# Sparse Reward

Hung-yi Lee

# Sparse Reward

## Reward Shaping

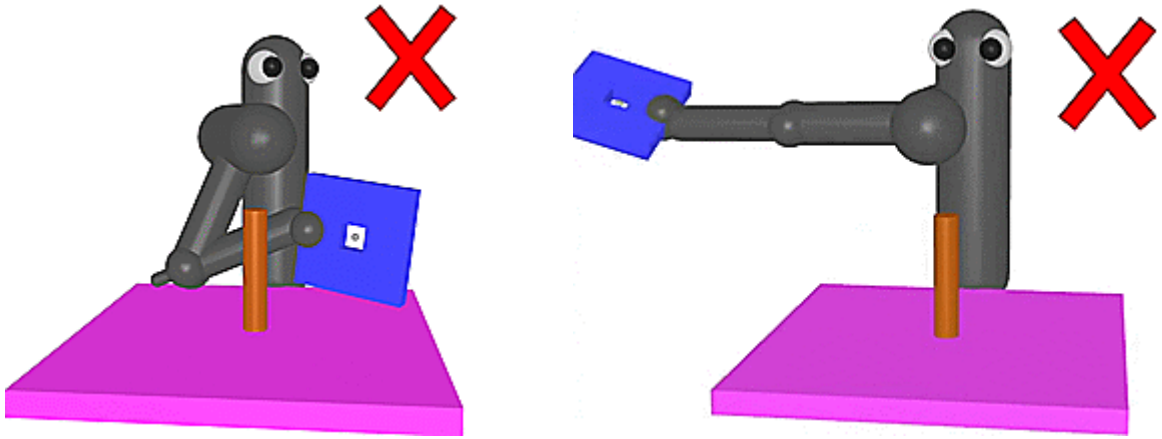
# Reward Shaping

為machine 示範action，引導學習

（更改真實的reward，引導machine做到我們希望他做的事情）

# Reward assigned by Developer

| Parameters | Description 讓他好戰一點ＸＤ | FlatMap \| CIGTrack1 |
|---|---|---|
| living 活著扣分 | Penalize agent who just lives | -0.008 / action |
| health_loss 掉血就扣分 | Penalize health decrement | -0.05 / unit |
| ammo_loss 減少彈藥扣分 | Penalize ammunition decrement | -0.04 / unit |
| health_pickup 補給包加分 | Reward for medkit pickup | 0.04 / unit |
| ammo_pickup | Reward for ammunition pickup | 0.15 / unit |
| dist_penalty idle扣分 | Penalize the agent when it stays | -0.03 / action |
| dist_reward 移動加分 | Reward the agent when it moves | 9e-5 / unit distance |



