发交易量足够大),温度和化学势确定.特征函数是马休(Massieu)函数 $J(\mu, V, T)$ .

```
root@master:~# python model_run.py
... ..
+++++
胜赔: 1.82
进球期望: 1.5 ± 0.25
失球期望: 0.5 ± 0.25
--- [0. 0.5 0.5] --- 置信度: 2
+++++
胜赔: 4.01
进球期望: 1.5 ± 0.25
失球期望: 2.25 ± 1.6875
--- [0.5 0.25 0.25] --- 置信度: 4
=====
厄斯特什 最近三场进球: [1.0, 0.0, 1.0] 最近三场失球: [2.0, 0.0, 1.0]
迪格弗斯 最近三场进球: [2.0, 4.0, 2.0] 最近三场失球: [1.0, 1.0, 1.0]
=====凯利指数=====
初始: 返还率: 93.29 凯利: 0.91 0.89 0.95
即时: 返还率: 93.33 凯利: 0.97 0.85 0.82
... ..
```

## SVM方法用于六爻断卦

**Core Idea:**周易六爻是通过打卦(相当于一种分布的采样)将客观事件映射到卦象的符号空间,其中涉及到时间的推演和取余等运算;SVM方法的引入相当于取代人工的断卦即符号空间到事件预测的过程;通过爬虫程序的收/发卦象符号数据以及体育比赛结果数据,用scikit-learn构建了"断卦"的SVM模型;

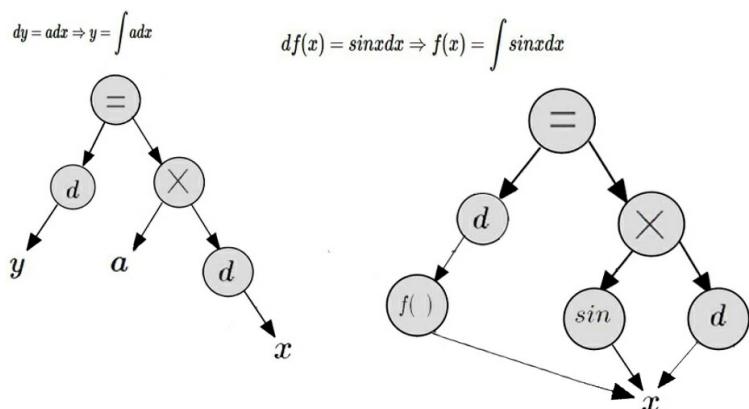
```
root@master:~# python main.py
... ..
#res:戊戌年 己未月 庚戌日 乙酉时
#res:日空:寅卯
#驿马—申 桃花—卯 日禄—申 贵人—丑未
#['腾蛇 父母戊子水 世',
# '勾陈 妻财戊戌土 世',
# '朱雀 子孙庚午火 官鬼戊申金',
# '青龙 官鬼辛酉金 应'...]
win draw lose
[0.25 0.25 0.5 ]
```

## 2.智能符号计算

符号计算的研究是我的硕士早期工作,使机器具备数理逻辑推导能力是一个极具挑战性的课题,一直以来主要靠基于规则的匹配算法并依赖大容量的先验知识库.我的主要目标是使机器具备类人的泛化能力很强的通用数理逻辑能力.

### 基于图嵌入和深度神经网络的公式推导

**Core Idea:**符号计算向来是基于代数方法的"硬推导",是否可以利用神经网络的拟合能力,实现模糊计算的公式推导?首先将公式表达为公式图形式,再基于公式图给出了对应的embedding方法及公式的feature space,最终神经网络 $\mathbb{P}(X_{t+1} | X_r, X_t)$ 可以根据最终目标公式 $X_r$ 和当前的公式 $X_t$ 给出预测下一步公式的 $X_{t+1}$ ;



## 3.构建从概念学习到逻辑推理的智能

随着我对智能符号计算的深入思考,我发现这个问题是概念学习和逻辑推理智能的一个**高阶特例**,因此我转入了这项研究.首先,我给"概念学习/逻辑推理"下了严格明确的定义;然后,为了训练/检验Agent的"概念学习/逻辑推理"能力,基于Python开发了Fuzzy World这样一个蕴含复杂任务的虚拟3D环境;到目前为止,提出了一种基于深度强化学习、融合语义和视觉信号的baseline model,并且有了下一步的研究思路.

