

# Free Energyによる Attention Control

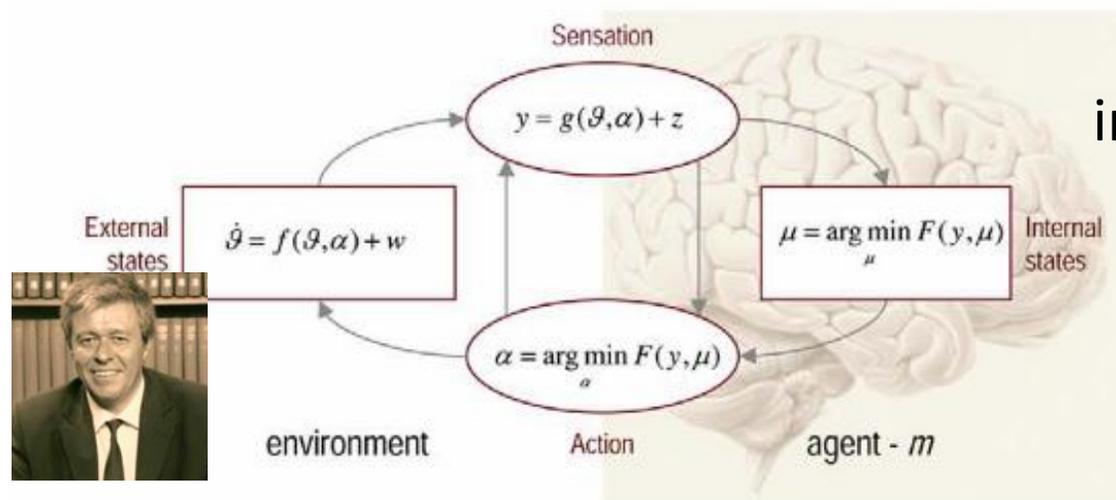
落合、都築チーム

# WBAIハッカソン報告

# 我々のやりたいことは何か？

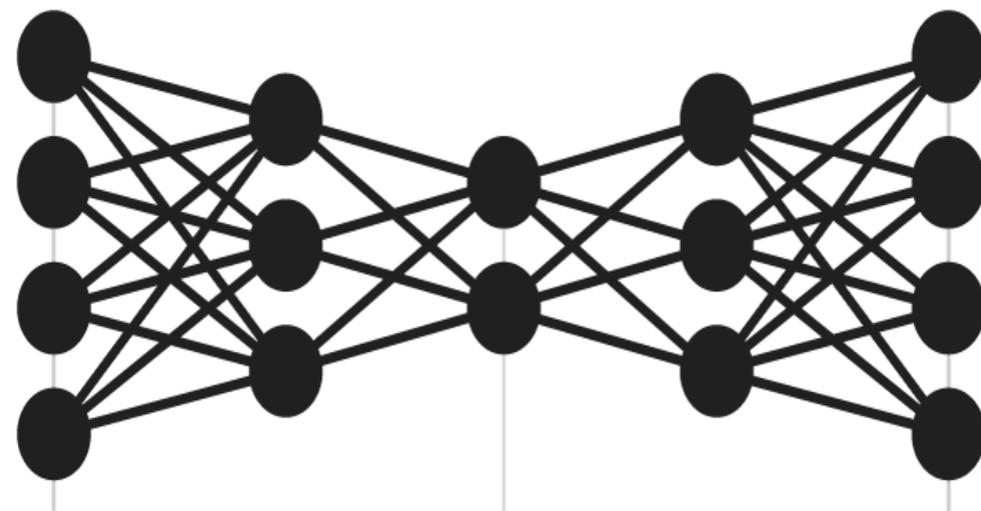
➡ Fristonの”Free Energy Principle”をANNとして実装する

## Free Energy Principle by Friston



implement

## Artificial Neural Network



$$F = \text{Energy} - \text{Entropy} = -\langle \ln p(y, \vartheta | m) \rangle_q + \langle \ln q(\vartheta) \rangle_q$$

# ”Free Energy Principle” とは何か？

- ➡ 環境中でのエージェントの学習、行動を決めるための枠組み
- ➡ エージェントは、自身のFree Energy を最小化するように内部状態と、行動を決定する

$$F(x, \mu) = \underbrace{-E_{q(z|x)}[\log p_{\theta}(x|z)]}_{\text{精度}} + \underbrace{D_{KL}[q(z|x)||p_{\theta}(z)]}_{\text{複雑度}}$$

# Internal state

$$\mu = \arg \min_{\mu} F(x, \mu)$$

予測誤差を最小化するよう  
内部パラメータを学習  
(通常の生成モデル学習)

## Internal state

$$\mu = \arg \min_{\mu} F(x, \mu)$$

予測誤差を最小化するよう  
内部パラメータを学習  
(通常の生成モデル学習)

## Action(Attention)

$$a = \arg \min_a F(x, \mu)$$

予測誤差を最小化する行動  
を生成  
予測＝注目  
(Active Inference)

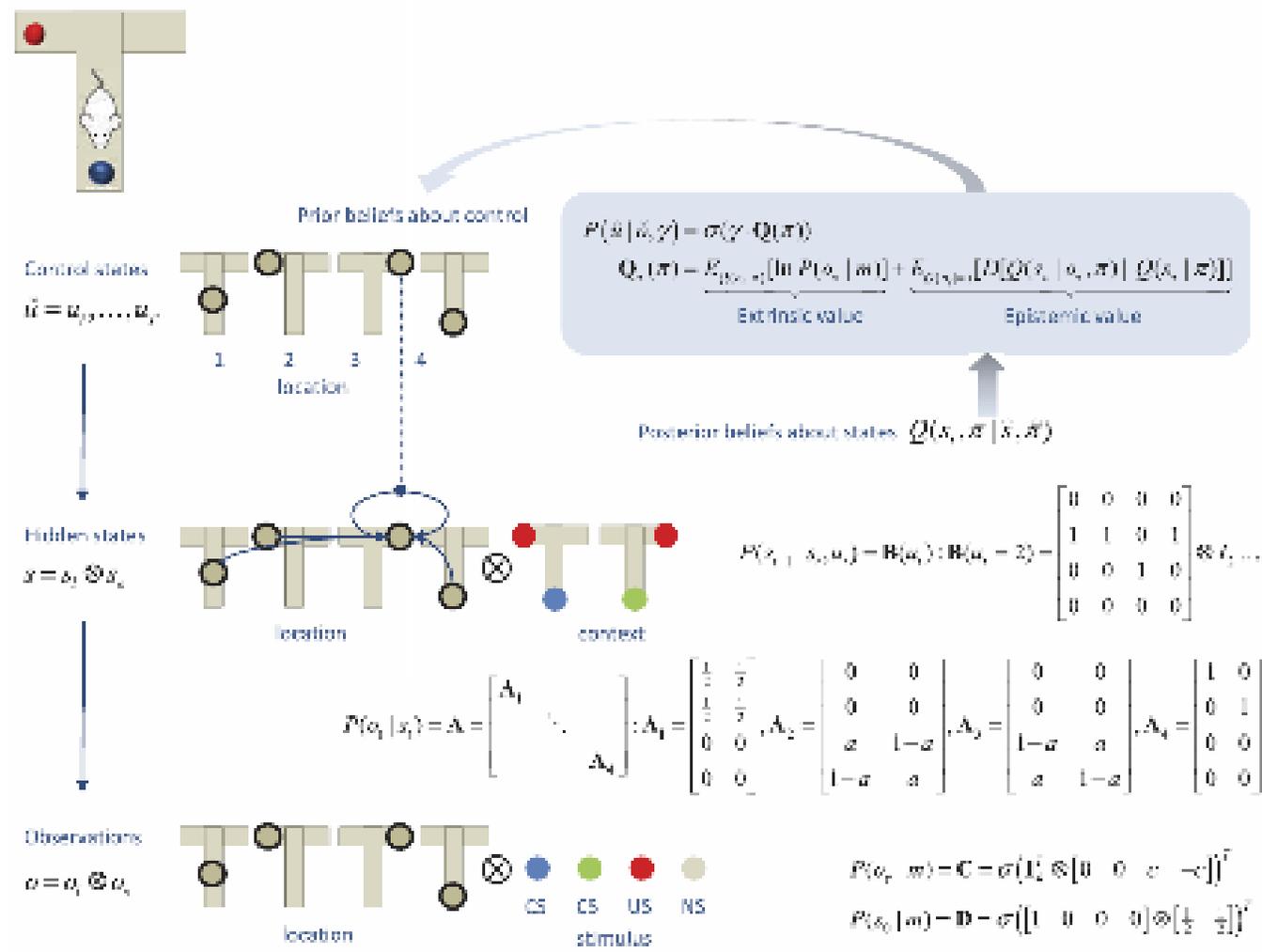
# “Free Energy Principle” の何が嬉しいか？

強化学習と比べて

- ➡ 恣意的な報酬の作り込みがない
- ➡ 広い範囲の現象に適用できる。

# Fristonによる実装

➡ 行列計算(線形モデル)による解析的な実験



# Fristonによる実装の問題点と改良案

単純な問題(Toy problem)にしか適用できない.



ANNで実装

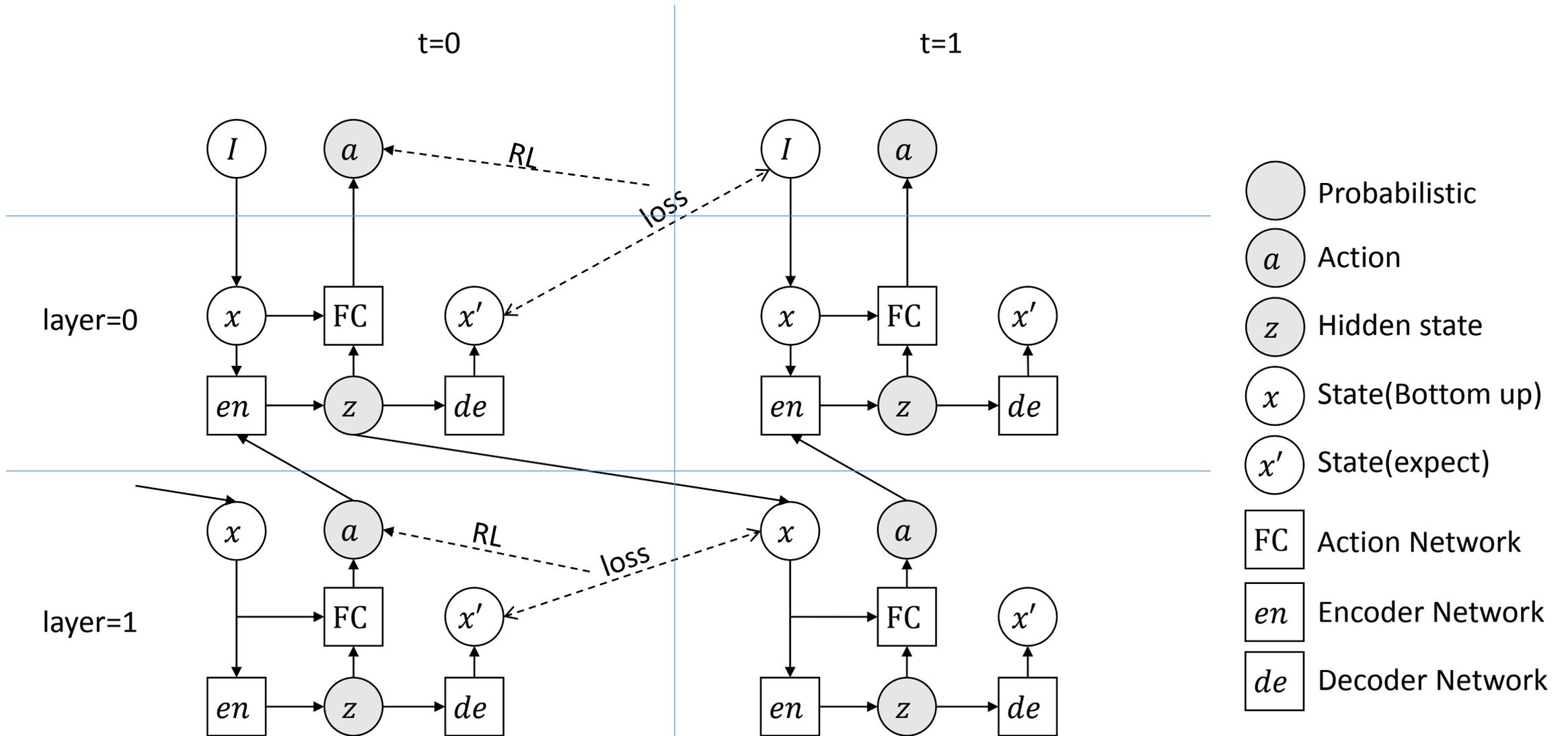
大規模かつ複雑な問題に適用できる

# 実装方針

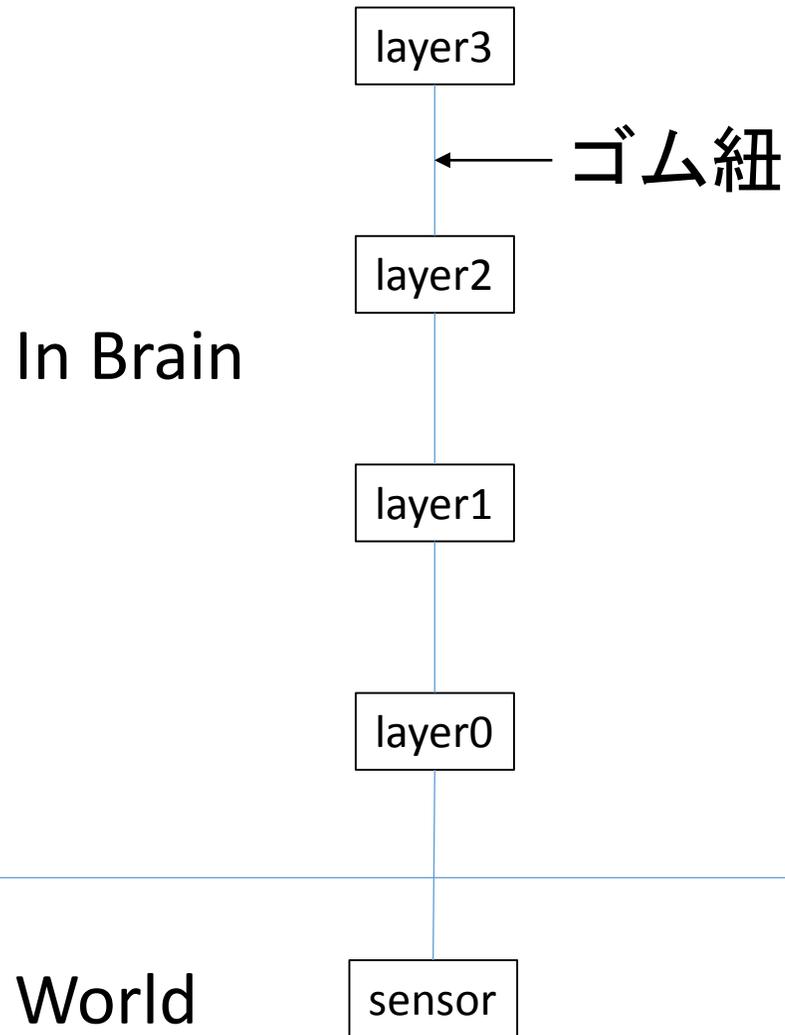
## Free Energy Principle = Variational Autoencoder

- Variational Autoencoder(VAE)は変分ベイズをニューラルネットワークで表現したもの
- 変分ベイズではFree Energyを最小化している
- Free Energyを内部パラメータ調節で最小化するNetwork = Variational Autoencoder
- **ハッカソン: Active InferenceをVAEに追加**
  - **Active Inference=Free Energyを最小化する動きを生成すること**