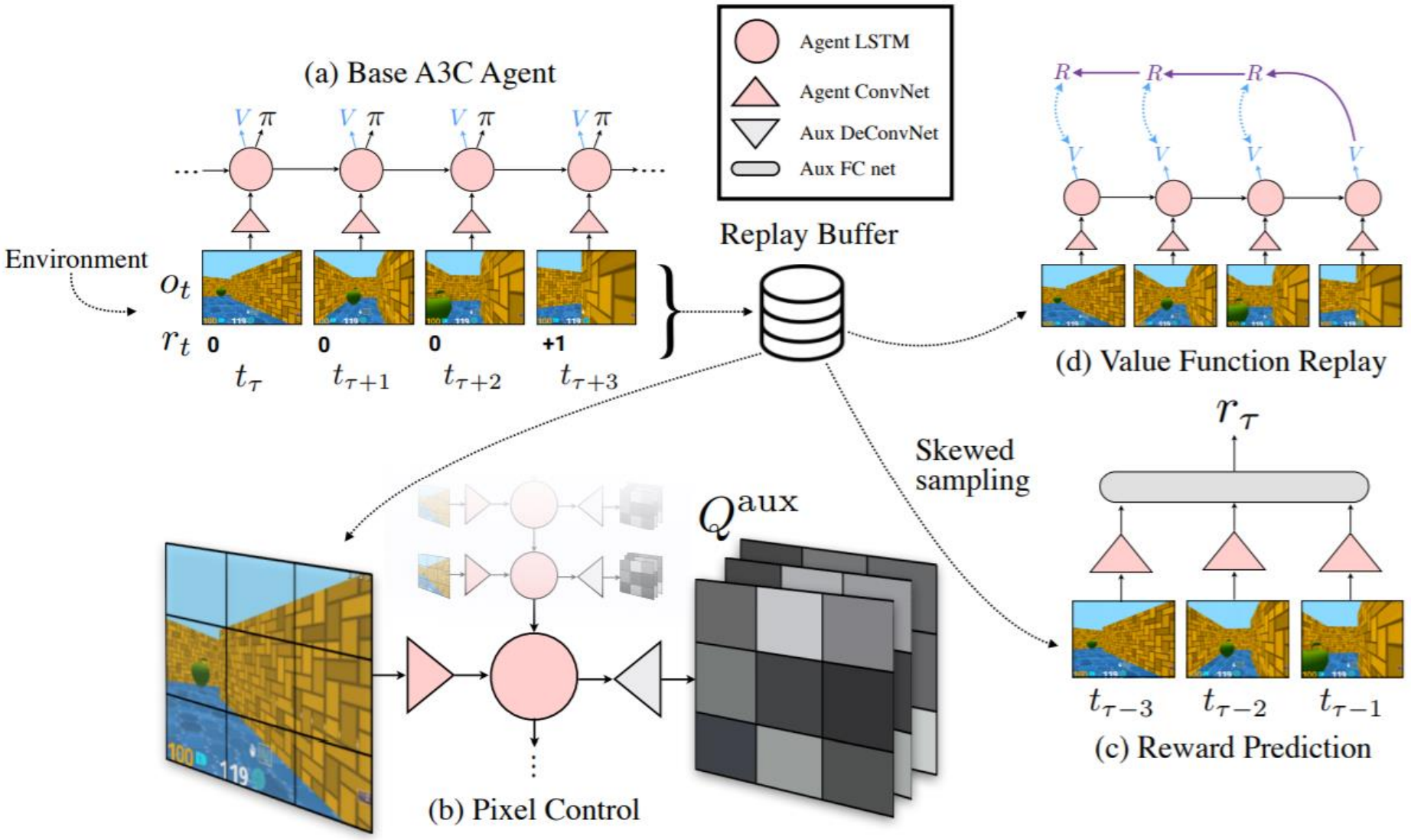
Get reward, when closer

Need domain knowledge

