

# AIDemo



# 模仿你的行为

李蓝天

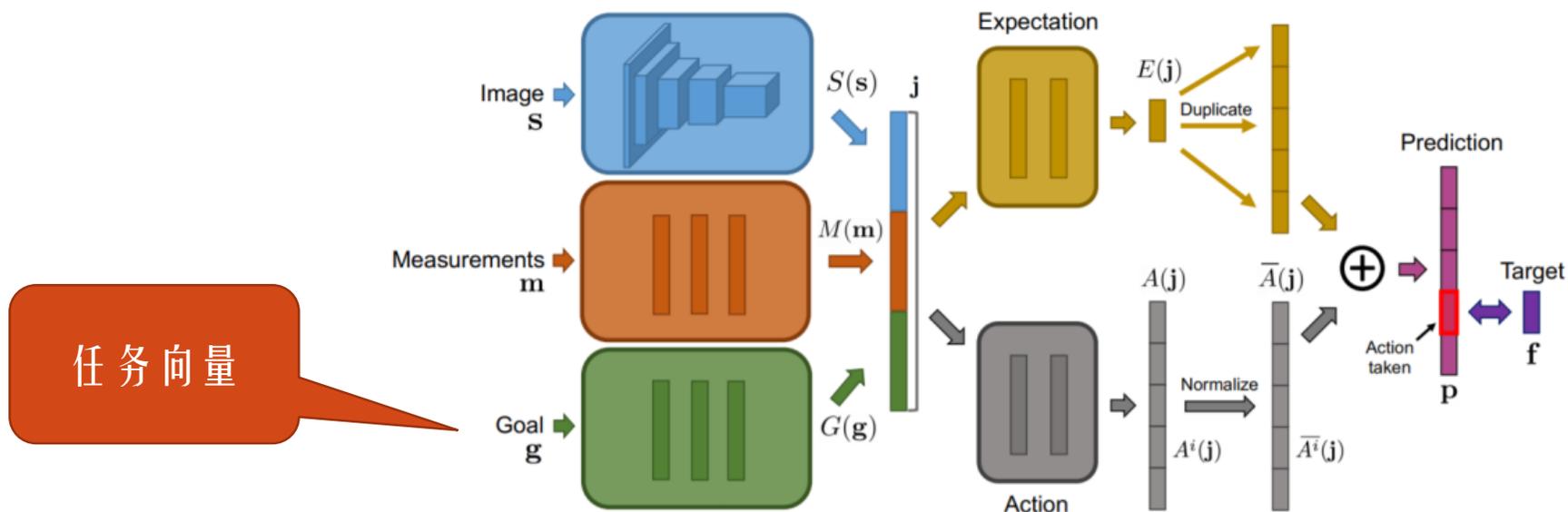
# 强化学习

强化学习是机器主动和环境进行交互，通过环境给出的反馈（奖励或惩罚）不断调整学习行为策略，直至得到最优回报。



# 实验：会打DOOM游戏的AI机器人

- 训练目标：基于当前状态预测某一动作的未来收益（DFP模型）
- 任务向量：弹药、生命和杀怪数，三者比重将定义动作收益，影响机器的行为方式



# 实验步骤

- 认真阅读 `robot/fps-game/doc/README`，了解实验步骤。
  - 可在图形界面下双击打开，也可在终端下用 `vim` 打开
  - 关于 `vim` 的操作，请 [aibook.csl.t.org](http://aibook.csl.t.org) 下 Linux Shell 简易教程
- 打开终端，进入 `fps-game` 的 `code` 目录。
- 运行缺省程序

**`sh run.sh`**

- 运行上述命令即可显示机器人自动打怪场景。

# 运行界面

- 进入到 fps-game 的子文件夹 code 中运行 run.sh

```
tutorial@localhost:~/aibook/demo/robot/fps-game/code
文件(F) 编辑(E) 查看(V) 搜索(S) 终端(T) 帮助(H)
[tutorial@localhost ~]$ ls
aibook 公共 模板 视频 图片 文档 下载 音乐 桌面
[tutorial@localhost ~]$ cd aibook/demo/robot/fps-game/code/
[tutorial@localhost code]$ ls
DFP example maps run.sh ViZDoom
[tutorial@localhost code]$ sh run.sh
```