### "概念学习/逻辑推理"问题的定义

首先需要引入**语义图**.根据大脑的功能性原理,即,大脑负责认知和推理的是不同的模块(那么负责学习概念学习和逻辑推理的模块应该也不一样),我们于是将语义图模块拆分为 $G_s$ 和 $G_l$ 分别负责概念认知和逻辑推理.

- 在**环境状态认知语义图** $G_s$ 中:谓词连接了实体或概念,所以 $G_s$ 是多个三元组组成的,比如 $is(moving, cow)$ 或者 $biggerThan(cow, desk)$ .
- 在**逻辑推断语义图** $G_l$ 中:是有因果联系的事件构成的,比如 $is(openning, light) \Rightarrow is(hot, room)$ 就是一个一阶逻辑的事件关联.

"概念学习/逻辑推理"的内涵相当广泛,在这里需要对其下一个约束性的明确定义;



**概念学习**:概念的定义即一阶谓词逻辑描述,比如例句**走到马的左边**,“马”是一个具象,非常容易识别,但是“走到”和“左边”这种抽象概念是较难学习泛化的,并且使建立在“具象”之上的;概念学习的目标是正确的状态认知语义图 $G_s$ .

我们抽象出下面三种范式(这些范式已经可以让Agent具备完成一些概念学习任务的能力,更复杂的任务也可以由这些范式组合来完成):

- 描述型范式 $d_m(o_i)$ :物体 $o_i$ 是 $d_m$ 的(例如,桌子是红色的);
- 二元型范式 $p_l(o_i, o_j)$ :物体 $o_i$ 是 $p_l$ 于物体 $o_j$ (例如,凳子相对东北方位于桌子/凳子矮于桌子);
- 描述型范式 $a_k(o_i)$ :对物体 $o_i$ 施加行动 $a_k$ (例如,转动凳子);

**逻辑推理**:逻辑推理是根据环境状态 $G_s^t$ 到 $G_s^{t+1}$ 的转移,得出环境规律知识图 $G_l^t$ ,并利用学得规律 $G_l^t$ 完成特定任务的过程;这个过程至少涉及到了观察、归纳、泛化推理;

符合定义的典型逻辑推理任务:

- 基于自然语言命令的逻辑推理:特别是长程的,比如命令是**使汽车的温度降低**,而当前环境蕴含逻辑:**打开(开关A)=>会转动(椅子B)=>会升温(电阻)=>会降温(汽车)**这样的长程逻辑.
- Agent之间的交互:通过和环境交互训练,Agent可以维护一个**逻辑规律语义graph**  $G_l$ ,既可以转化为自然语言传达给人类,也可以传递给另一个Agent;

## "概念学习/逻辑推理"训练/测试工具Fuzzy World的开发

**主要目的**:按照我们对"概念学习/逻辑推理"问题的定义,构建一个易用、易验证算法、干扰少的虚拟3D环境,专用于检测Agent的"概念学习/逻辑推理"能力;

在Fuzzy World中,Agent通过第一人称视角观察世界和交互操作后世界的变化,自然语言被用于引导和辅助Agent来认知世界完成Task;所以这个世界至少包含以下内容:

- **object**  $O$ :环境里的物体实体,可以是自定义导入的.obj格式模型.
- **state**  $S^t$ :物体在时刻 $t$ 的状态 $S^t$ :物体是否运动,其温度或者位置等等.
- **rule set**  $\{R_i\}$ :规则集合,比如“开关2打开则物体5的温度升高”这样的一阶逻辑,一个场景中的规则集合 $\{R_i\}$ 是给定的有限的,可以描述为物体及状态的转移映射即:

$$R_i : S^t(O_m) \rightarrow S^{t+1}(O_n)$$

以上是环境自身客观存在的性质,Agent可以通过**语言提示**和**主动尝试**探索物体性质和这些规则,来完成任务.一个关键的疑虑就是:**难道使用这个工具需要手工去编写大量的场景及其物体分布、逻辑规则?**不需要,因为工具已经内置了自动环境生成功能,可以生成海量的交互环境供Agent训练.