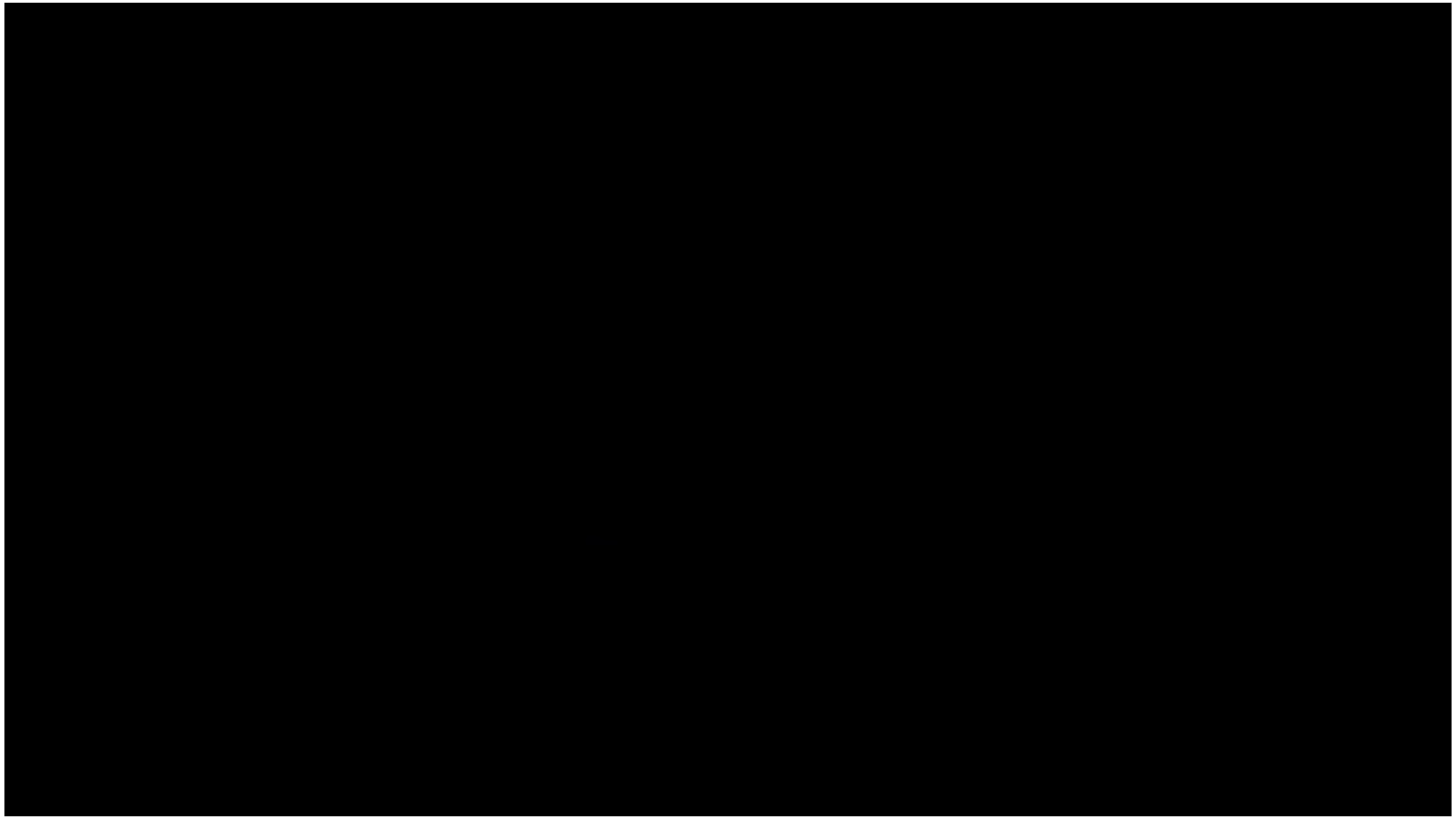
https://openreview.net/pdf?id=Hk3mPK5gg

https://arxiv.org/abs/1611.05397