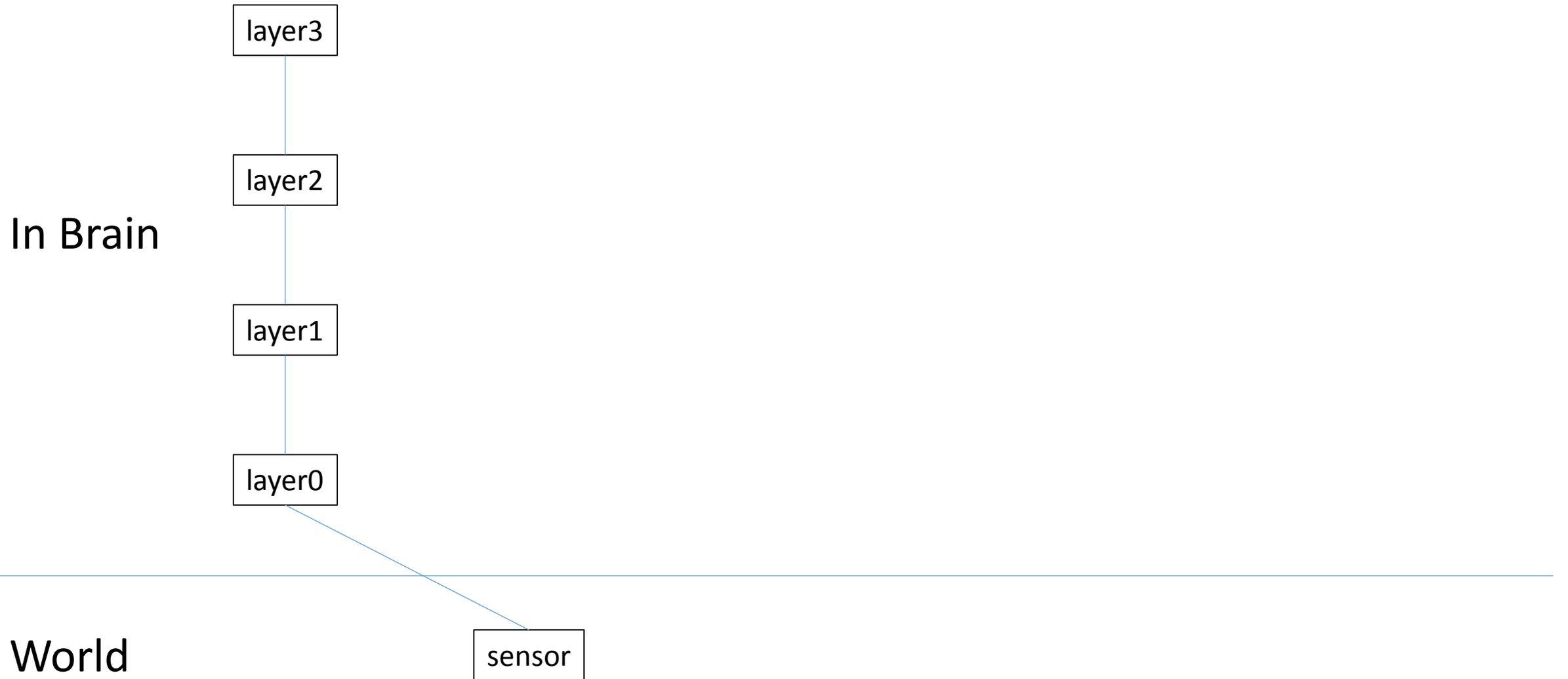
# Active Inference Network



# Active Inference Networkのイメージ



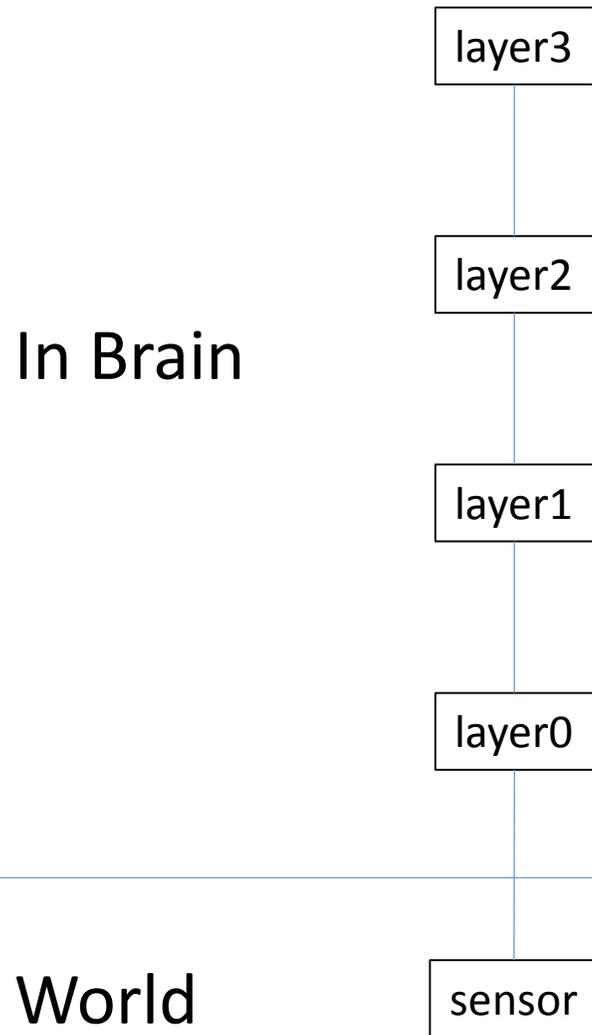
# Active Inference Networkのイメージ



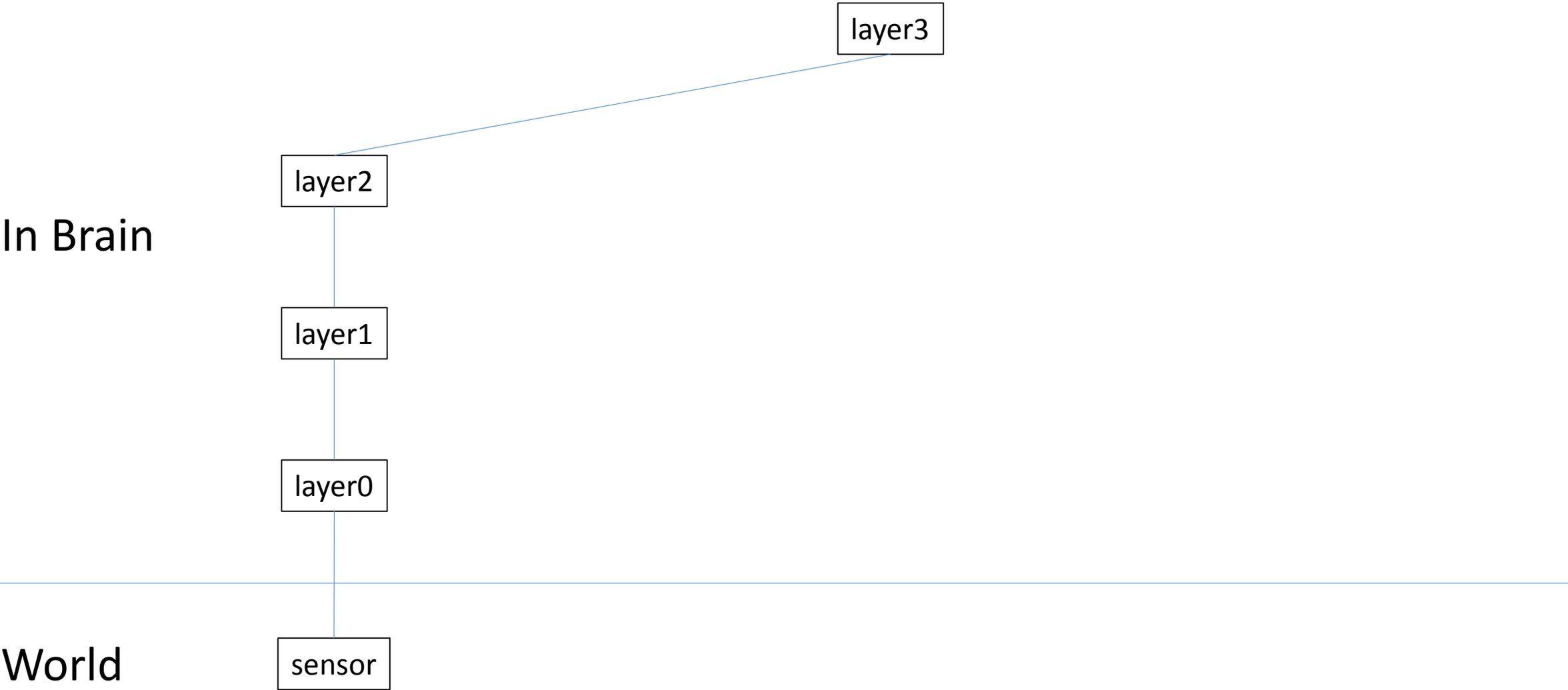
# Active Inference Networkのイメージ



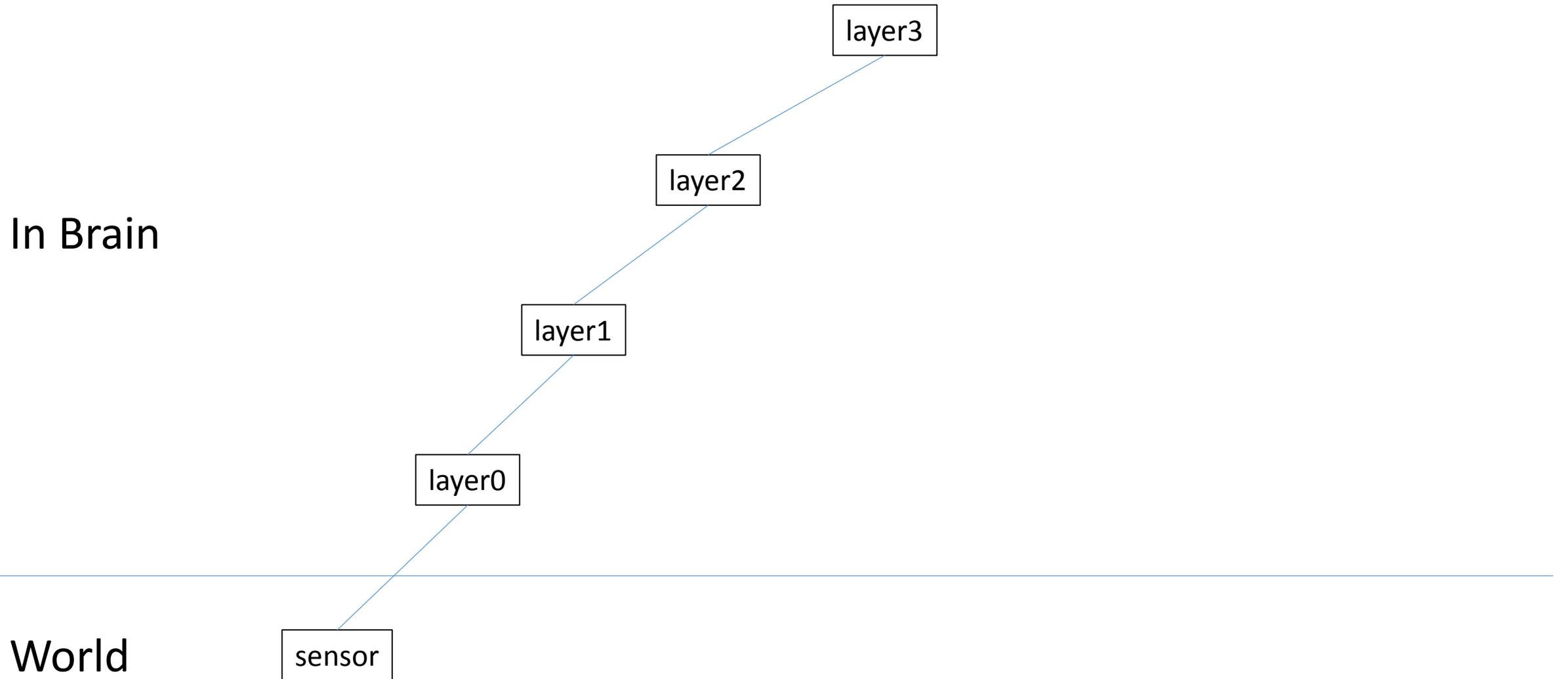
# Active Inference Networkのイメージ



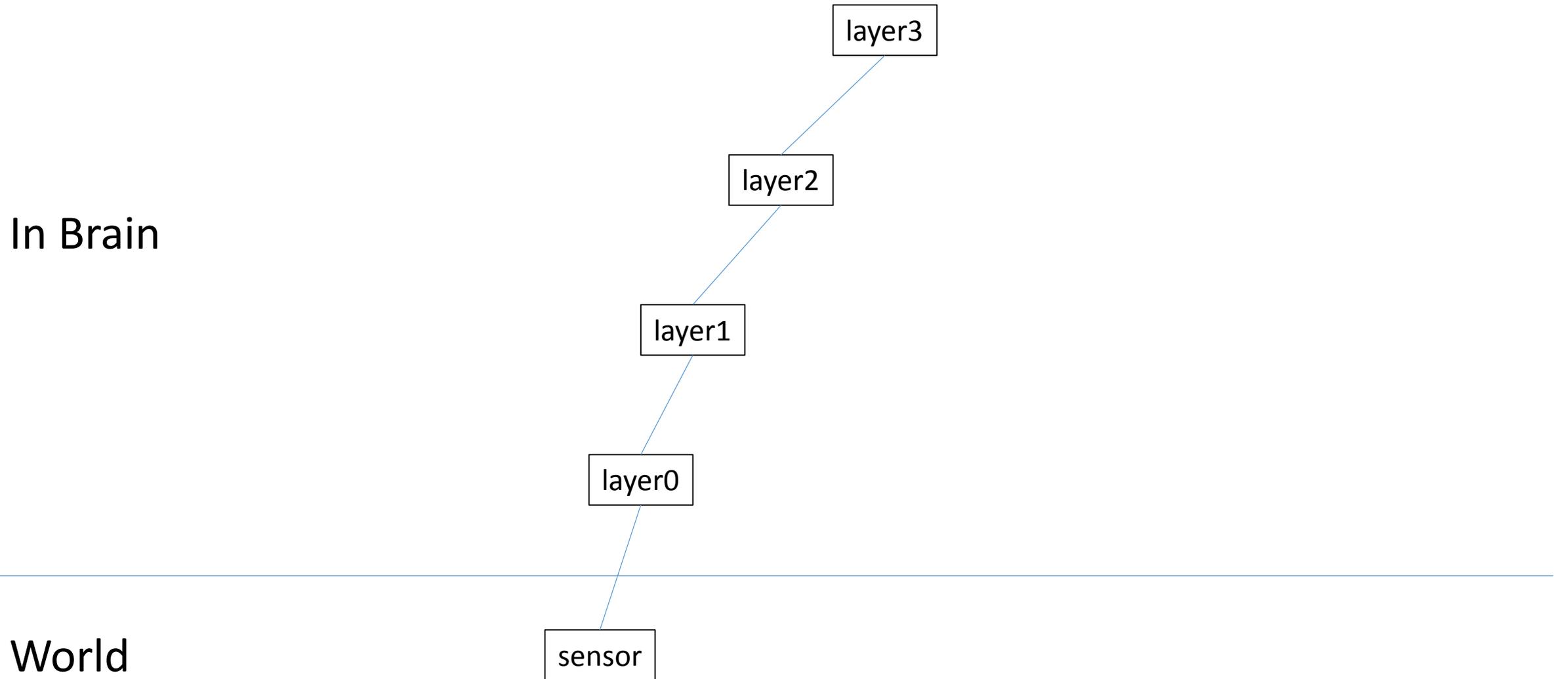
# Active Inference Networkのイメージ



# Active Inference Networkのイメージ



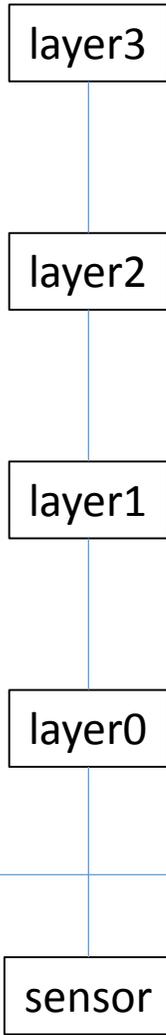
# Active Inference Networkのイメージ



# Active Inference Networkのイメージ

In Brain

World



# Active Inference Network

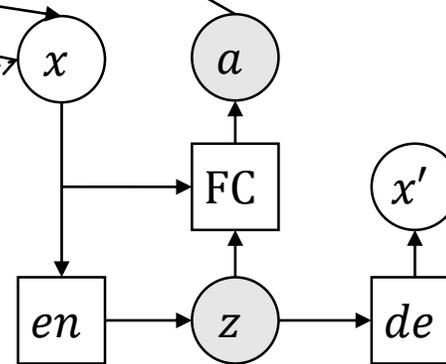
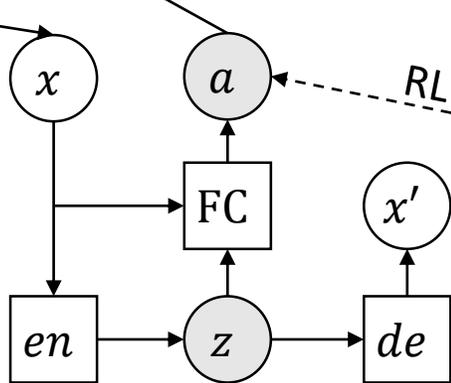
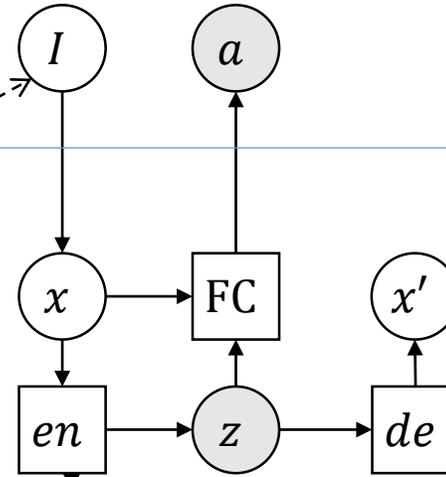
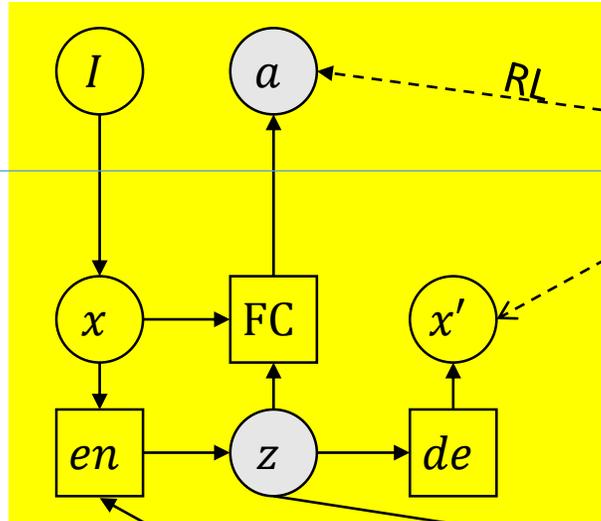
ハッカソン作成範囲

t=0

t=1

layer=0

layer=1



- Probabilistic
- $a$  Action
- $z$  Hidden state
- $x$  State(Bottom up)
- $x'$  State(expect)
- FC Action Network
- $en$  Encoder Network
- $de$  Decoder Network

# Active Inference Network

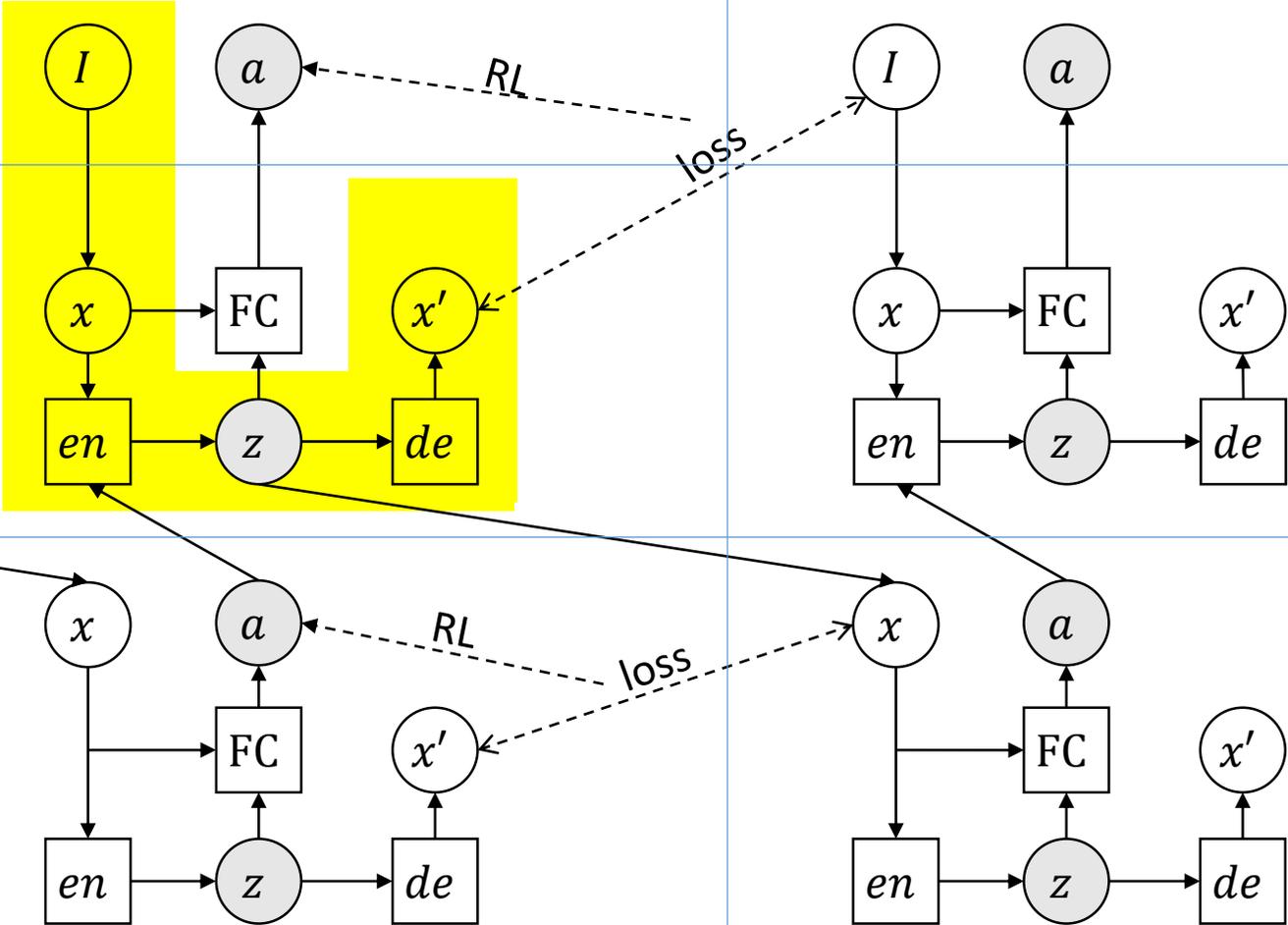
## Variational Autoencoder

t=0

t=1

layer=0

layer=1



- Probabilistic
- $a$  Action
- $z$  Hidden state
- $x$  State(Bottom up)
- $x'$  State(expect)
- FC Action Network
- $en$  Encoder Network
- $de$  Decoder Network