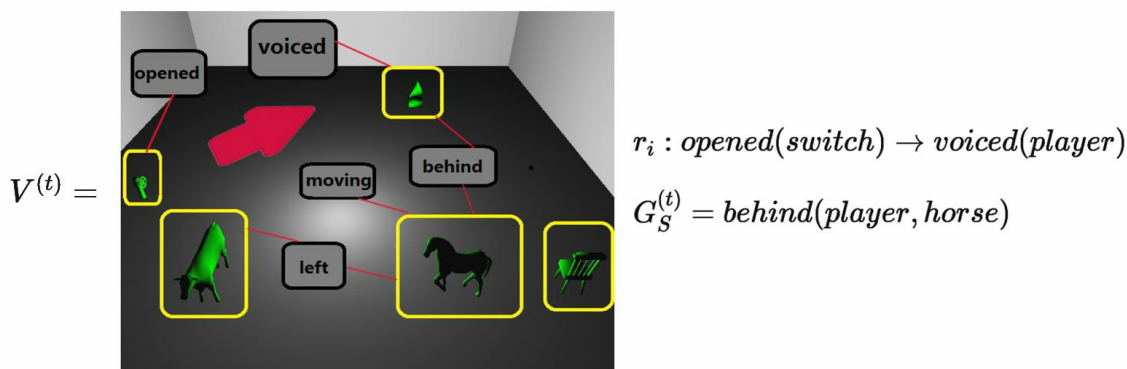
具体来说我规定了一种名为".fw"的文件来文本化地描述一个FuzzyWorld的场景物体及世界逻辑,一个".fw"的文件格式如下:

```

WORLD_NAME:samle_1;
WORLD_SIZE:44;
OBJ_NUM:5;
OBJ:ID:"0";MODEL:"./obj/box/jeep.obj";MESH:"jeep";POS:-1,-0.5;TEMP:"TEMP_NO
OBJ:ID:"1";MODEL:"./obj/box/music.obj";MESH:"music";POS:0,0;TEMP:"TEMP_NORM
OBJ:ID:"2";MODEL:"./obj/box/bed.obj";MESH:"bed";POS:-0.5,0;TEMP:"TEMP_NORMA
OBJ:ID:"3";MODEL:"./obj/box/cow.obj";MESH:"cow";POS:-0.5,-1;TEMP:"TEMP_NORM
OBJ:ID:"4";MODEL:"./obj/box/bed.obj";MESH:"bed";POS:0.5,-0.5;TEMP:"TEMP_NOR
RULE_NUM:4;
RULE:"3":"STOP"=>"4":"TEMP_HIGH";
RULE:"2":"MOVE"=>"2":"STOP";
RULE:"0":"TEMP_NORMAL"=>"3":"TEMP_HIGH";
RULE:"0":"MOVE"=>"2":"STOP";

```

所以通过这种文本格式,程序可以由文本描述来创建一个场景,亦可以由程序快速生成多个文本描述供创建大量的场景和逻辑规则来自动地构建海量训练集.



$l^{(t)} =$ “make the thing behind horse voiced;”

## "概念学习"任务的baseline model

**用于concept learning的静态数据集:**我们根据规则算法生成了大量的包含3D物体的场景图片 $V$ ,以及对应的语言提问 $L_Q$ ,语言回答 $L_A$ ,命令文本 $L_C$ ,行动表征 $A^{(t)}$ ;这个数据集适合用于端到端的神经网络训练.