# Reward from Auxiliary Task

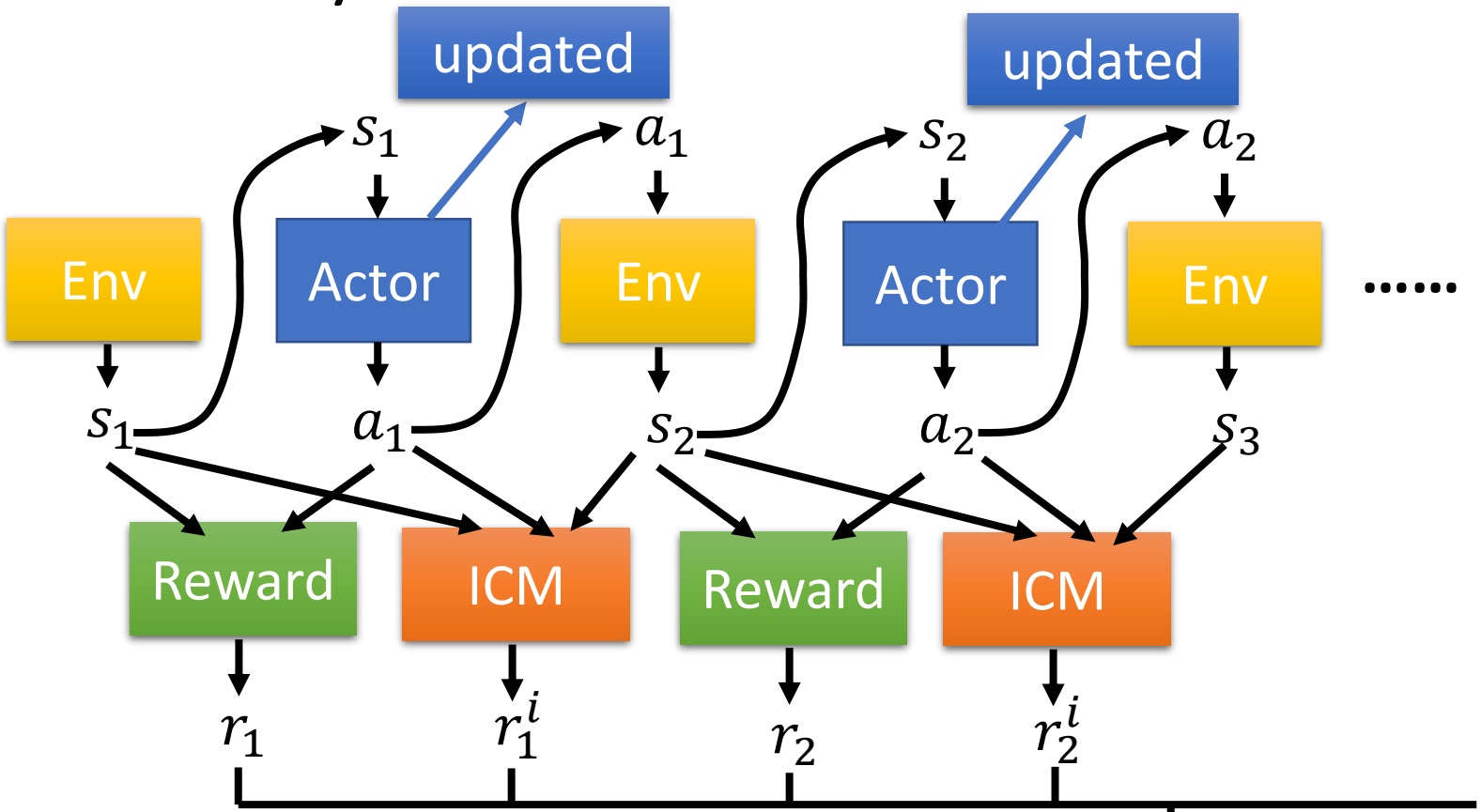


# Demo