# Active Inference Network

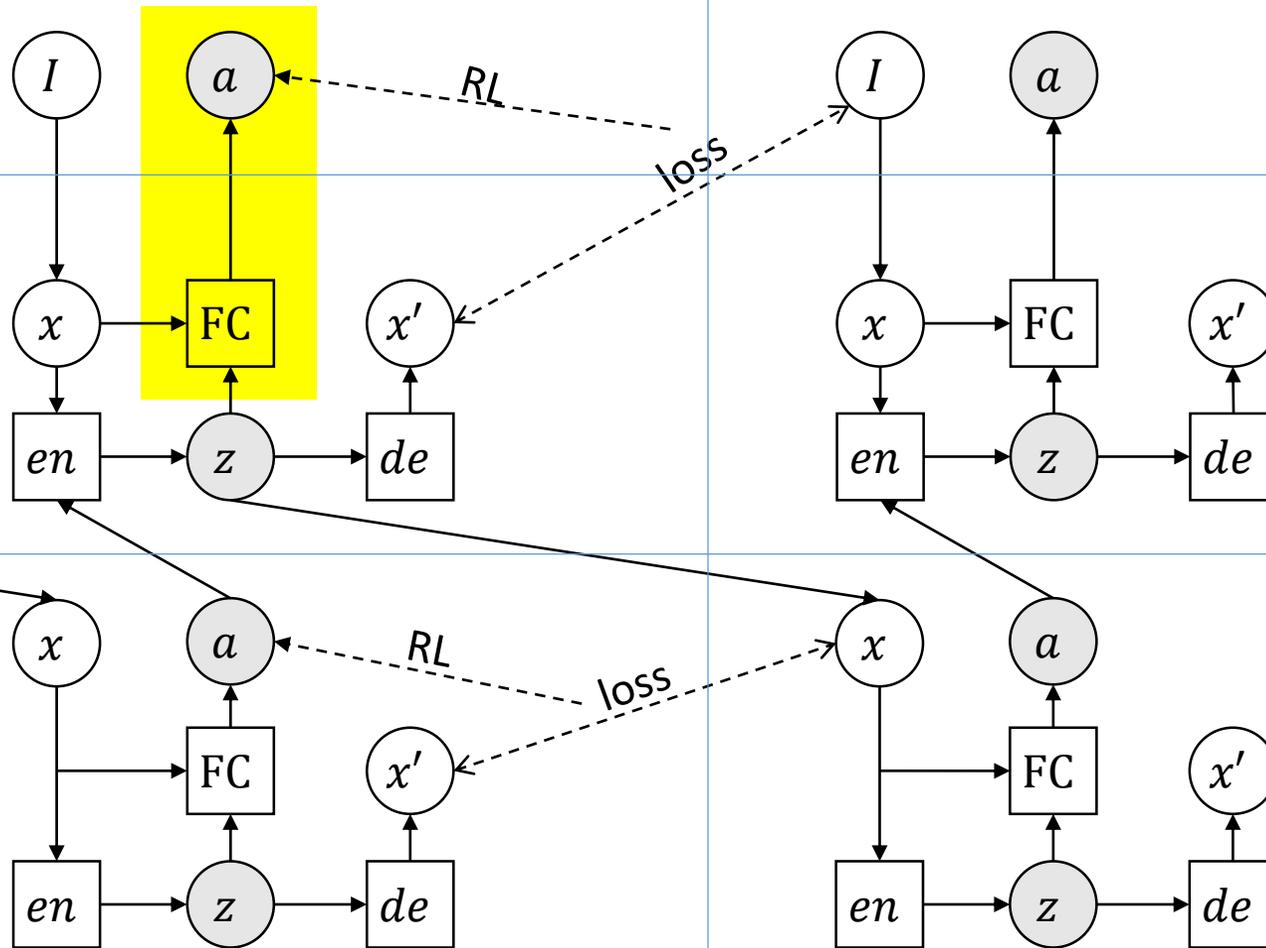
Action

t=0

t=1

layer=0

layer=1



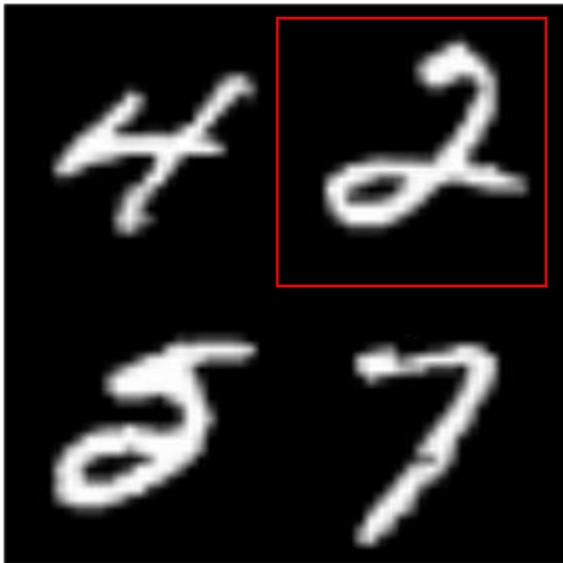
- Probabilistic
- Action**
- Hidden state
- State(Bottom up)
- State(expect)
- Action Network**
- Encoder Network
- Decoder Network

# WBAIハッカソンで何をやったか？

➡ FEPを用いて、視線移動課題を解くエージェントを実装

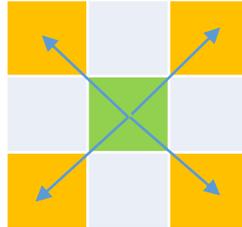
環境

着目領域



MNIST画像を並べた空間  
(トーラス状に回り込み)

ピクセル単位の移動



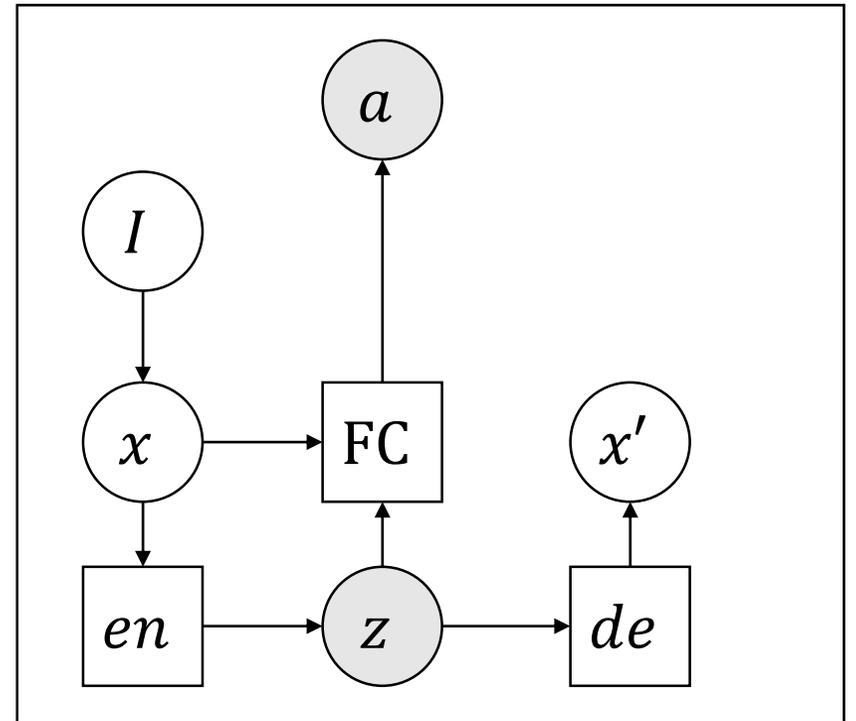
$dx = -1 \text{ or } 1, dy = -1 \text{ or } 1$