A terminal window showing the steps to run a script. The terminal title is 'tutorial@localhost:~/aibook/demo/robot/fps-game/code'. The menu bar includes '文件(F)', '编辑(E)', '查看(V)', '搜索(S)', '终端(T)', and '帮助(H)'. The user enters 'ls' and sees a list of files including 'aibook', '公共', '模板', '视频', '图片', '文档', '下载', '音乐', and '桌面'. Then, they enter 'cd aibook/demo/robot/fps-game/code/'. Next, they enter 'ls' and see 'DFP', 'example', 'maps', 'run.sh', and 'ViZDoom'. Finally, they enter 'sh run.sh'. Two orange callout boxes with white text are present: one labeled '路径' (Path) pointing to the 'cd' command, and another labeled '运行' (Run) pointing to the 'sh run.sh' command.

路径

运行

# 游戏结果

- 弹药 ( ammo ) : 36.87
- 健康 ( health ) : 96.43
- 打怪 ( frags ) : 21.02
- 平均得分 : 38.0

A screenshot of a terminal window. The title bar reads "tutorial@localhost:~". The menu bar contains "文件(F)", "编辑(E)", "查看(V)", "搜索(S)", "终端(T)", and "帮助(H)". The main content area shows a shell prompt "|tutorial@localhost ~|\$" followed by a blank line.

```
tutorial@localhost:~  
文件(F) 编辑(E) 查看(V) 搜索(S) 终端(T) 帮助(H)  
|tutorial@localhost ~|$
```

# 改变任务向量

- 修改 code/examples/run\_exp.py 中的任务向量，如下图红色箭头所指。
- 任务向量中的三个数值 [0.5, 0.5, 1.0] 分别代表三个目标：弹药、生命和杀怪数，设定值在 -1 到 1 之间。

```
tutorial@localhost:~/aibook/demo/robot/fps-game/code
文件(F) 编辑(E) 查看(V) 搜索(S) 终端(T) 帮助(H)
110     experiment_args['test_objective_coeffs_temporal'] = np.array([0., 0., 0., 0.5, 0.5, 1.])
111     experiment_args['test_random_prob'] = 0.
112     experiment_args['test_checkpoint'] = 'checkpoints/2017_04_09_09_13_17'
113     experiment_args['test_policy_num_steps'] = 500
114     experiment_args['show_predictions'] = False
115     experiment_args['multiplayer'] = False
116
117     # ===== adjust this parameter to achieve the different goal =====
118     # three goal are ammo health frags respectively, and the larger the value is, the better the goal can be achieved
119     # for example, when setting [1.0 1.0 -1.0], the AI agent will cherish its bullets, increase its life and do not kill the monsters
120     # Note Note Note !!! that the value can not be zero, otherwise there will be error.
121     experiment_args['test_objective_coeffs_meas'] = np.array([0.5, 0.5, 1.0])
122     # experiment_args['test_objective_coeffs_meas'] = np.array([0.01, 0.01, 1.0])
123     # experiment_args['test_objective_coeffs_meas'] = np.array([1.0, 1.0, -1.0])
124     # experiment_args['test_objective_coeffs_meas'] = np.array([-1.0, 0.01, 0.01])
125     # experiment_args['test_objective_coeffs_meas'] = np.array([0.01, 1.0, 0.01])
126
```

# 改变任务向量

- 任务向量中的某一数值设定值越大，对应目标将被实现的越好。
- 例如，将三个参数设定为  $[1.0, 1.0, -1.0]$  时，机器人将会更加珍惜子弹，维持生命，但尽量不击杀怪物。
  
- 请学生自行修改任务向量，并重新执行 `run.sh`。
- 观察改变任务向量对游戏结果产生的影响。

**The end !**