https://arxiv.org/abs/1705.05363

# Curiosity



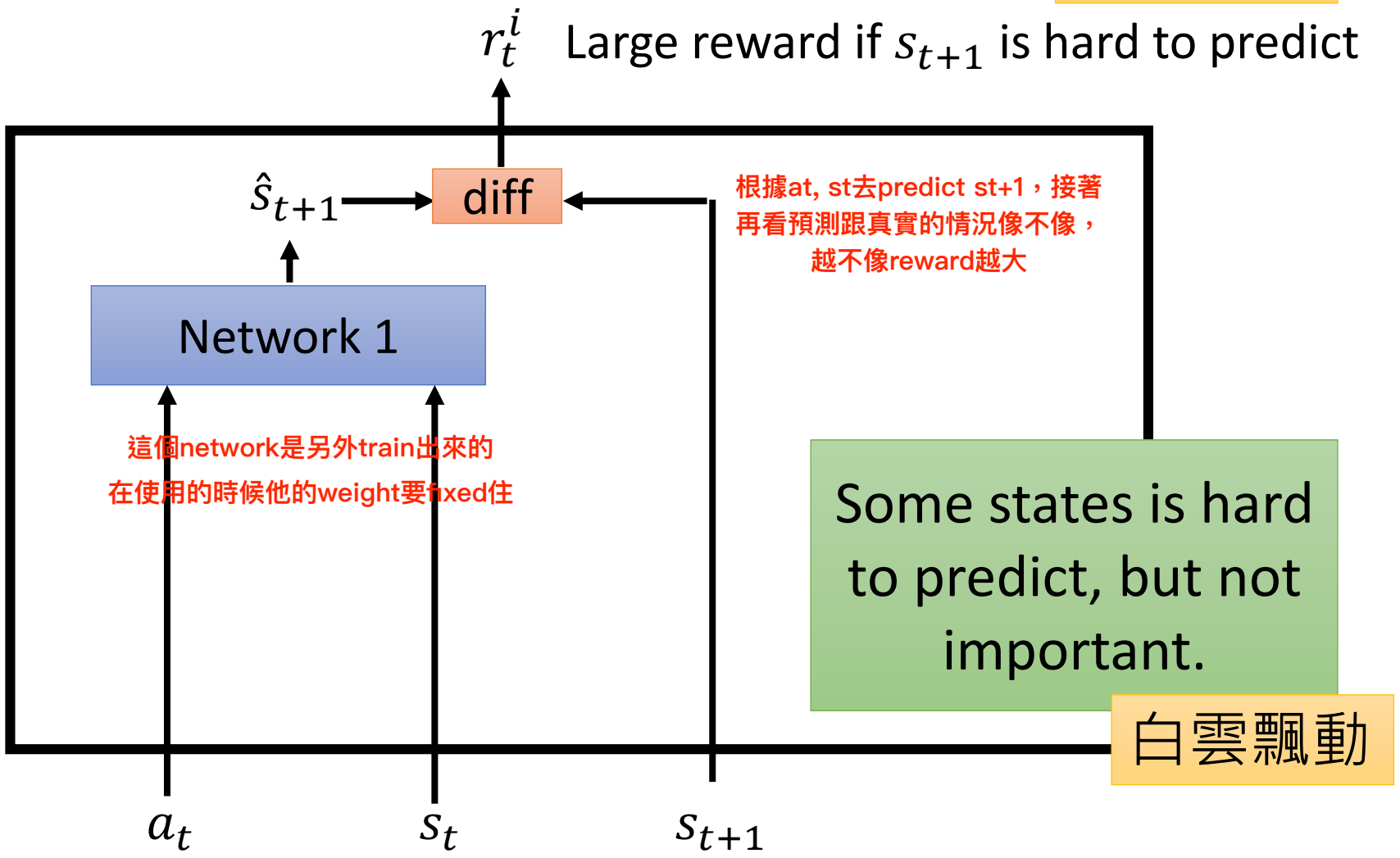

ICM = intrinsic curiosity module