移動

画像



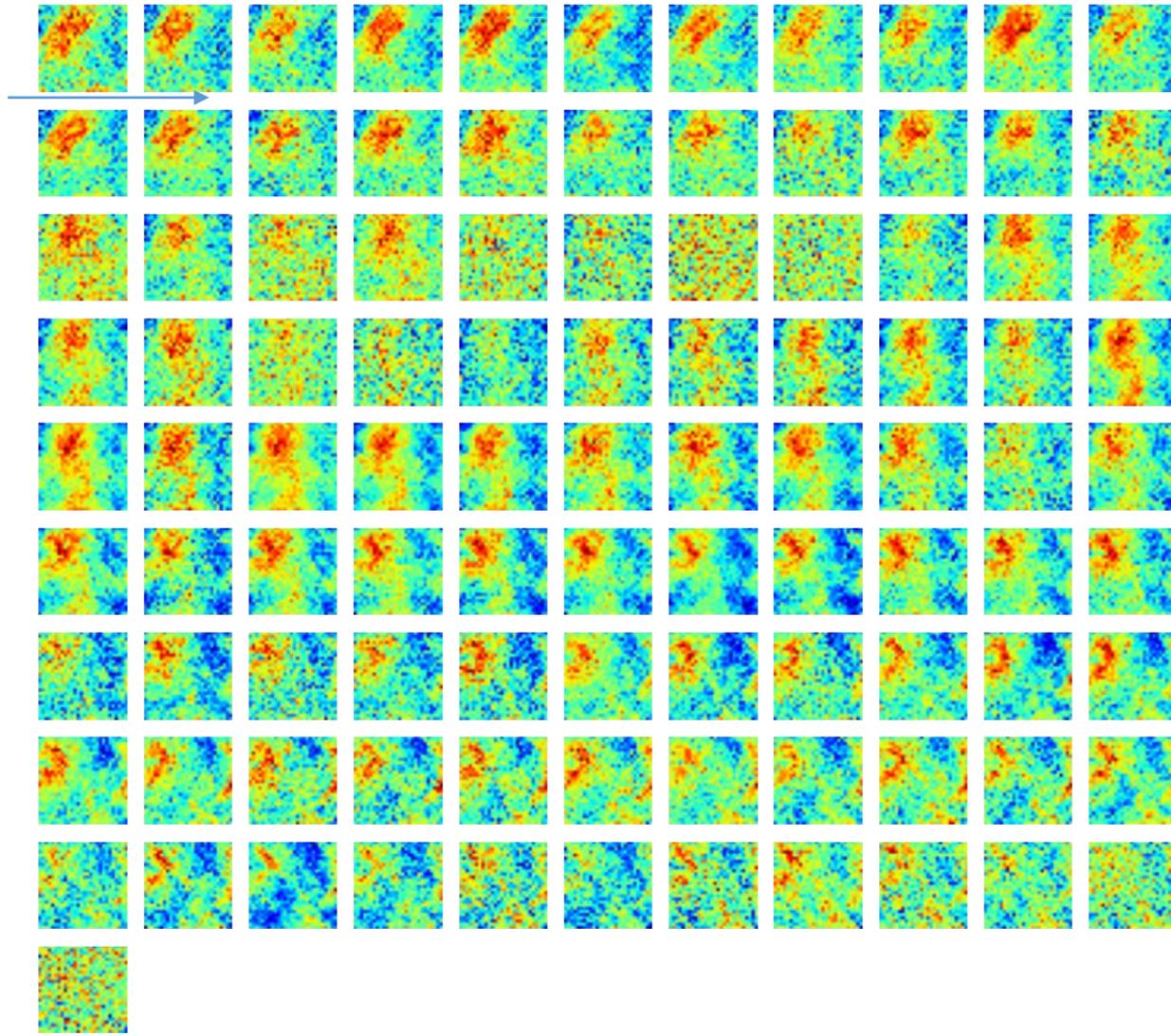
エージェント



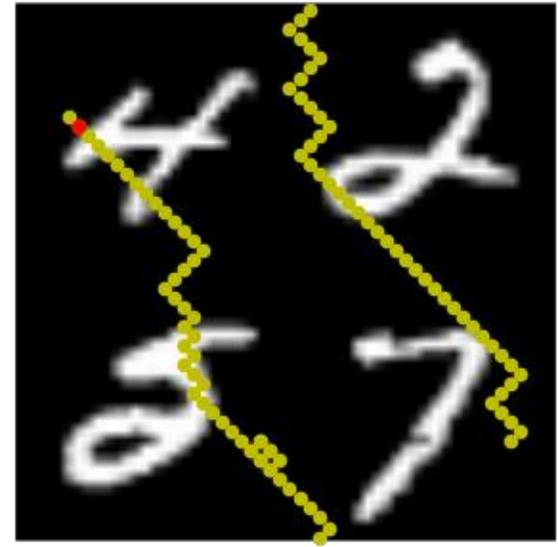
VAE+行動生成器

Step 0 ~ 100

予測画像



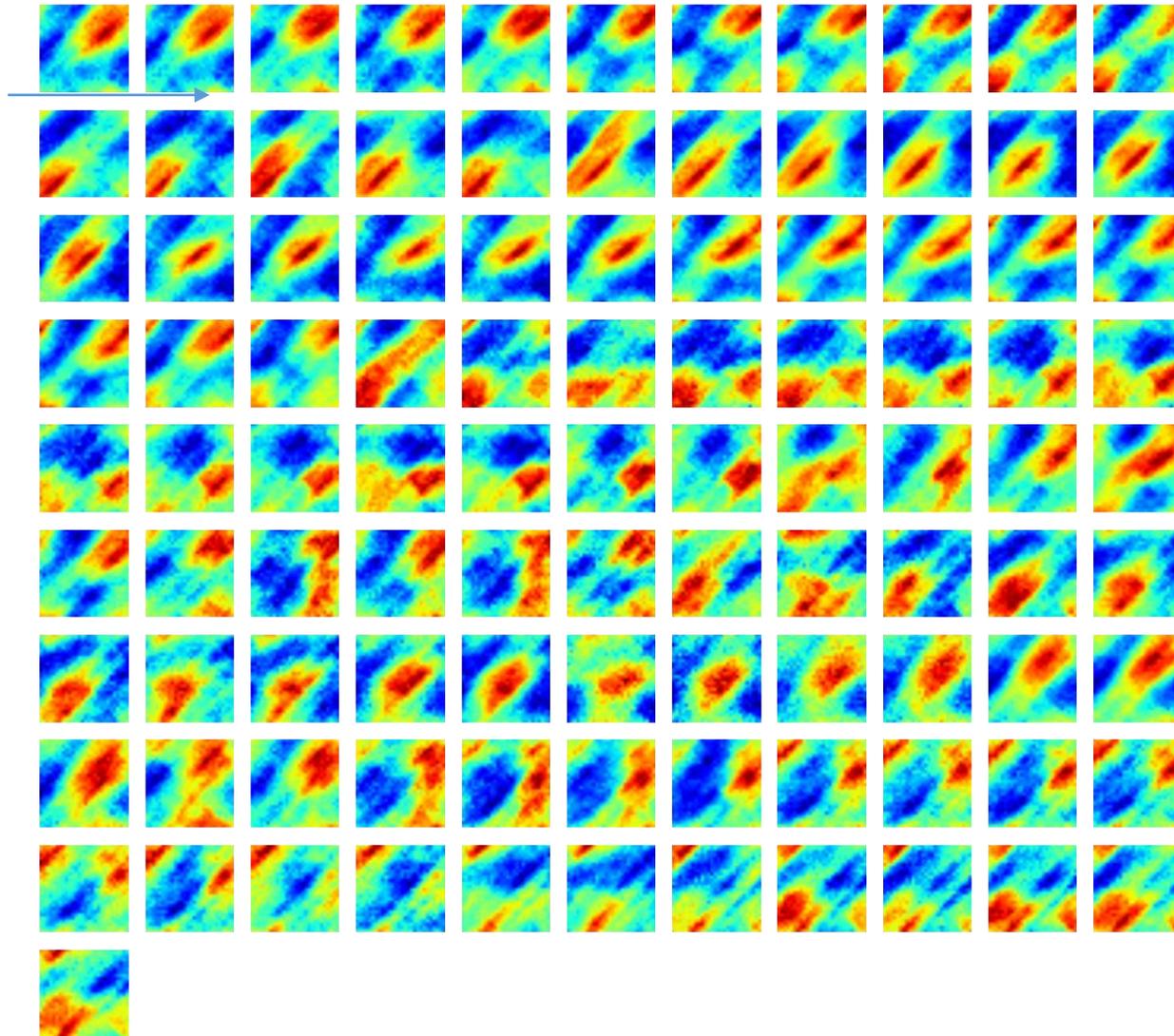
軌跡



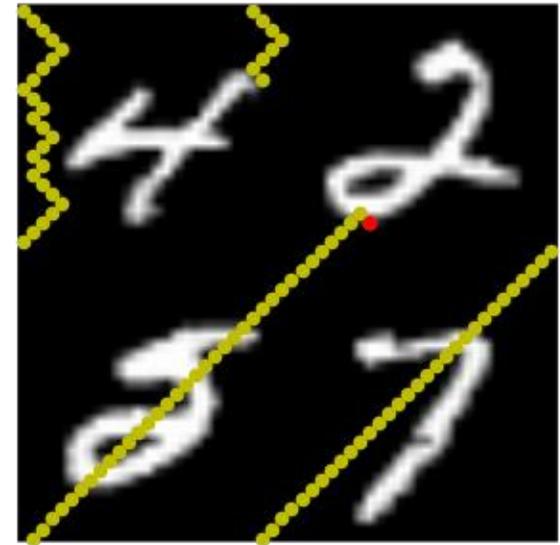
● スタート

Step 4000 ~ 4100

予測画像



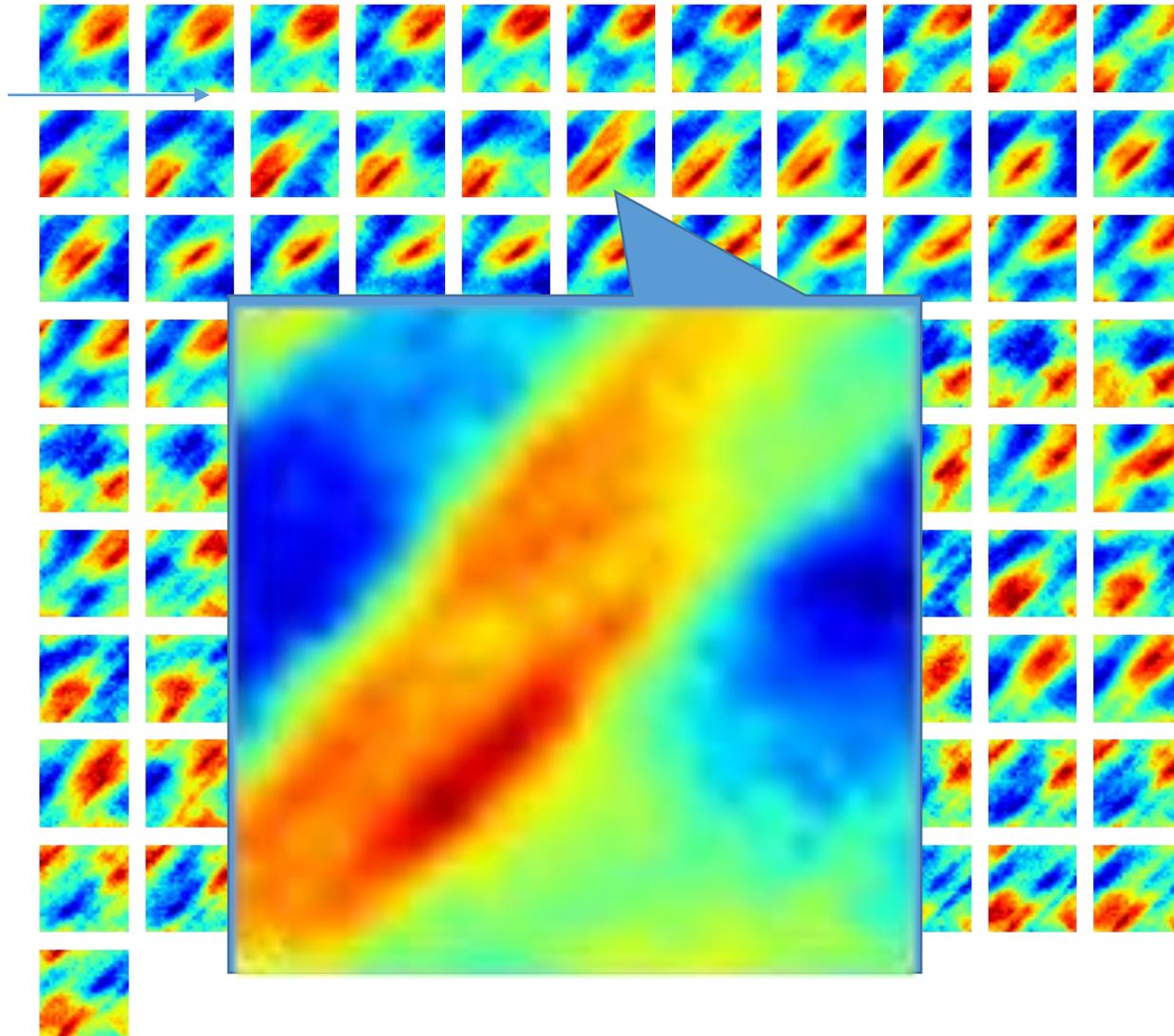
軌跡



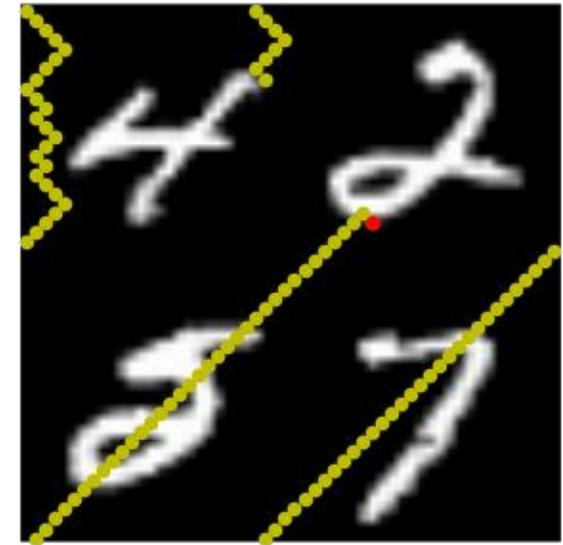
● スタート

Step 4000 ~ 4100

予測画像



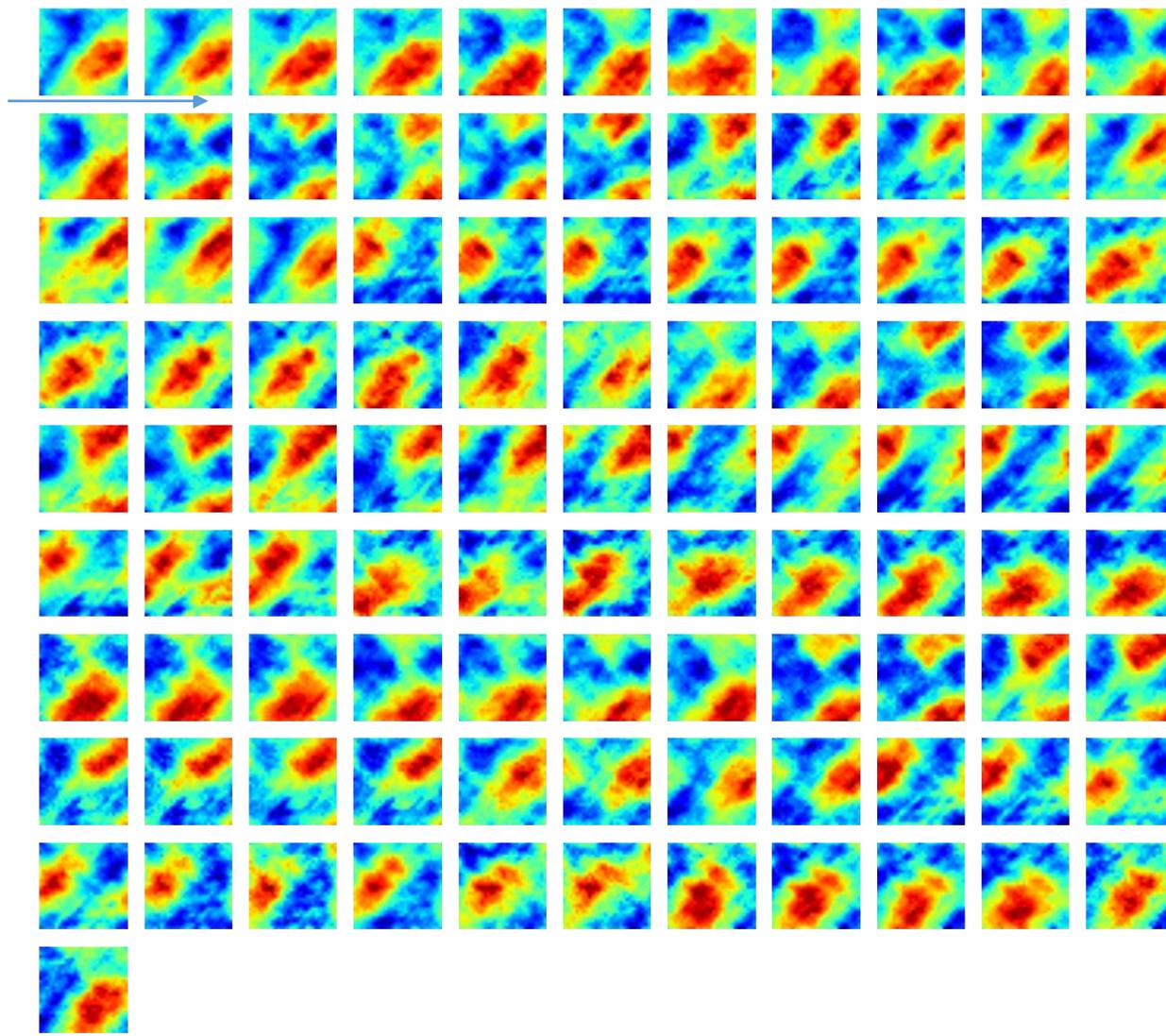
軌跡



● スタート

# Step 8000 ~ 8100

## 予測画像



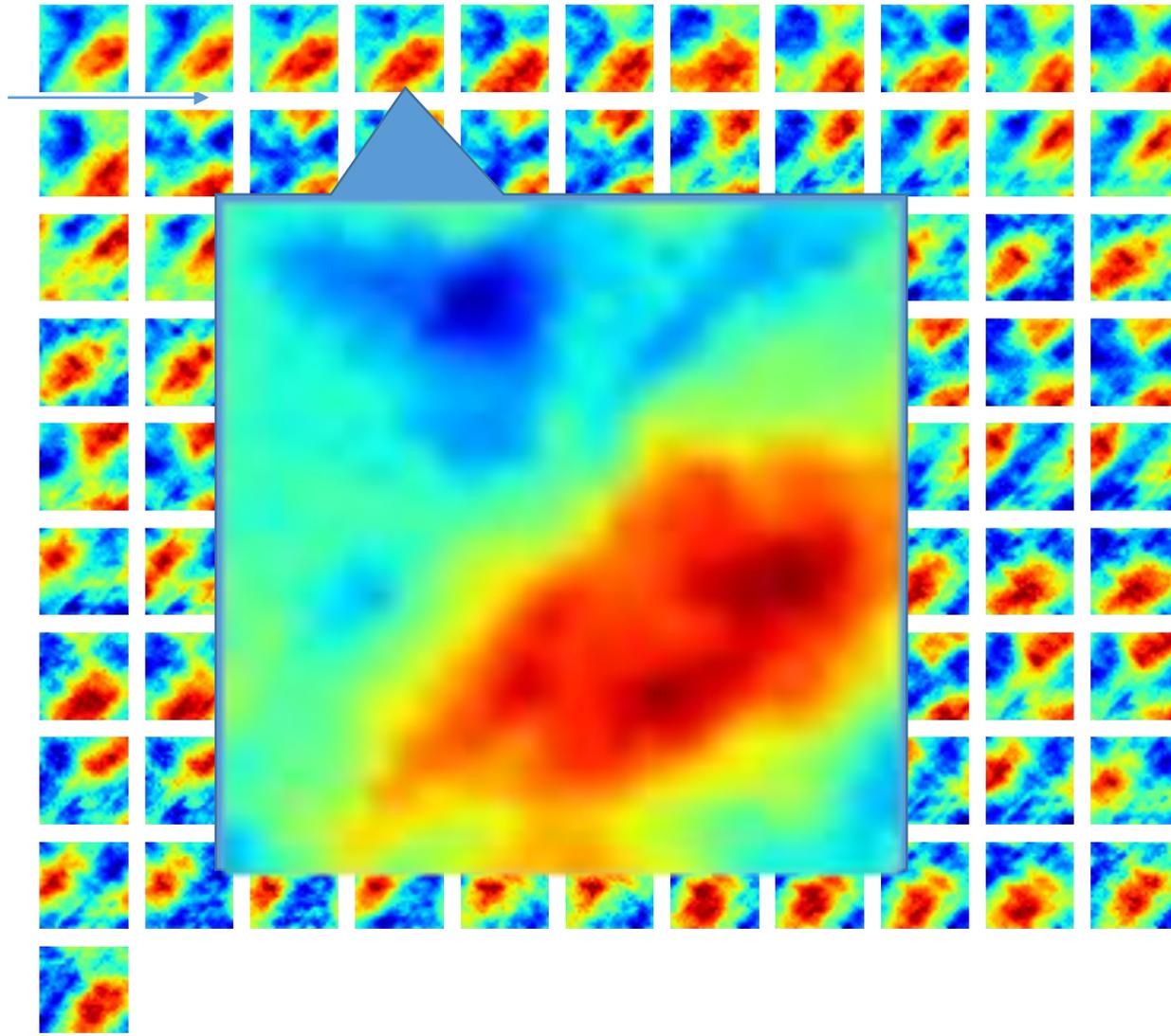
## 軌跡



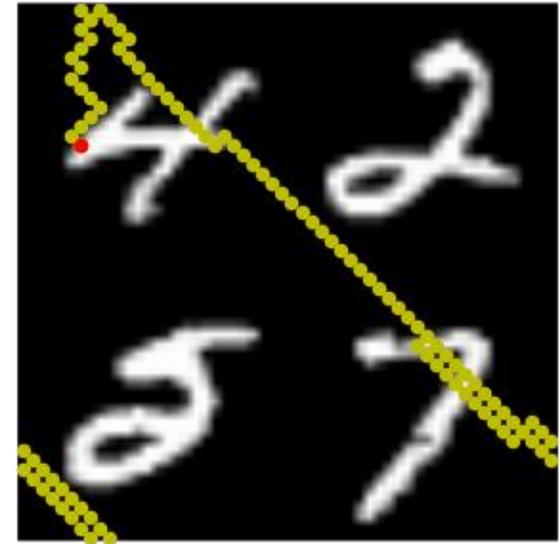
● スタート

Step 8000~8100

予測画像



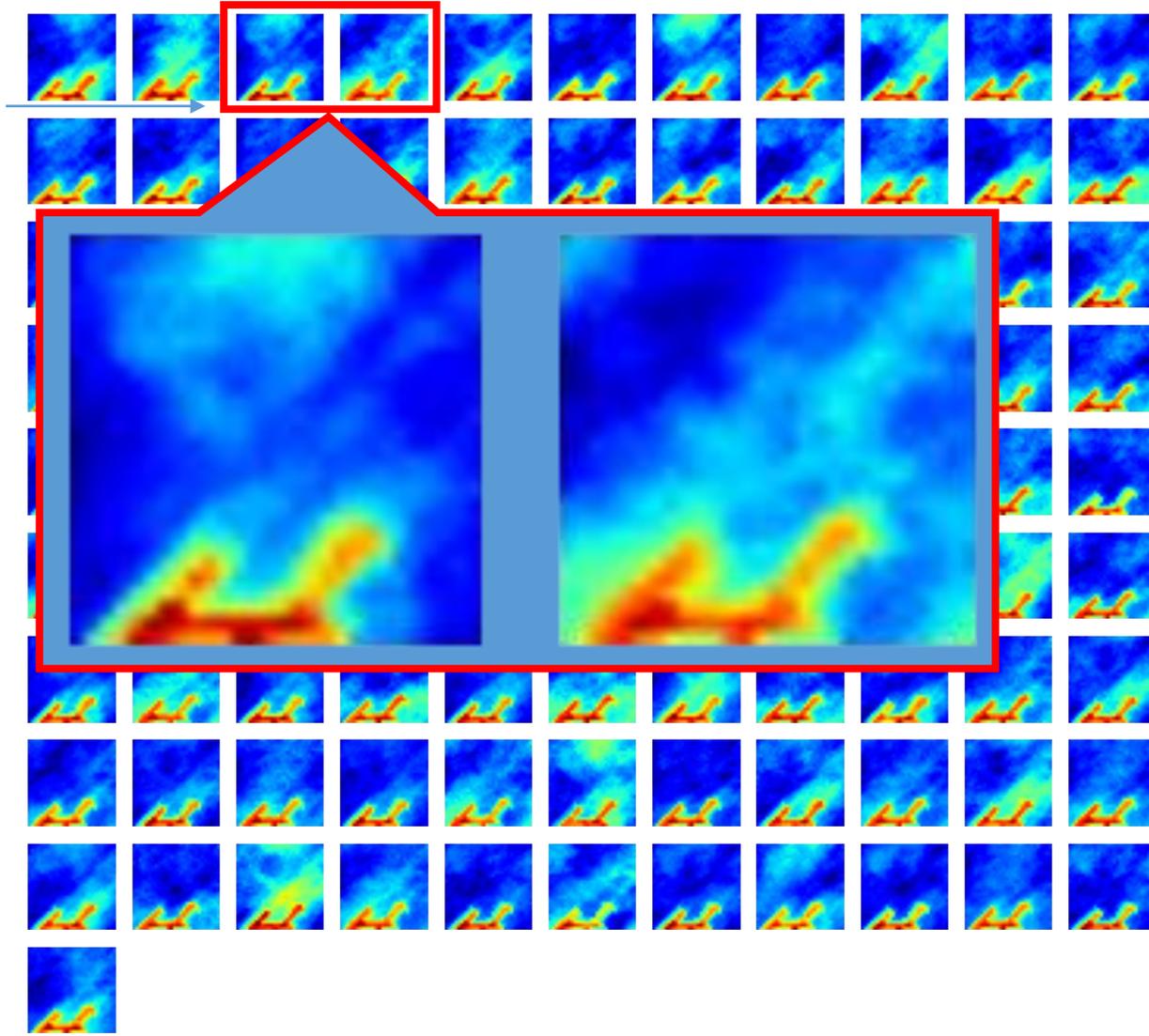
軌跡



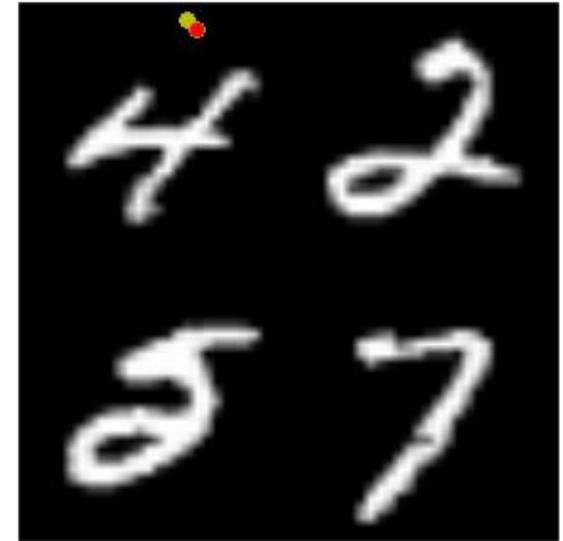
● スタート