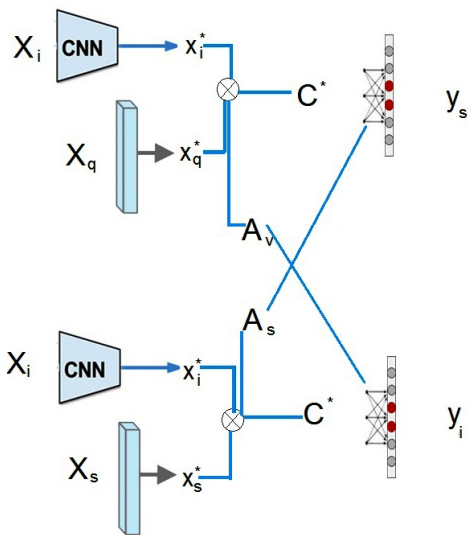
我们的目标是基于语义和视觉来学习概念(更high-level的目标是做逻辑推理).概念学习的考察有两点:第一,能否识别这个概念;第二,能否根据概念操作实物;因此第一步自然是需要生成大量谓词概念相关的**语料数据**和**图像数据**:我使用的方式是用算法基于某种rule来生成满足任务需求的数据,称为**Fuzzy World的静态数据集**;概念学习任务主要完成两个子任务:识别任务 $\mathcal{P}(L_A | V, L_Q)$ 和生成任务 $\mathcal{P}(A^{(t)} | V, L_C)$ :

- **识别任务** $\mathcal{P}(L_A | V, L_Q)$ :根据问题文本 $L_Q$ 和视觉图像 $V$ 给出回答 $L_A$ ;
- **生成任务** $\mathcal{P}(A^{(t)} | V, L_C)$ :根据命令文本 $L_C$ 和视觉图像 $V$ 给出下一步的行动表征 $A^{(t)}$ ;

**Fuzzy World的静态数据集**中另外一些有用的数据如下:

- 视觉图像  $V \in \mathbb{R}^{l \times 96 \times 96}$ : 图像的维度是  $96 \times 96$ ,  $l$  是数据集的容量, 下同;
- 问题文本  $L_Q$ : 文本数据, 对图像中内容的问题;
- 描述文本  $L_Q$ : 文本数据, 对图像中内容问题的描述型回答文本;
- 命令文本  $L_A$ : 文本数据, 对图像中内容的指令;
- 行动表征  $A^{(t)}$ : 主要是空间坐标信息, 因为指令文本主要是寻找空间位置;
- 物体实体id列表  $D_o$ : 每张图  $V_i$  内物体的id列表;
- 物体实体检测框列表  $D_b$ : 每张图  $V_i$  内物体的检测框列表;
- 物体实体尺寸列表  $D_s$ : 每张图  $V_i$  内物体的尺寸列表;
- 物体实体颜色列表  $D_c$ : 每张图  $V_i$  内物体的颜色列表;
- 语义图列表  $D_G$ : 每张图  $V_i$  内的状态语义图边集: 形如  $e_k(v_i, v_j)$ ;

baseline将概念特征  $A$  分解表征为  $A = [A_v, A_s]$  分别是其视觉和语义表征, 他们来自这两个网络最后一层和特征提取层之间的权重, 网络架构如下(当前的方法采用keras写, 正确率40%左右):