好奇心

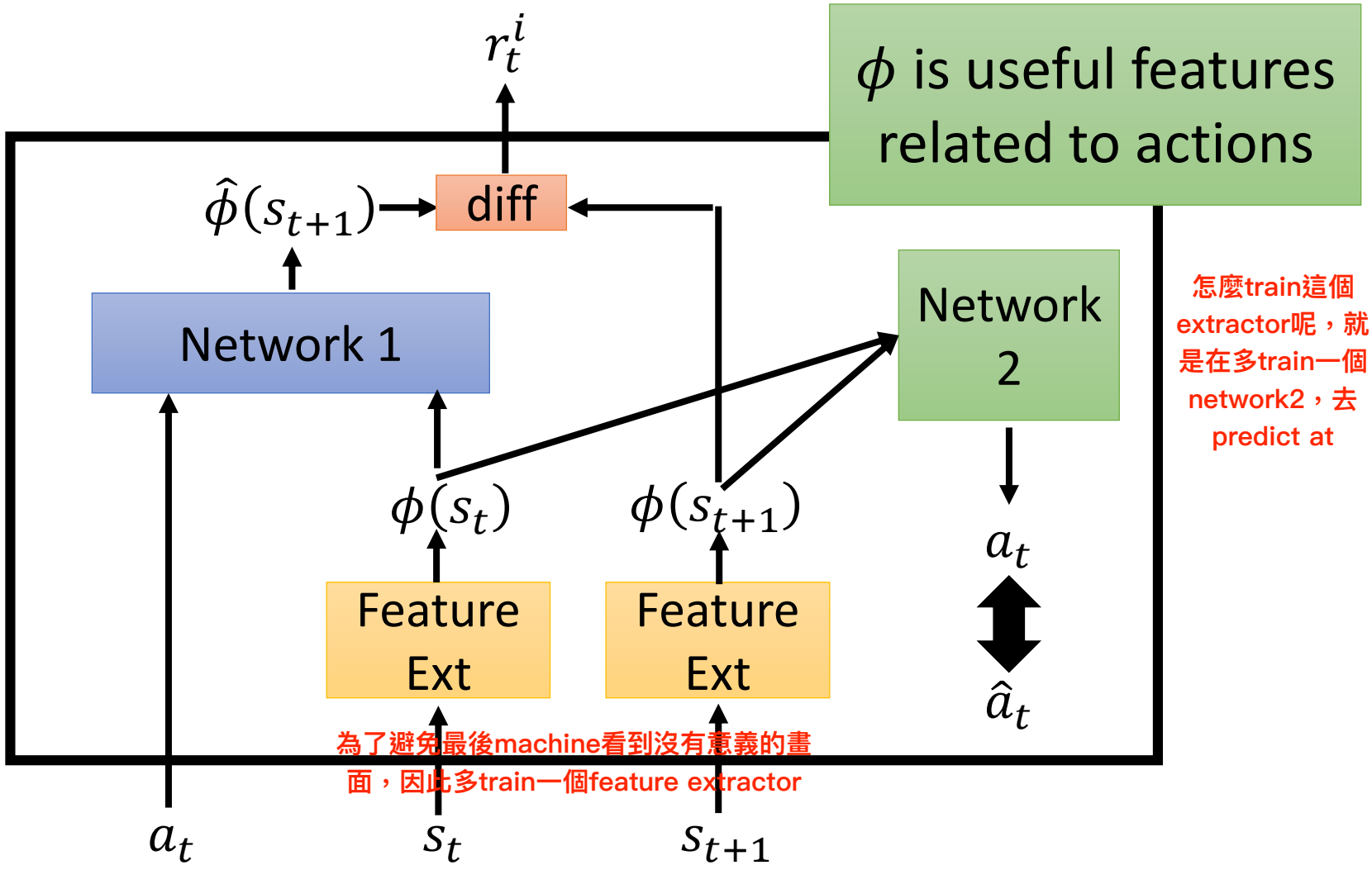
$$R(\tau) = \sum_{t=1}^{T} r_t + r_t^i$$