Step 12000~12100

予測画像



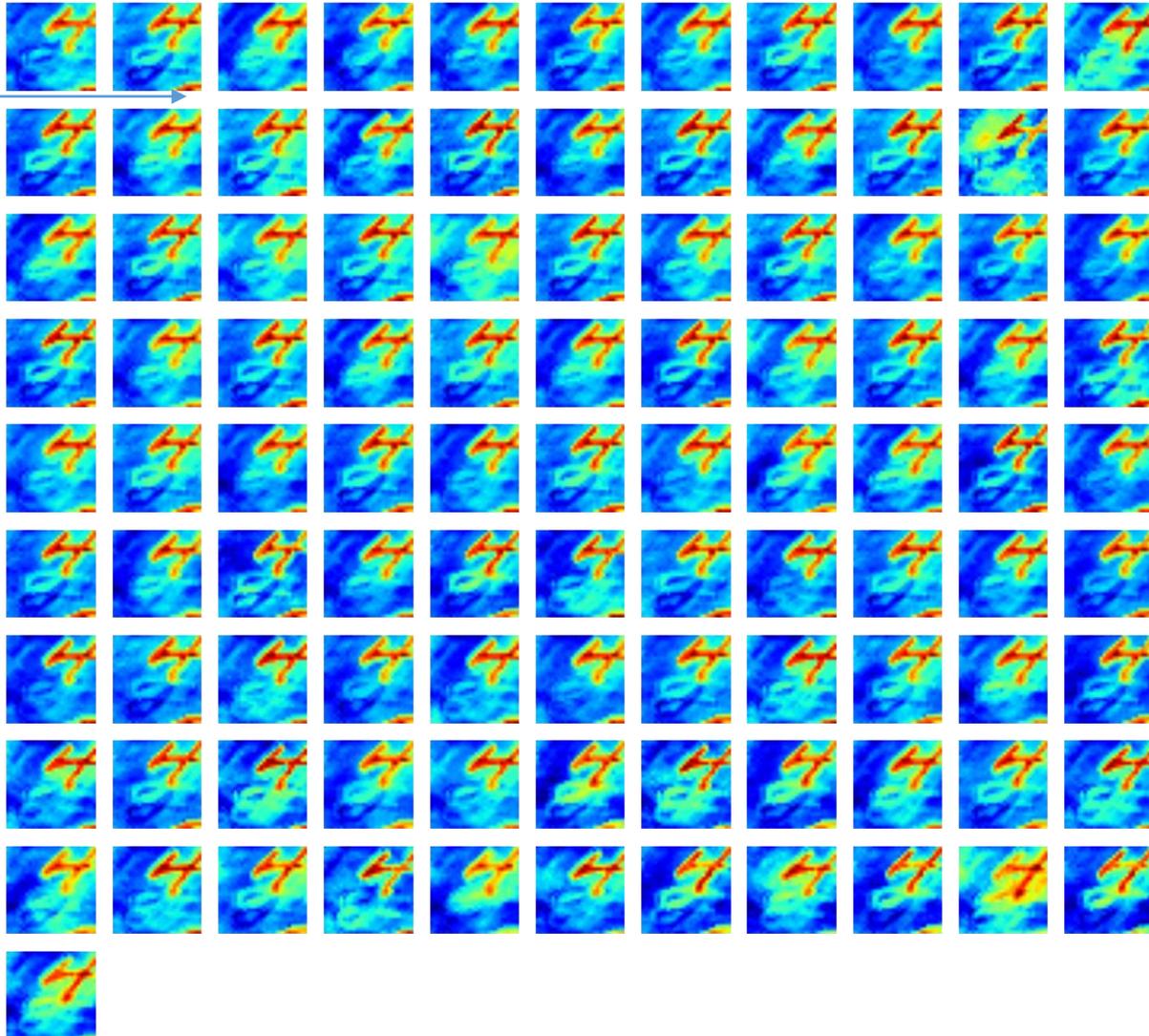
軌跡



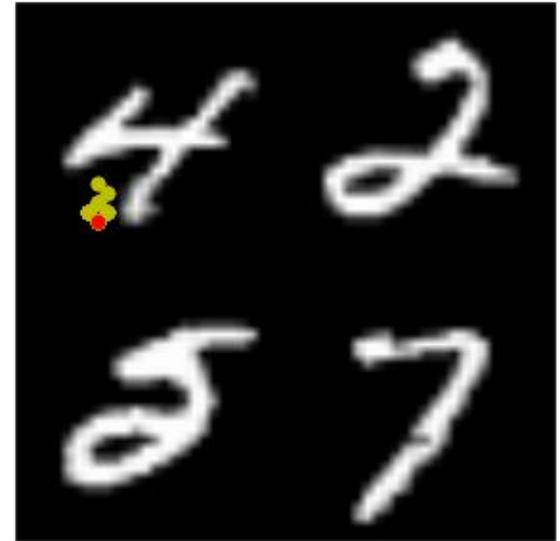
● スタート

Step 16000~16100

予測画像



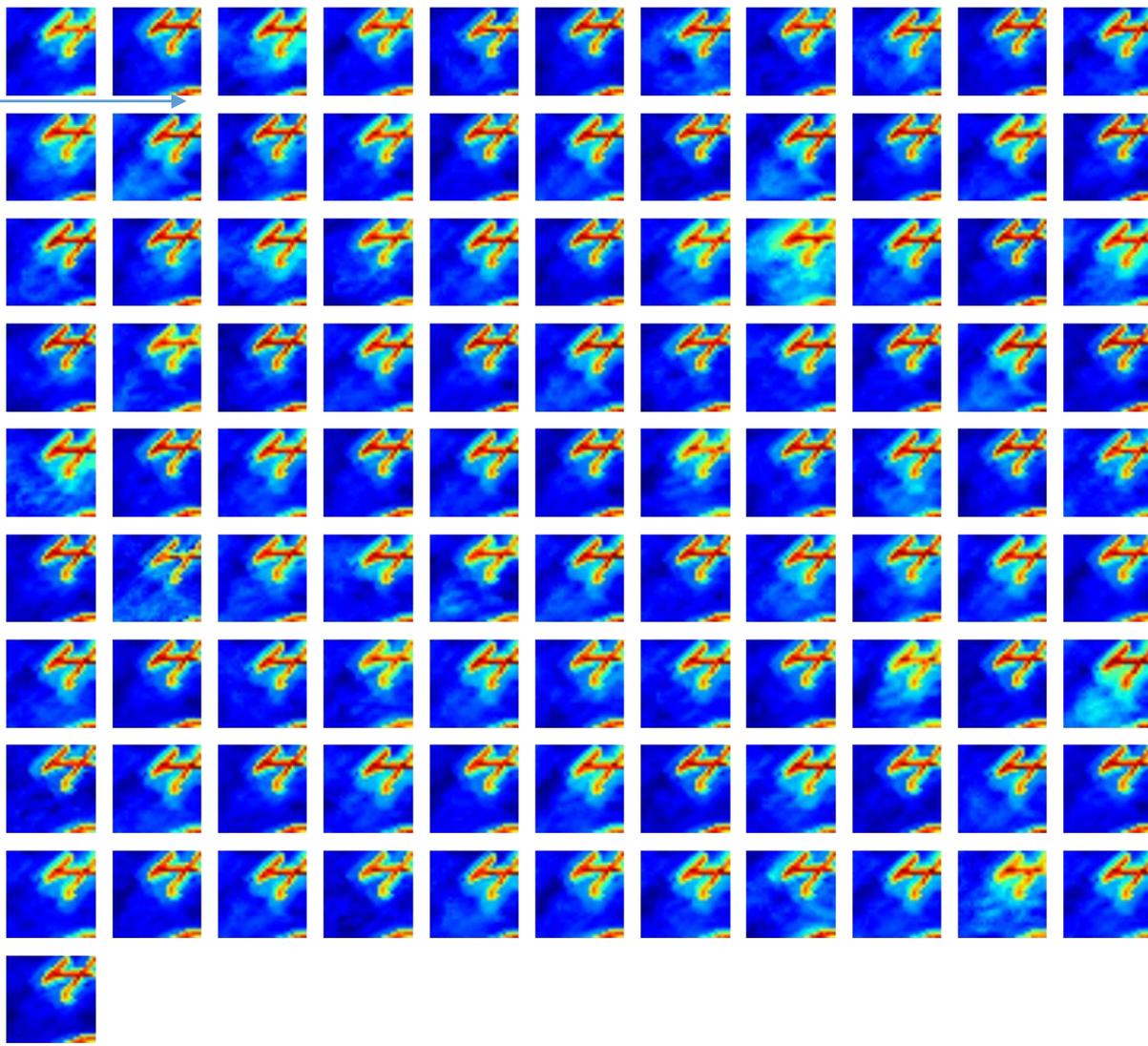
軌跡



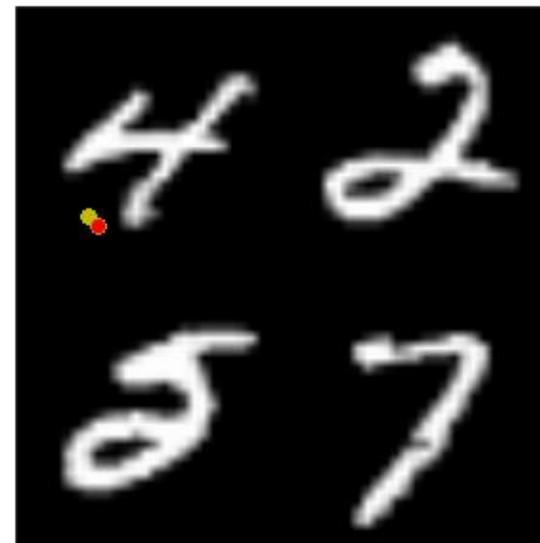
● スタート

# Step 20000 ~ 20100

予測画像

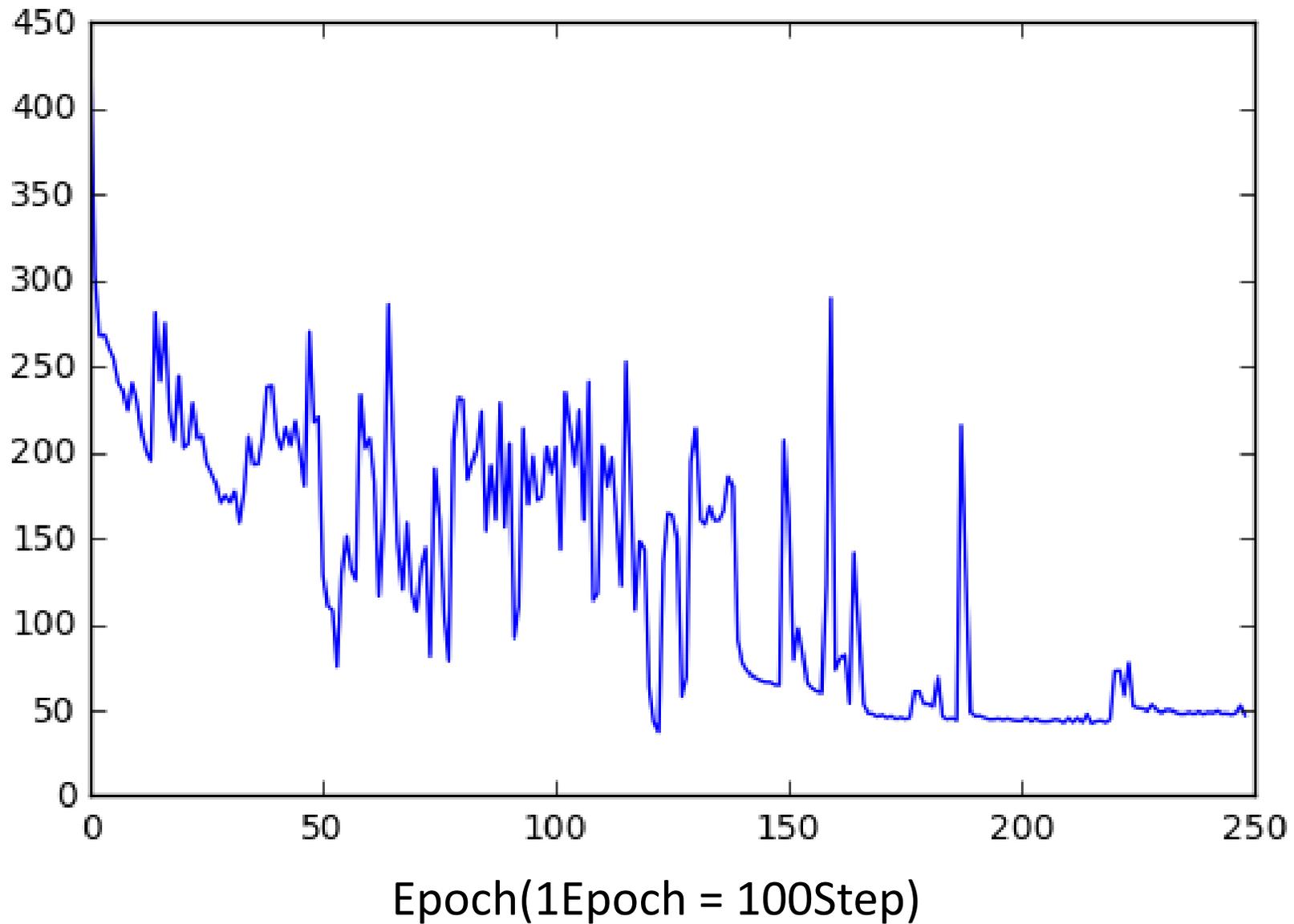


軌跡



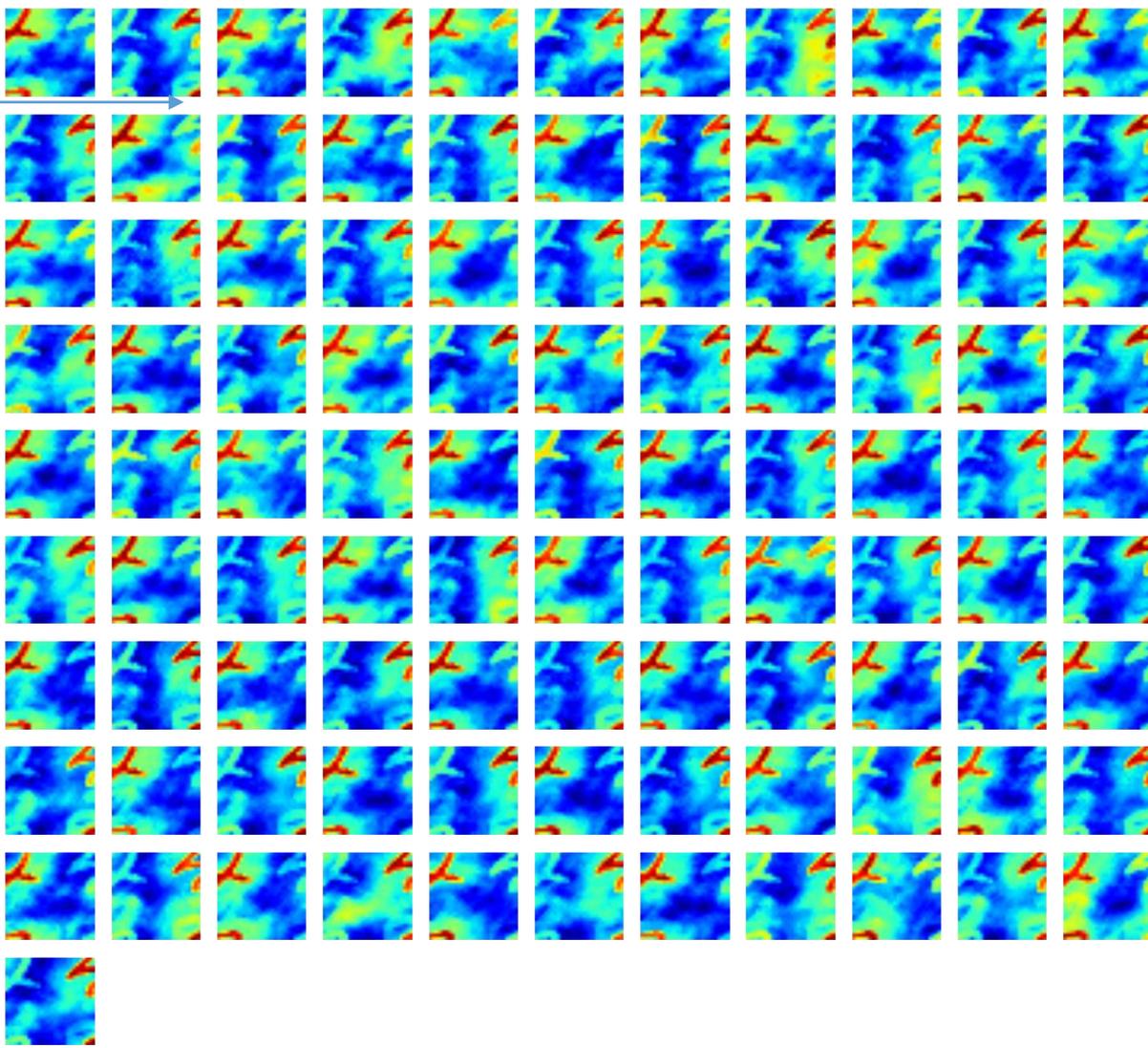
● スタート

# loss

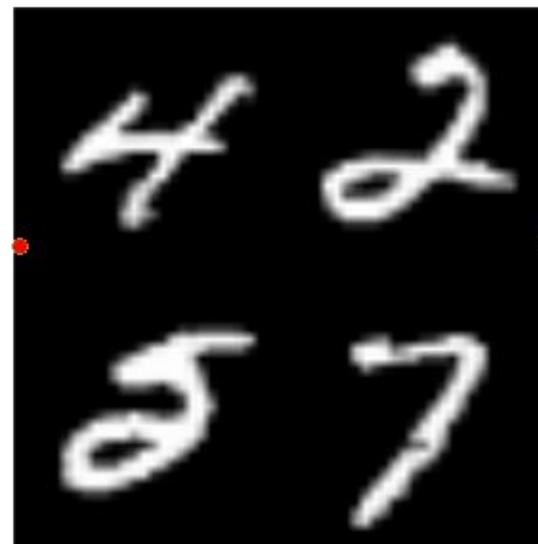


# LSTMを追加: Step 16000~16100

予測画像



軌跡

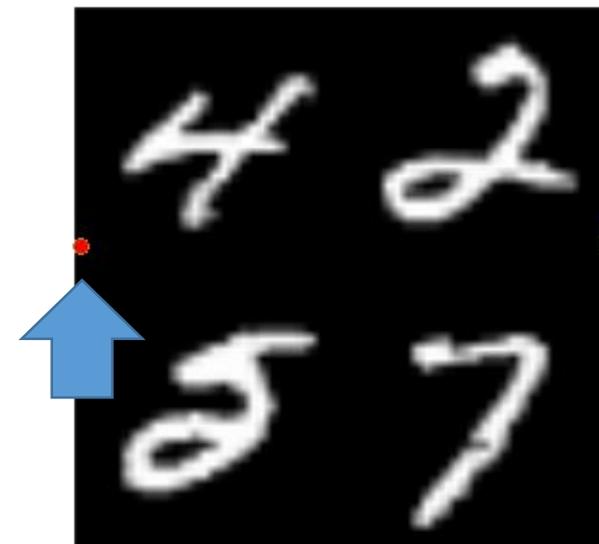
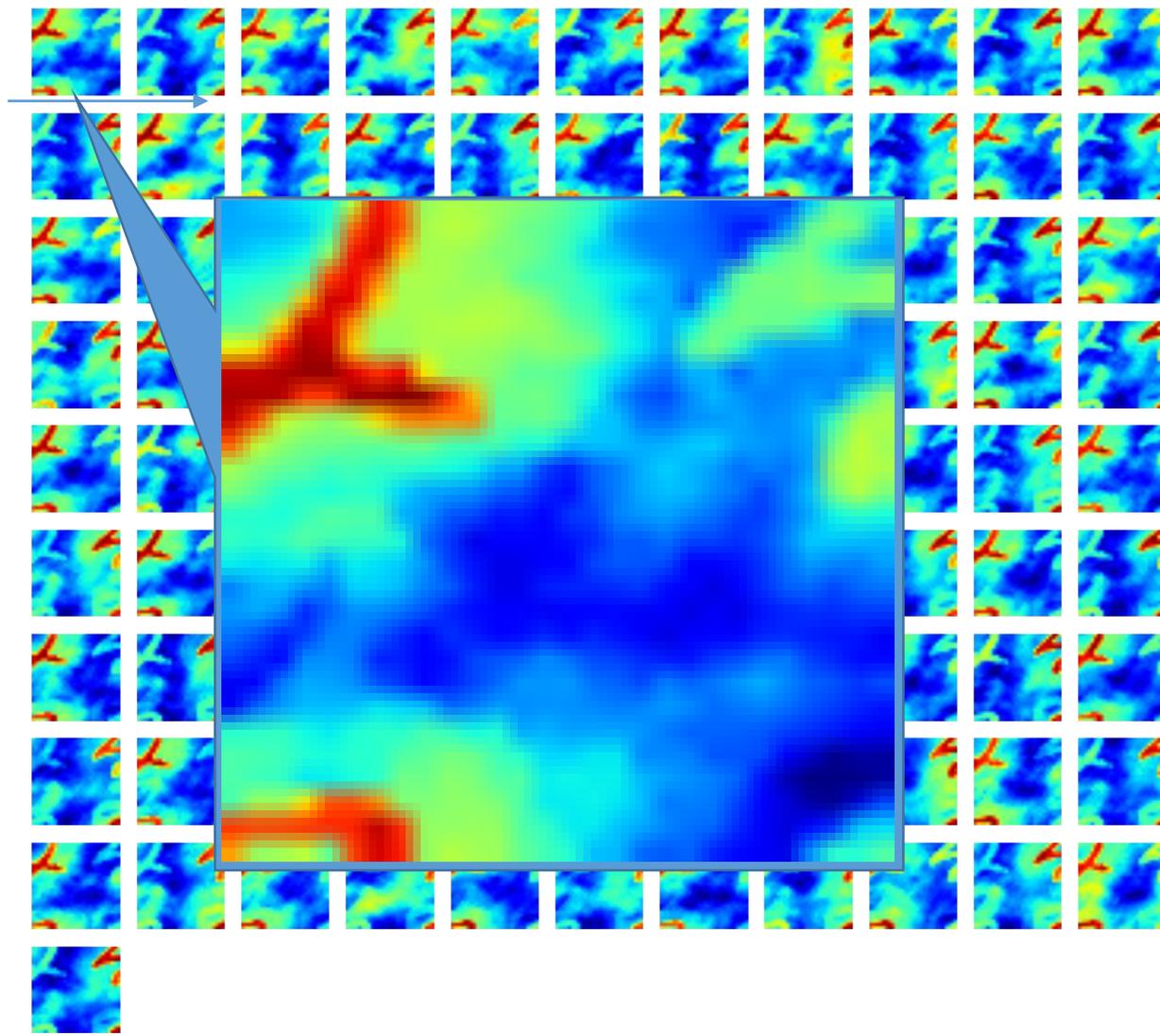


● スタート

# LSTMを追加: Step 16000~16100

予測画像

軌跡

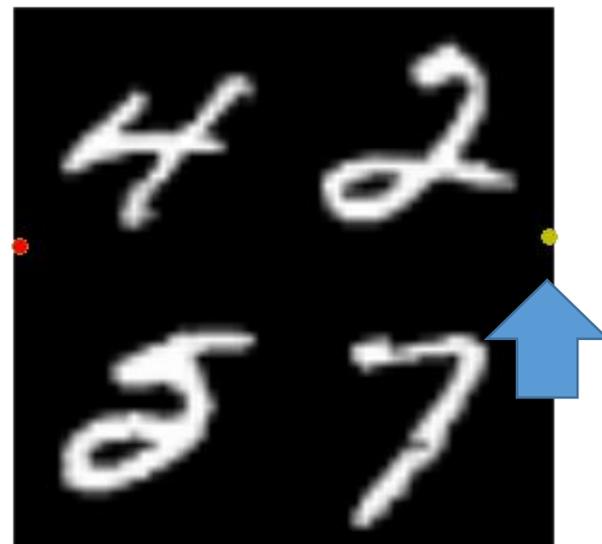
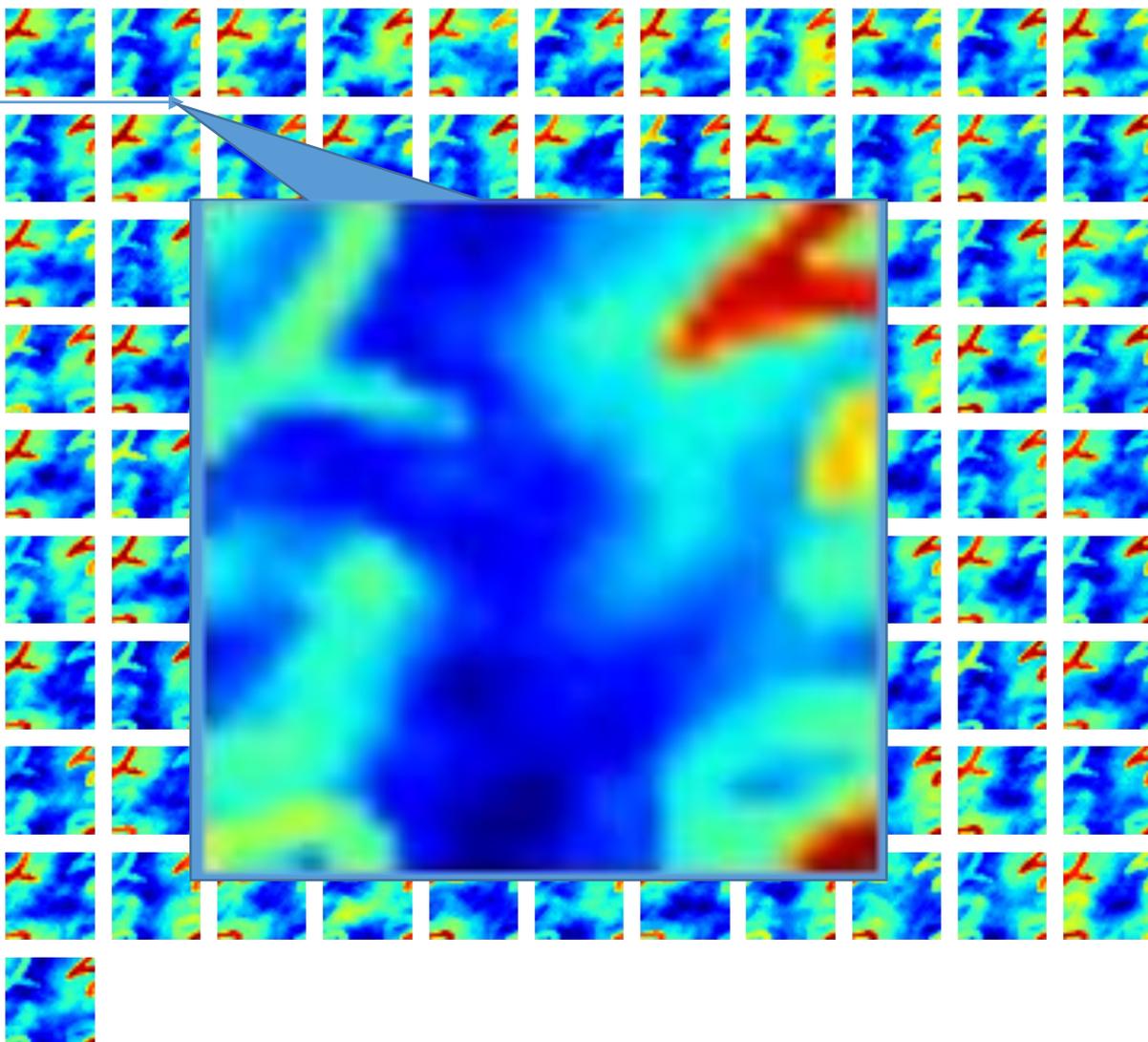