## "逻辑推理"任务的baseline model

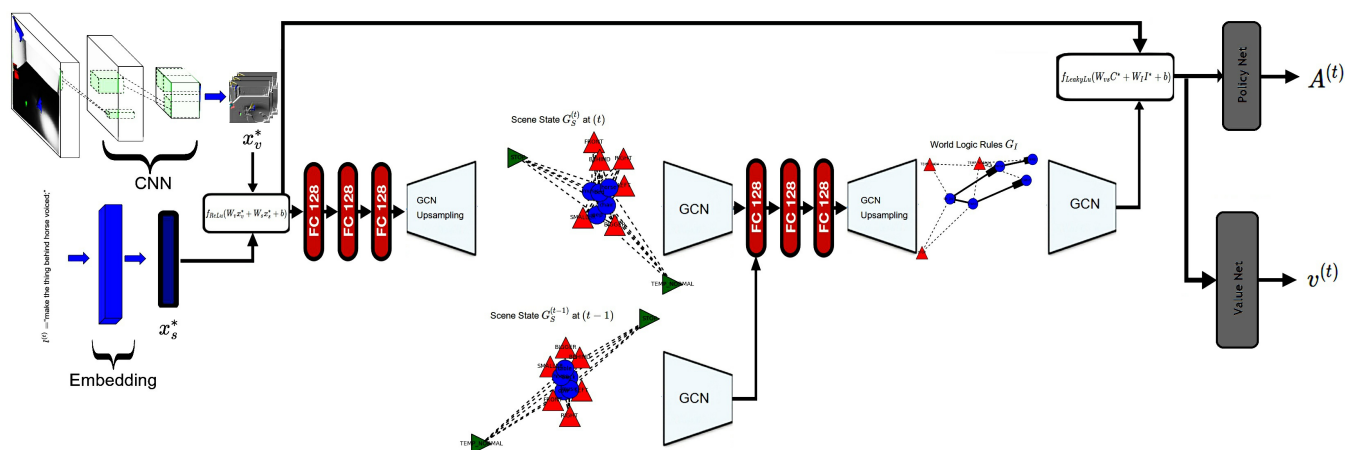
**用于logic inference的动态3D环境:** 我们根据规则算法生成了大量的包含3D物体的场景, 逻辑规则, 和特定任务的场景(平均每个场景有8条一阶逻辑规则, 30多个任务), 以及对应的任务命令  $\{l^{(t)}\}$ ; 这个环境适合被用于使用强化学习的架构来训练Agent完成任务;

传统的强化学习训练应当是将Agent代入环境一步步完成任务且学习  $\pi(a|s)$  和  $Q(a, s)$  (连续任务: 让Agent一直做下去), 但是为了方便神经网络的训练采用的方式是"离散任务", 也就是先在环境上跑一边正确的任务示范, 收集数据

$\{V^{(t)}, l^{(t)}, G_s^{(t)}, A^{(t)}, V^{(t)}\}$  然后训练Agent网络, 对应的baseline网络包含的模块:

- 用于决策的policy network  $\pi(a^{(t)} | G_s, G_l, l^{(t)})$ ;
- 用于环境物体状态识别的recognition network  $R(v^{(t)}, l^{(t)})$ ;
- 用于构建逻辑关系的inference network  $I(G_s^{(t)}, G_s^{(t-1)}, v^{(t)})$ ;
- 用于评估策略的value network  $Q(A, G_s, G_l, l^{(t)})$ ;





## 下一步的研究思路/计划

**当前baseline的缺陷分析及展望:**当前的baseline虽然结合进了语义图,但是仍然是依赖深度学习的方法做梯度下降和数据集拟合,难以具备真正智能的泛化性推理能力(智能非拟合).因此问题的重点是如何构建具备one-shot learning级别的概念学习/逻辑推理智能,这样的答案我倾向于去认知心理学、神经科学、以及数学理论中寻找.

目前我正在沿着这样的思路进行探究,更细节的内容还未敢妄下断言,但是有如下宏观观点:

- 类脑的认知智能不必是很强的编码解码架构模型:各种神经科学表征规律的研究证明我们的大脑擅长编码,但是没有理论支撑大脑是类似VAE或者GAN那样强的生成解码模型,这符合直觉,因为我们的大脑擅长认知推理,而非拟合某种复杂的分布;
- 反馈机制的具体形式:智能系统一定有反馈,但不一定是梯度下降;神经科学、控制科学等任何涉及智能的模型都离不开反馈机制,涉及反馈则又一定涉及代价函数,代价函数和反馈机制的具体形式值得深究!!
- 数学理论蕴含了一些对学习机制有价值的理论待发掘:由于学习理论离不开embedding,也就是需要将客观世界映射到表征空间,而诸如泛函分析等学科中的理论就有助于对映射进行严格讨论(比如机器学习中的SVM的核函数的定义式基于泛函分析中的希尔伯特空间).但是需要避免的另一个缺陷就是为了使用数学理论而去"套用"数学,这也本末倒置.