# Intrinsic Curiosity Module

鼓勵冒險



# Intrinsic Curiosity Module




怎麼train這個extractor呢，就是在多train一個network2，去predict at

# Sparse Reward

## Curriculum Learning

不能一次教太難
從簡單的題目交到難的題目

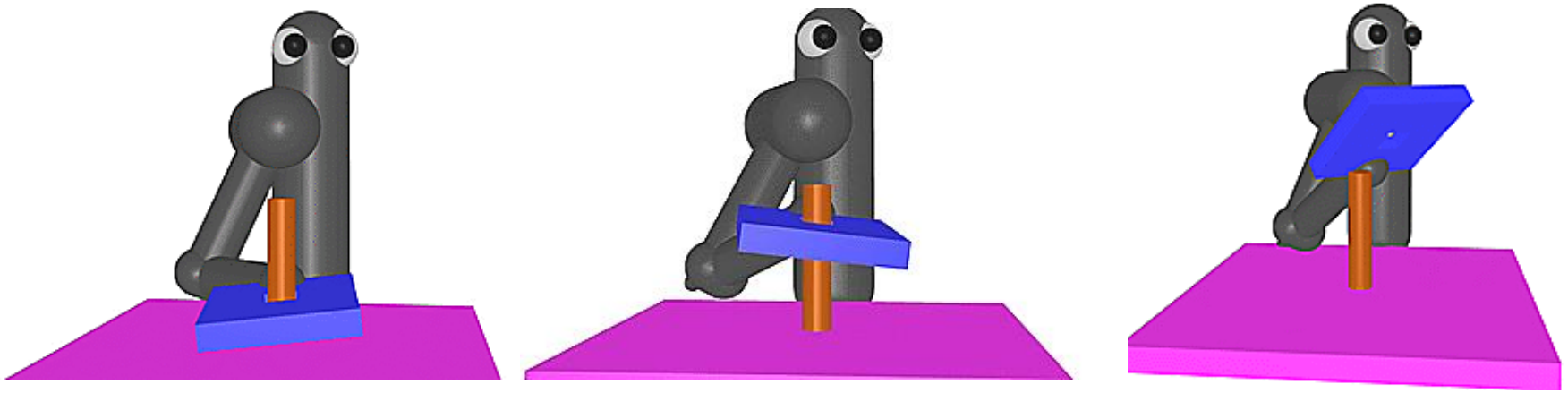
對data做處理（簡單到難）

# Curriculum Learning

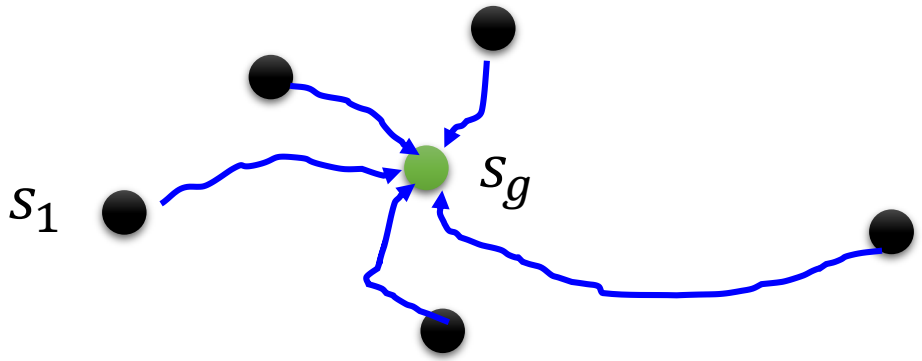
- Starting from simple training examples, and then becoming harder and harder.

幫機器規劃一下課程

| | Class 0 | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 | Class 6 | Class 7 |
|---|---|---|---|---|---|---|---|---|
| Speed | 0.2 | 0.2 | 0.4 | 0.4 | 0.6 | 0.8 | 0.8 | 1.0 |
| Health | 40 | 40 | 40 | 60 | 60 | 60 | 80 | 100 |

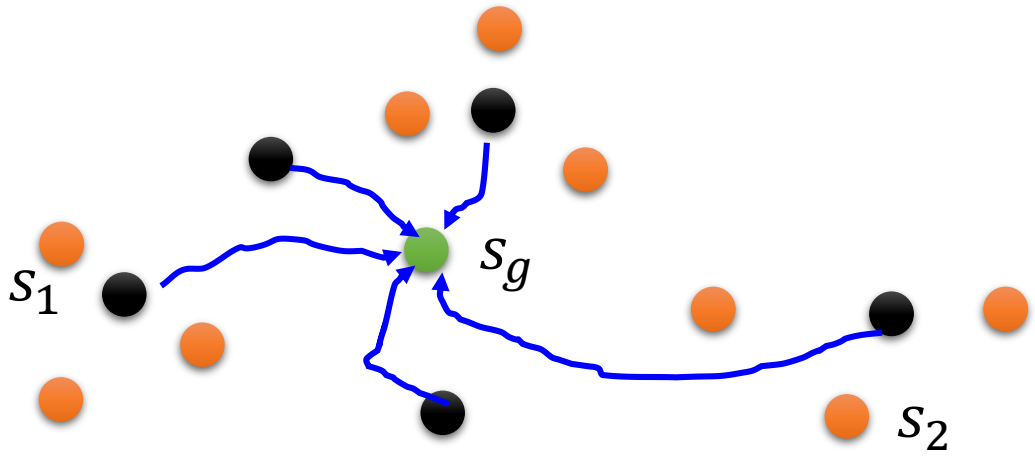


# Reverse Curriculum Generation




- Given a goal state $s_g$.
- Sample some states $s_1$ “close” to $s_g$
- Start from states $s_1$, each trajectory has reward $\mathrm{R}(s_1)$

# Reverse Curriculum Generation




- Delete $s_1$ who reward is too large (already learned) or too small (too difficult at this moment)
- Sample $s_2$ from $s_1$, start from $s_2$

# Sparse Reward

階層式的reinforcement learning

# Hierarchical Reinforcement Learning

對network做處理