● スタート

# LSTMを追加: Step 16000~16100

予測画像

軌跡



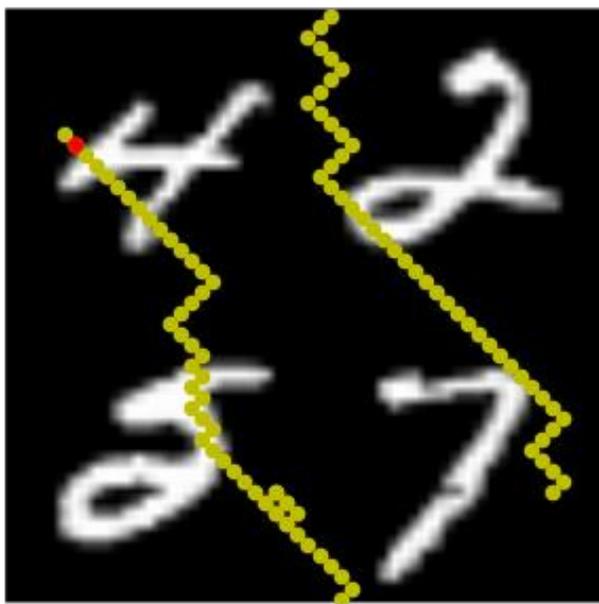
● スタート

# WBAIハッカソンでの結果

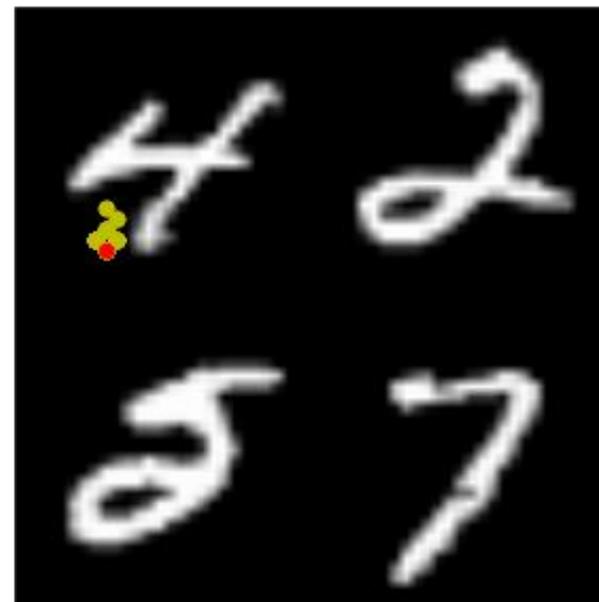
➡ 学習が進むと、視線移動が停止した

- スタート

Step 0~100



Step 16000~16100



学習

# ハッカソンの結果と課題

- ①生物のモデルとして静止し続けるのは不自然
- ②タスクの設計が，性能評価に適していない

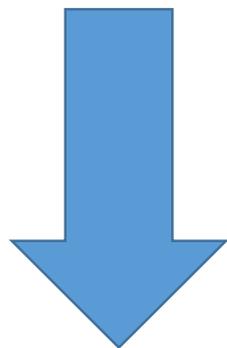


What's Next?

探索行動を取らせたい

追加実験  
(ハッカソン終了後)

目標：探索行動を取らせたい



方針：Epistemic Value

# Epistemic Value

$$\begin{aligned} \ln P(\hat{u}|\gamma) &= \gamma \cdot \mathbf{Q}(\pi) \\ &= \gamma \cdot (\mathbf{Q}_{t+1}(\pi) + \dots + \mathbf{Q}_T(\pi)) \end{aligned}$$

← 行動  $\hat{u}$  の選択確率は  $\mathbf{Q}_\tau(\pi)$  の時間総和

$$\mathbf{Q}_\tau(\pi) = \underbrace{E_{Q(o_t|\pi)}[\ln P(o_t|m)]}_{\text{Extrinsic Value}} + \underbrace{E_{Q(o_t|\pi)}[D[Q(s_t|o_t, \pi) || Q(s_t|\pi)]]}_{\text{Epistemic Value}}$$

Extrinsic Value



予想(期待)している観測

Epistemic Value



隠れ状態の曖昧さ減少量

# Epistemic Valueの近似

Epistemic Value

$$E_{Q(o_t|\pi)} [D[Q(s_t|o_t, \pi) || Q(s_t|\pi)]]$$



サンプル1つで近似

$$D[Q(s_t|o_t, \pi) || Q(s_t|\pi)]$$



方策  $\pi$  を直前の行動で近似

$$D[Q(s_t|o_t, a_{t-1}, o_{t-1}) || Q(s_t|a_{t-1}, o_{t-1})]$$

Epistemic ValueをいかにANNで実装するか？

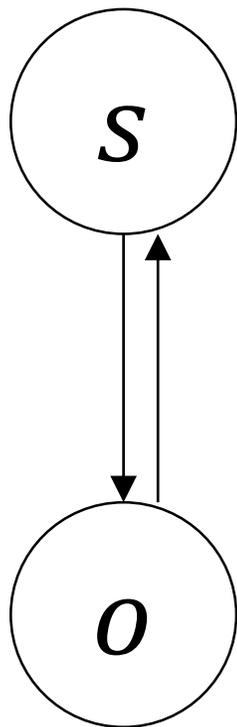
マルチモーダルVAE

# マルチモーダルVAE

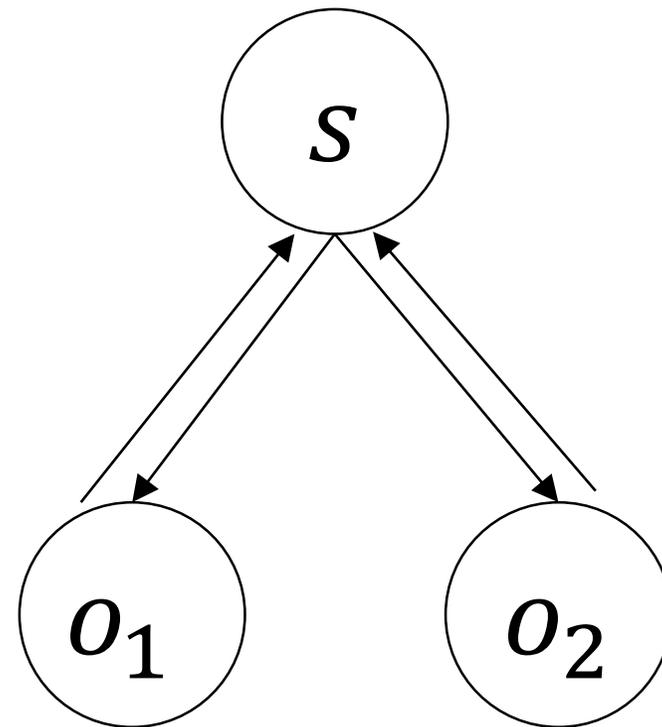
隠れ変数(分布)

ニューラルネット

観測変数(分布)

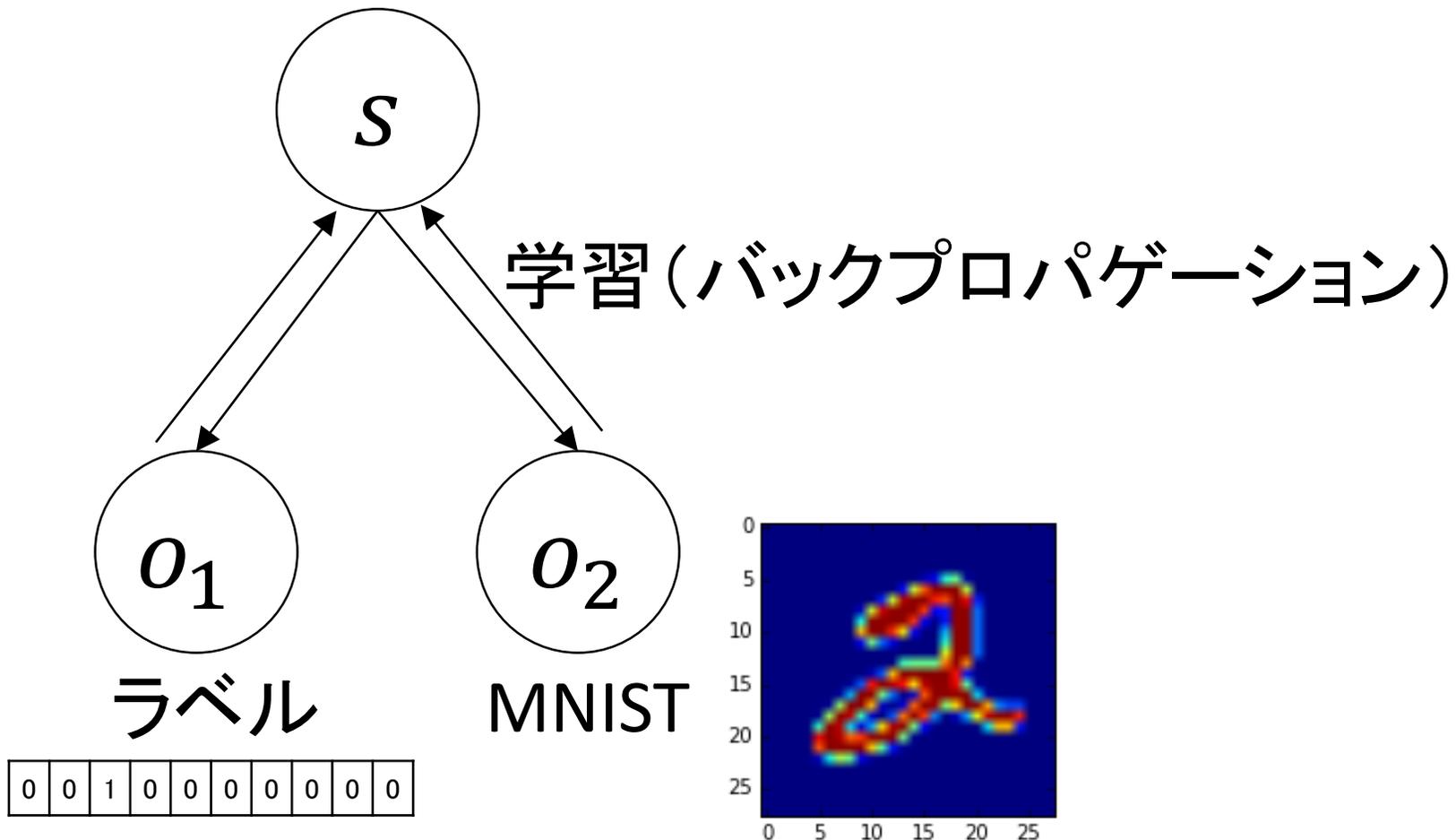


VAE

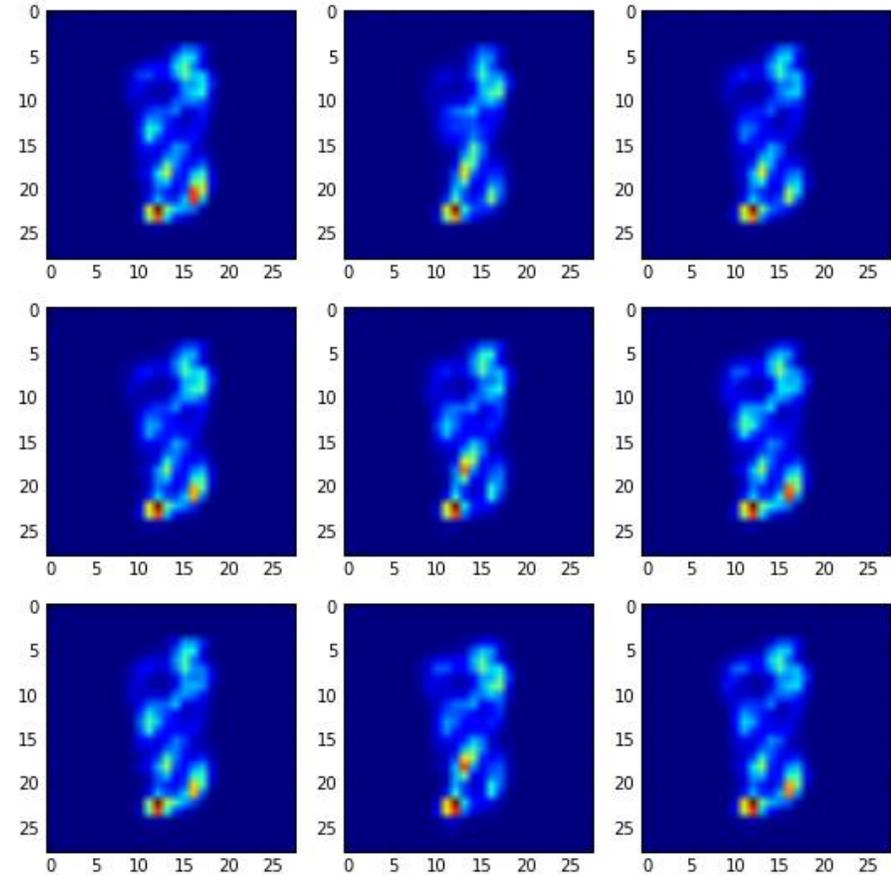
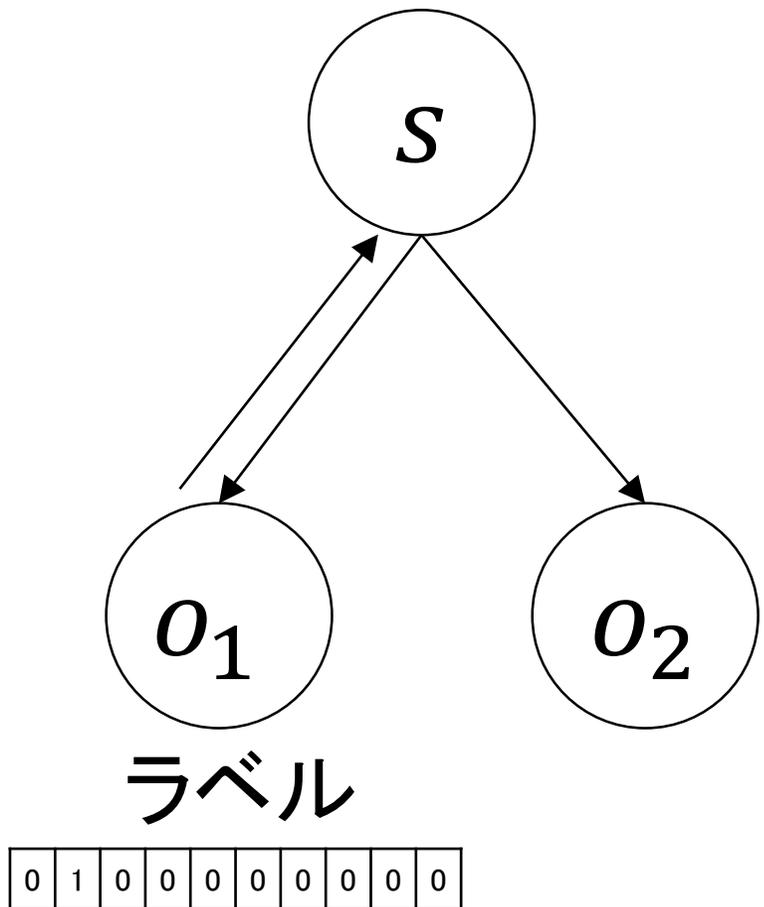


Multimodal VAE

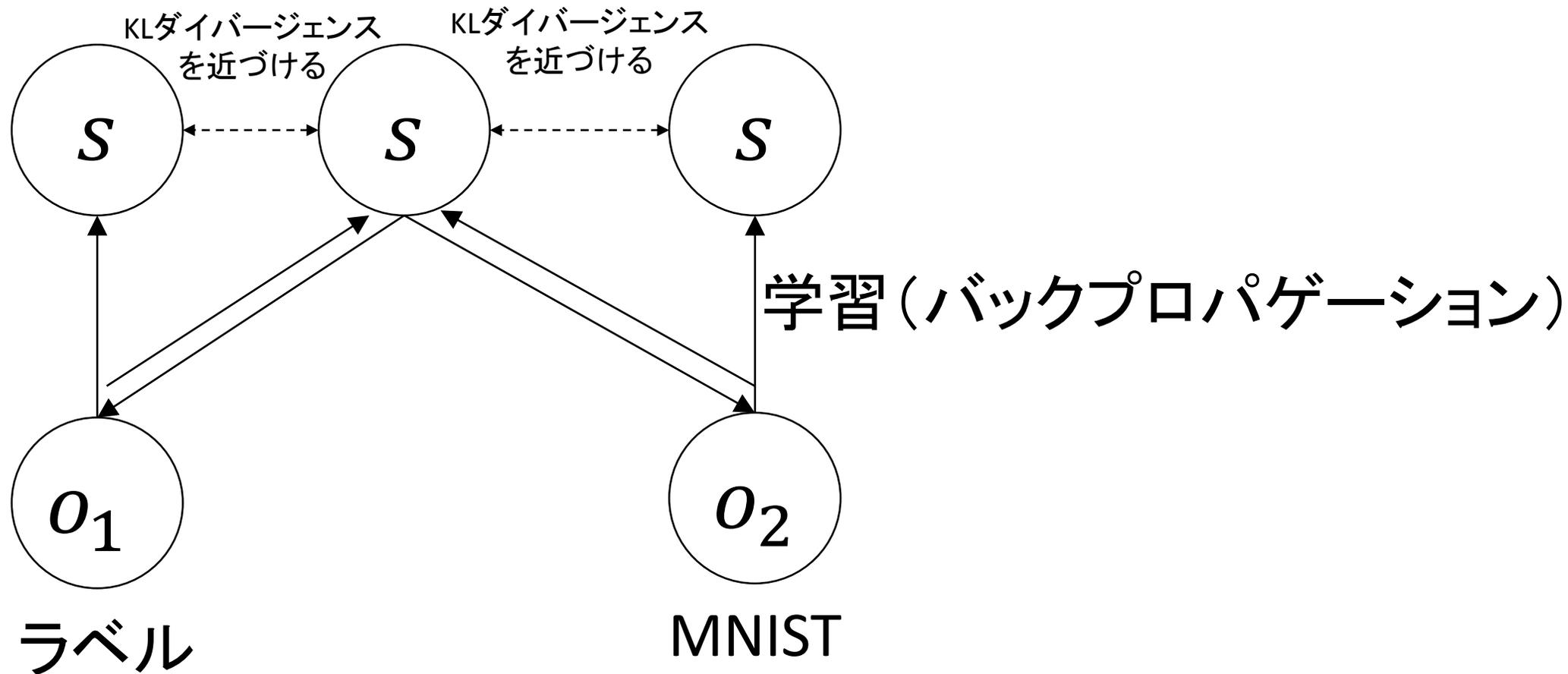
# マルチモーダルVAE(単純に可視層を分割)



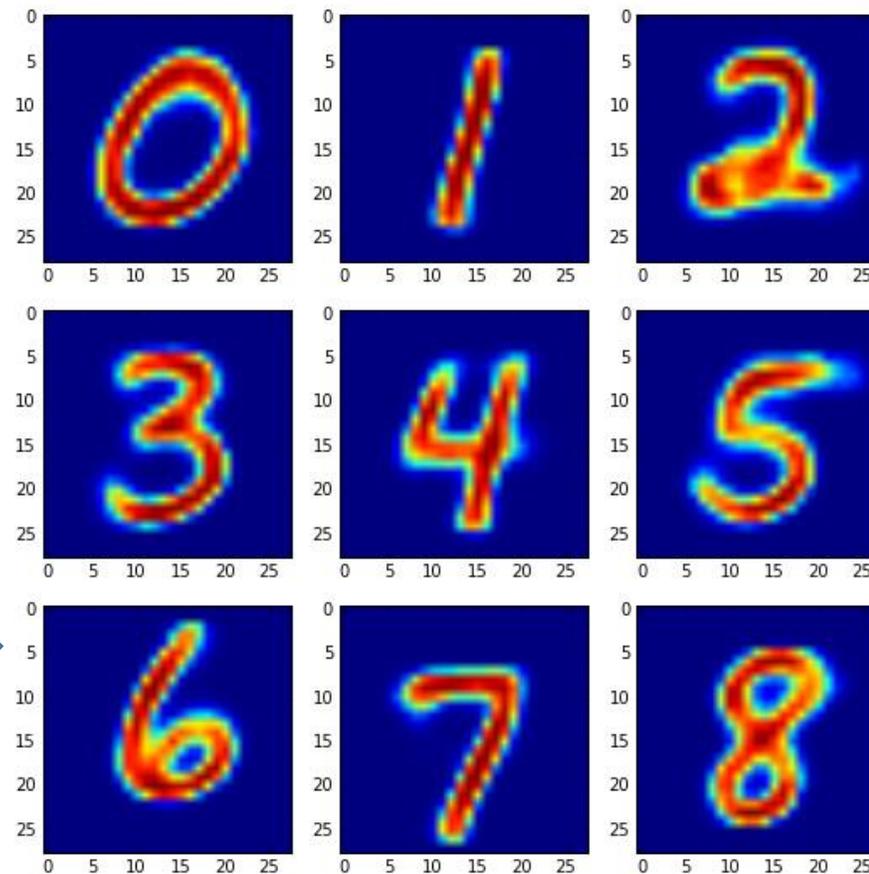
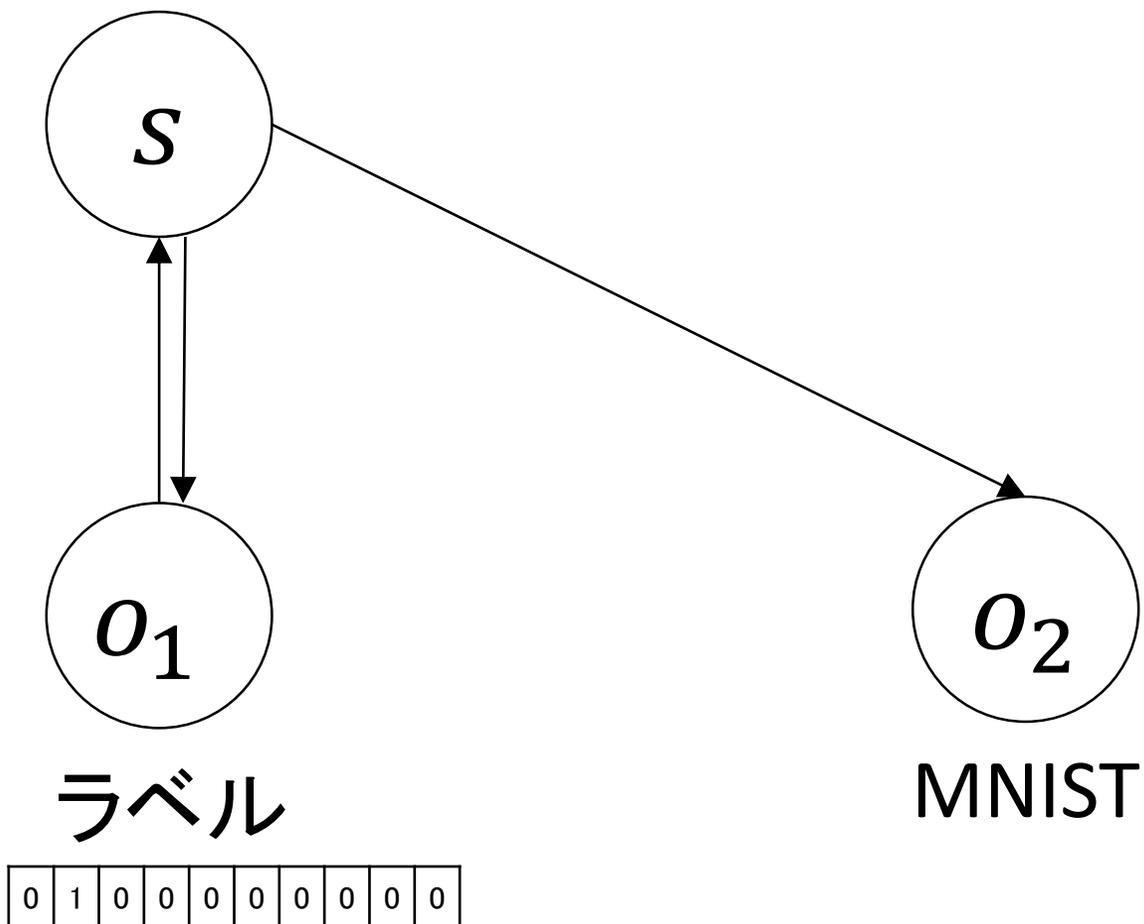
# マルチモーダルVAE(単純に可視層を分割)



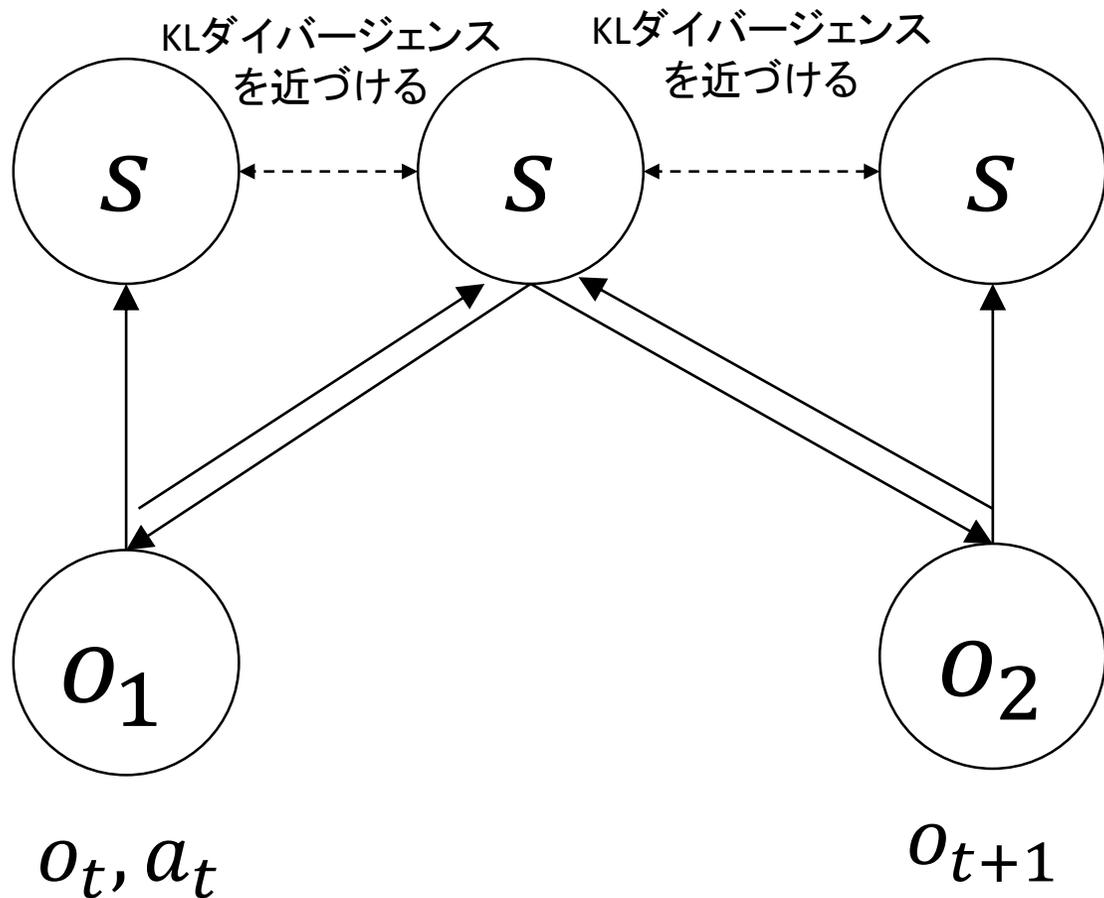
# マルチモーダルVAE



# マルチモーダルVAE



# マルチモーダルVAE(Epistemic Value用)

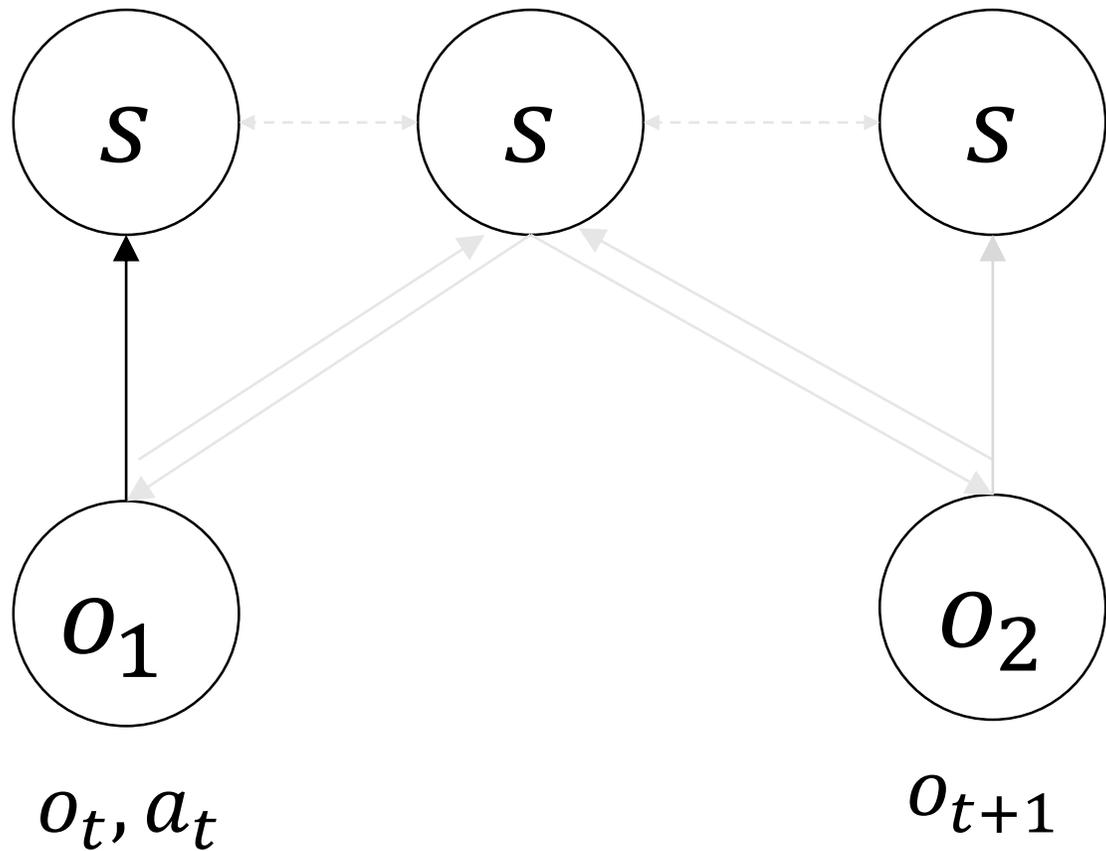


$O_t$  :現在の観測値

$a_t$  :現在の行動

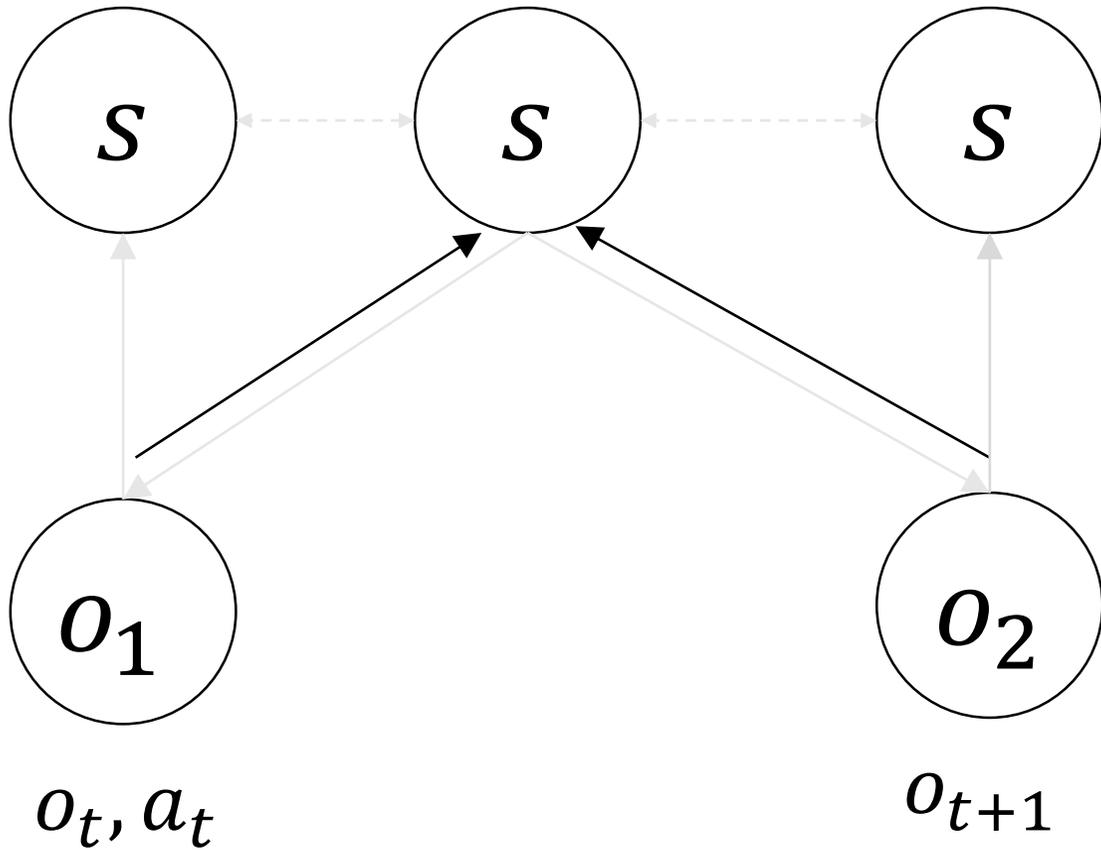
$O_{t+1}$  :次時刻の観測値

# マルチモーダルVAE(Epistemic Value用)



$$Q(s_t | a_{t-1}, o_{t-1})$$

# マルチモーダルVAE(Epistemic Value用)



$$Q(s_t | a_{t-1}, o_{t-1})$$

$$Q(s_t | o_t, a_{t-1}, o_{t-1})$$

# マルチモーダルVAE(Epistemic Value用)

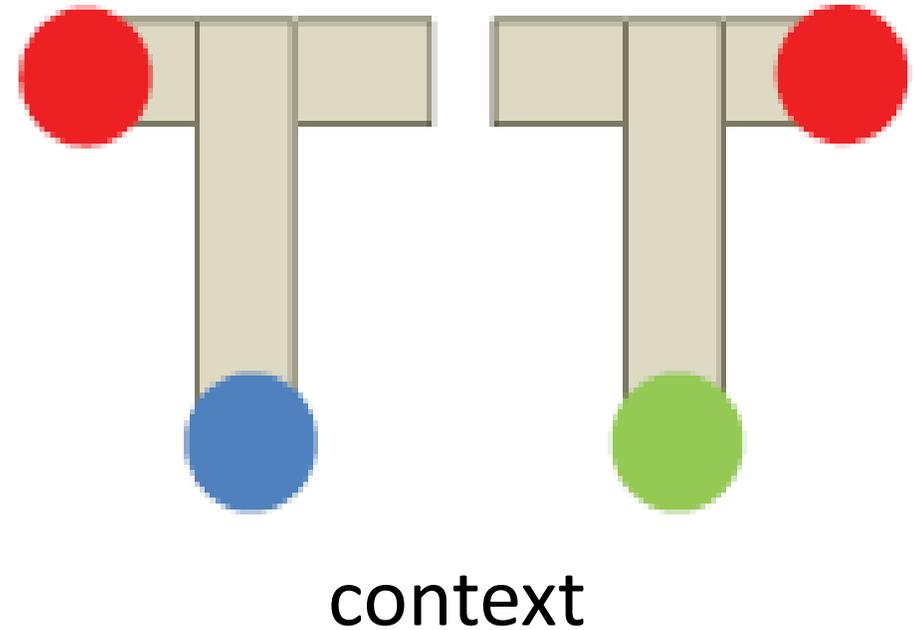
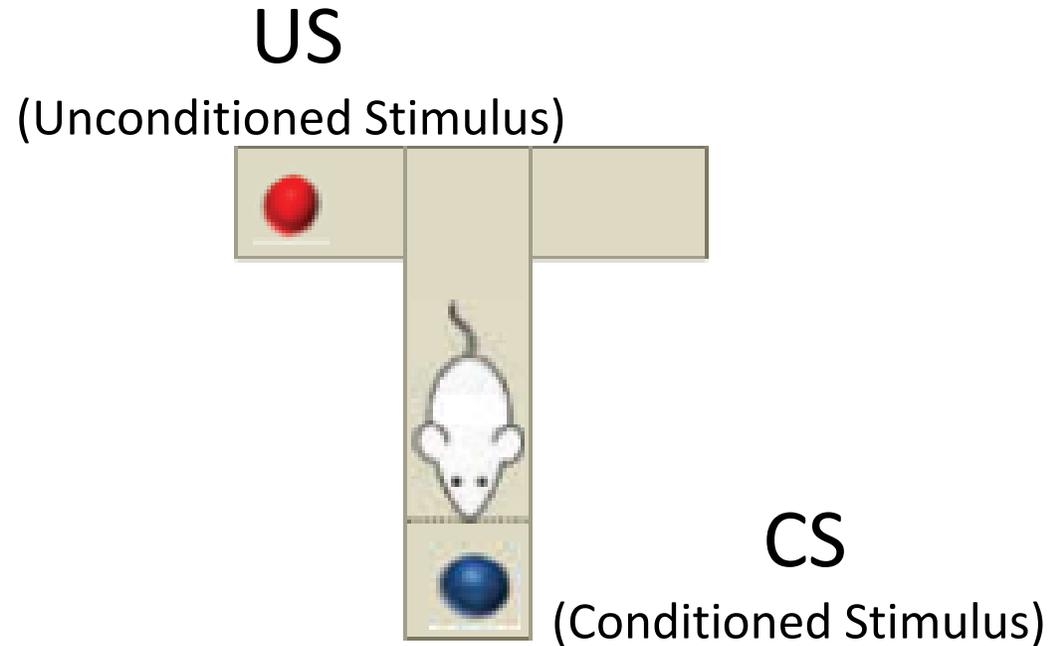
$$Q(s_t | a_{t-1}, o_{t-1})$$

$$Q(s_t | o_t, a_{t-1}, o_{t-1})$$

Epistemic Value の近似値

$$D[Q(s_t | o_t, a_{t-1}, o_{t-1}) || Q(s_t | a_{t-1}, o_{t-1})]$$

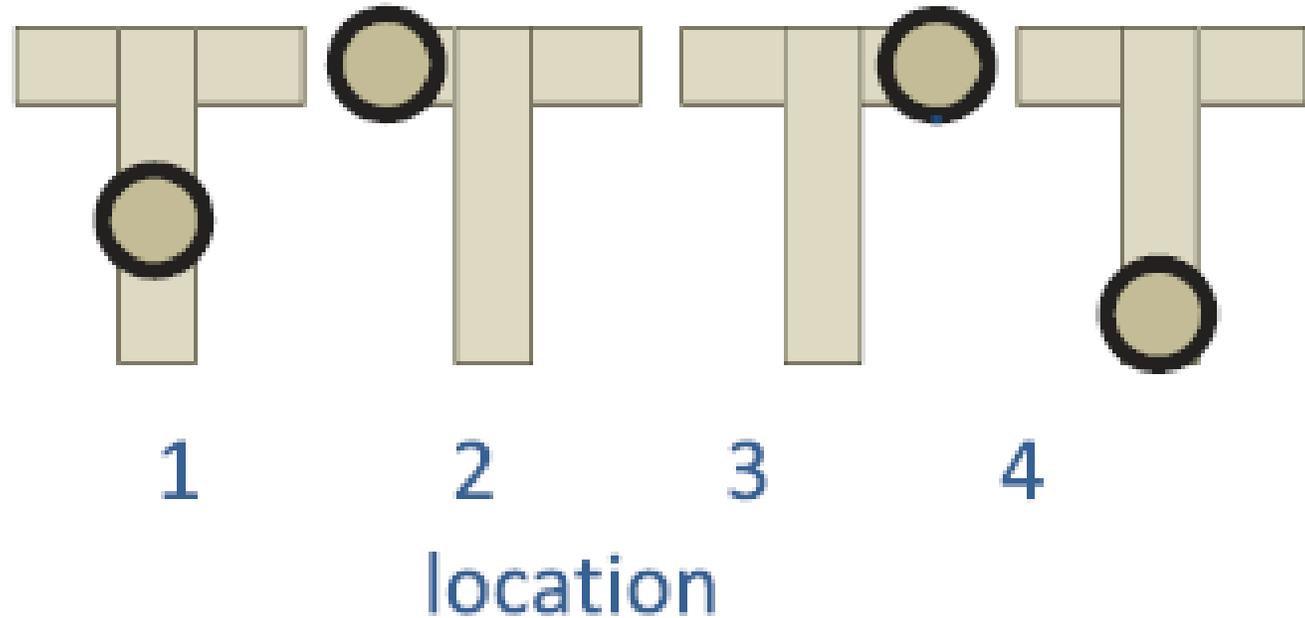
# 課題設定(論文の条件)



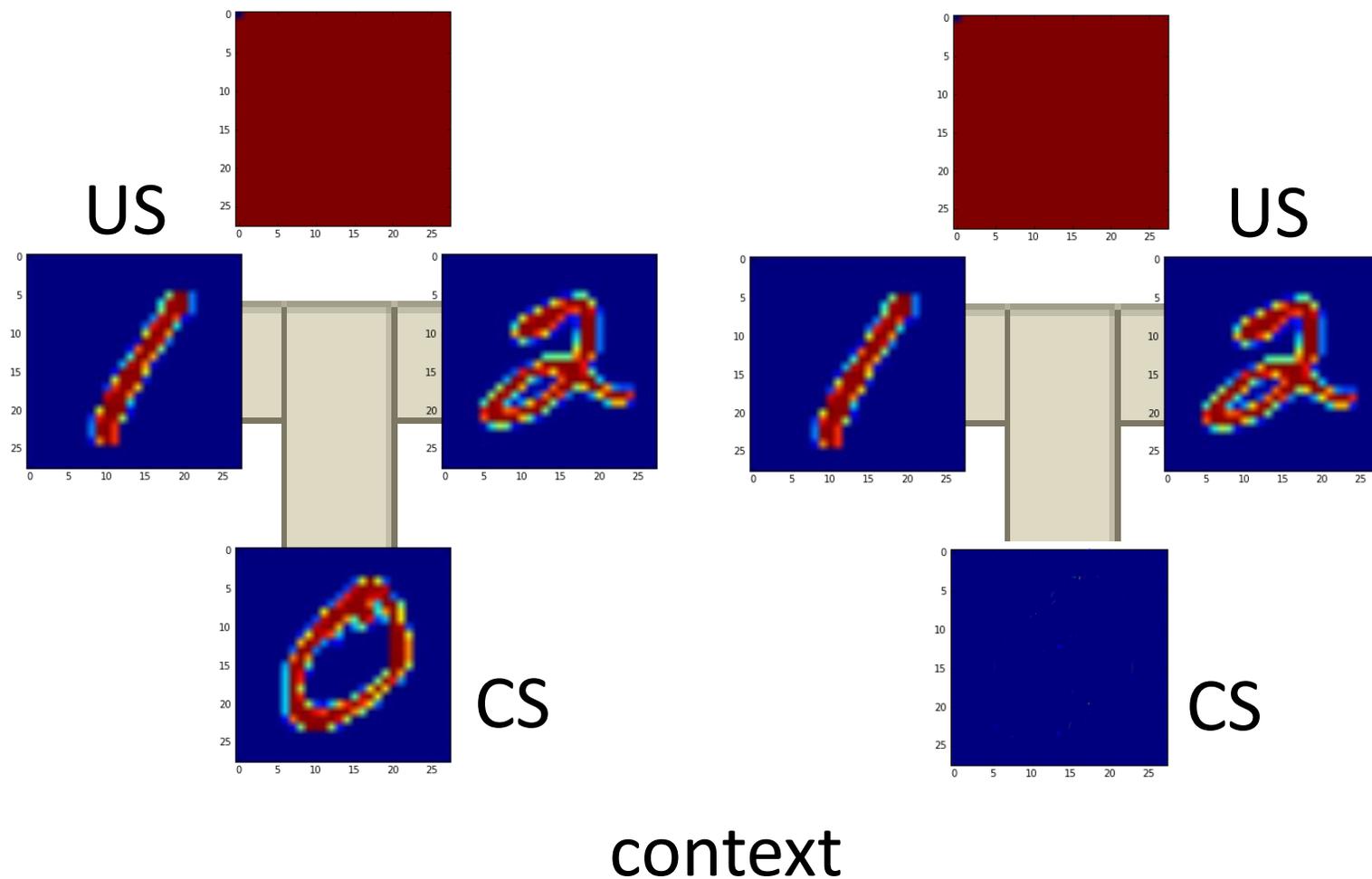
# 課題設定(論文の条件)

Control states

$$\tilde{u} = u_t, \dots, u_T$$



# 課題設定(今回の条件)

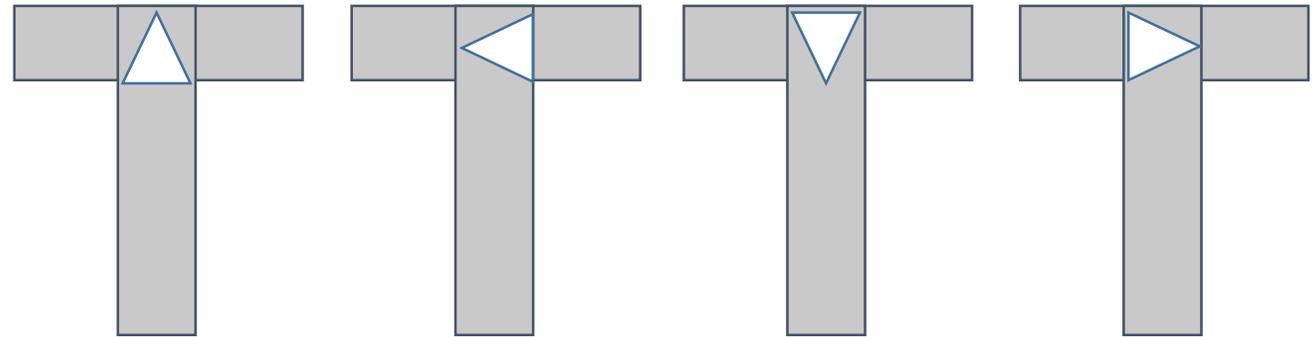


- 位置情報の代わりに画像を入力
- 報酬は左右どちらかの通路にある
- CSを先に観測する必要がある

# 課題設定(今回の条件)

Control states

$$\tilde{u} = u_1, \dots, u_T$$



1 ←→ 2 ←→ 3 ←→ 4



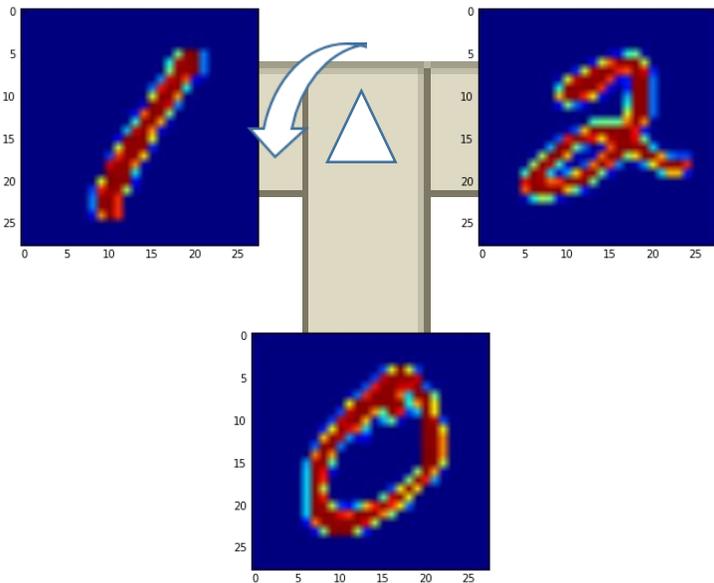
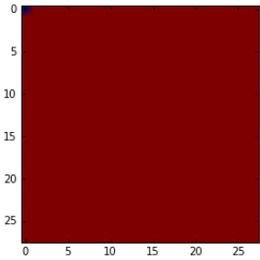
隣接する状態でのみ変化

# 実験方法

1. ランダムにアクションを取り  $o_t, a_{t-1}, o_{t-1}$  をサンプリング
2. マルチモーダルVAEにサンプルを学習
3. 学習済みのマルチモーダルVAEの隠れ層を使い Epistemic Valueを計算

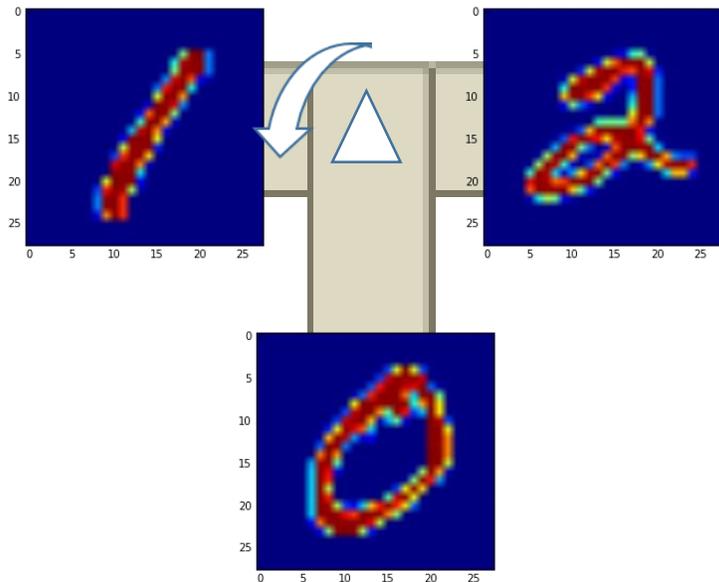
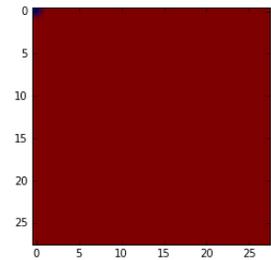
# Epistemic Value 1 : 上向きから左向き

$$D[Q(s_t|o_t, a_{t-1}, o_{t-1})||Q(s_t|a_{t-1}, o_{t-1})] = 0.501$$

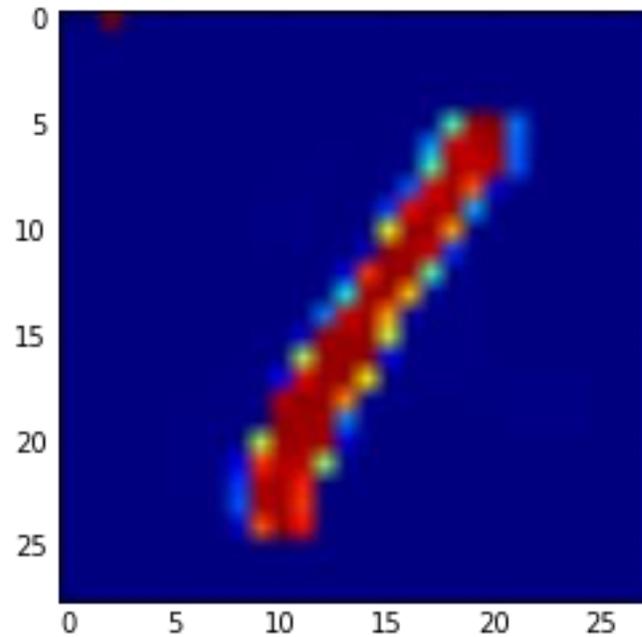


# Epistemic Value 1 : 上向きから左向き

$$D[Q(s_t|o_t, a_{t-1}, o_{t-1}) || Q(s_t|a_{t-1}, o_{t-1})] = 0.501$$



$Q(o_t|s_t)$

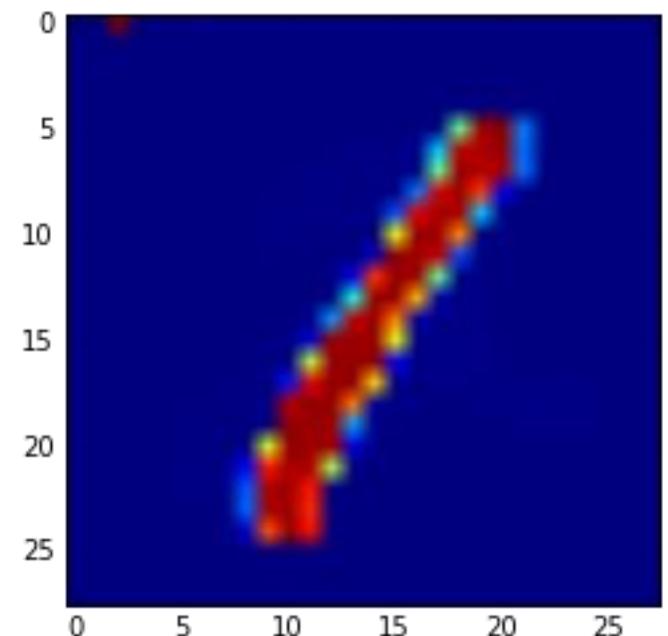
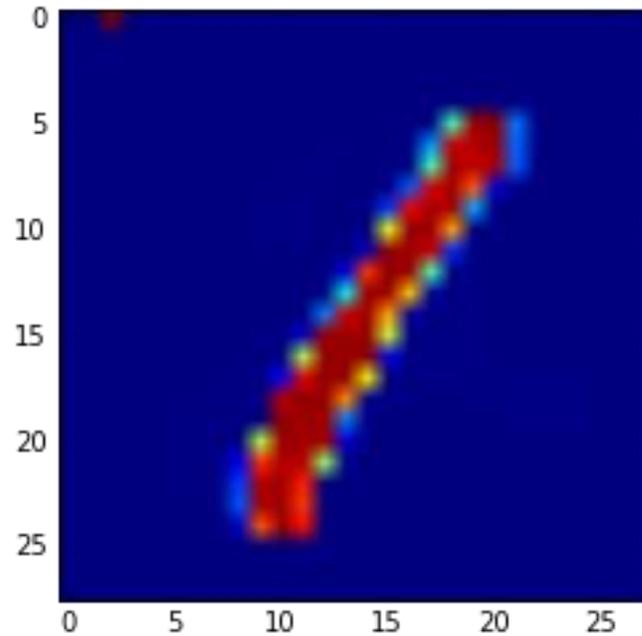
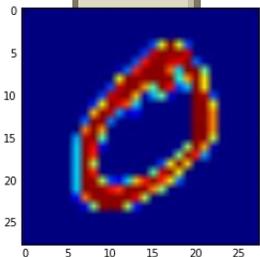
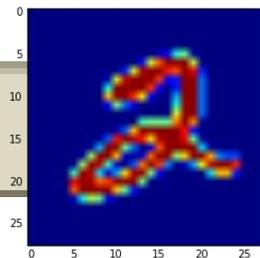
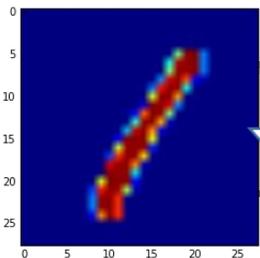
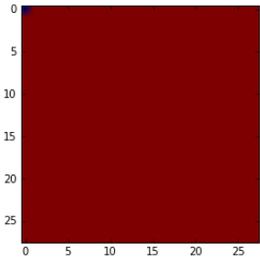


# Epistemic Value 1: 上向きから左向き

$$D[Q(s_t|o_t, a_{t-1}, o_{t-1}) || Q(s_t|a_{t-1}, o_{t-1})] = 0.501$$

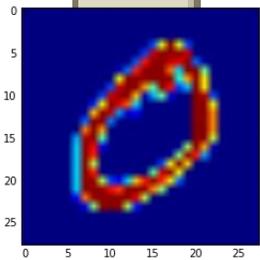
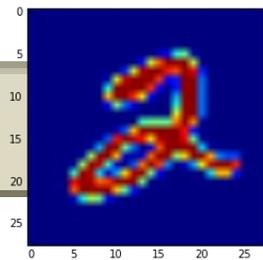
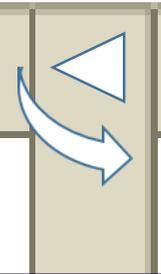
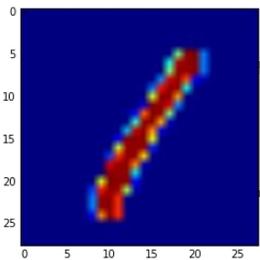
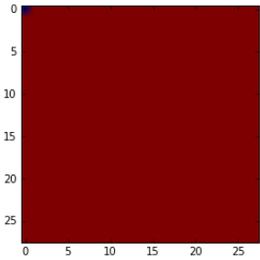
$Q(o_t|s_t)$

$Q(o_t|s_t)$



# Epistemic Value 1: 左向きから下向き

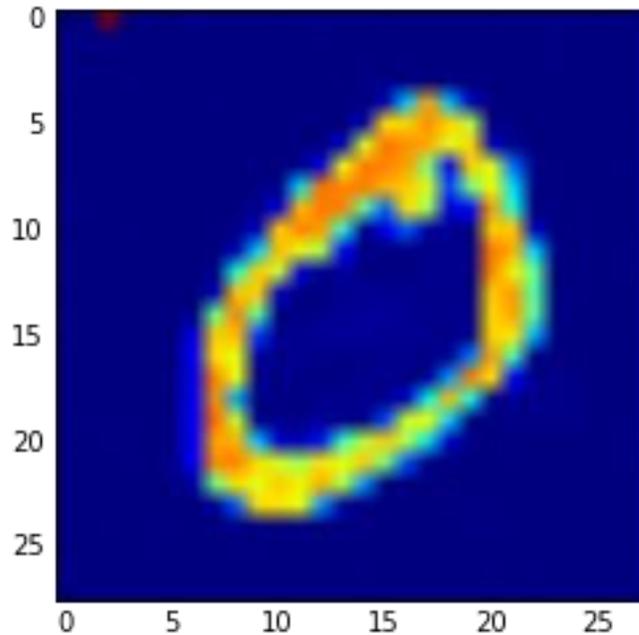
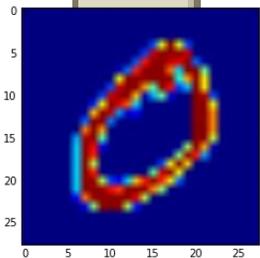
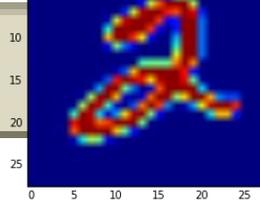
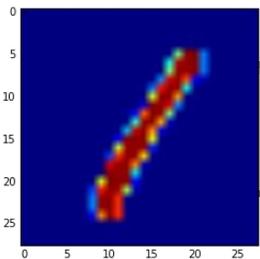
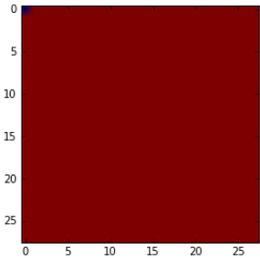
$$D[Q(s_t|o_t, a_{t-1}, o_{t-1})||Q(s_t|a_{t-1}, o_{t-1})] = 7.262$$



# Epistemic Value 2: 左向きから下向き

$$D[Q(s_t|o_t, a_{t-1}, o_{t-1}) || Q(s_t|a_{t-1}, o_{t-1})] = 7.262$$

$Q(o_t|s_t)$

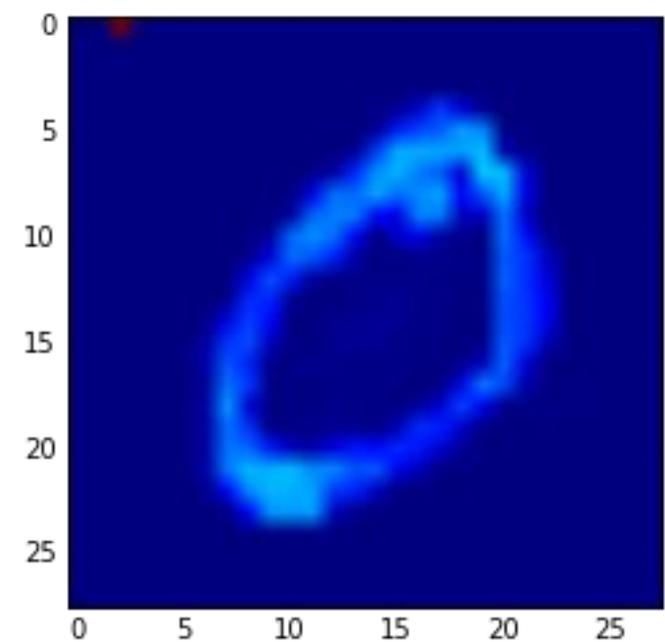
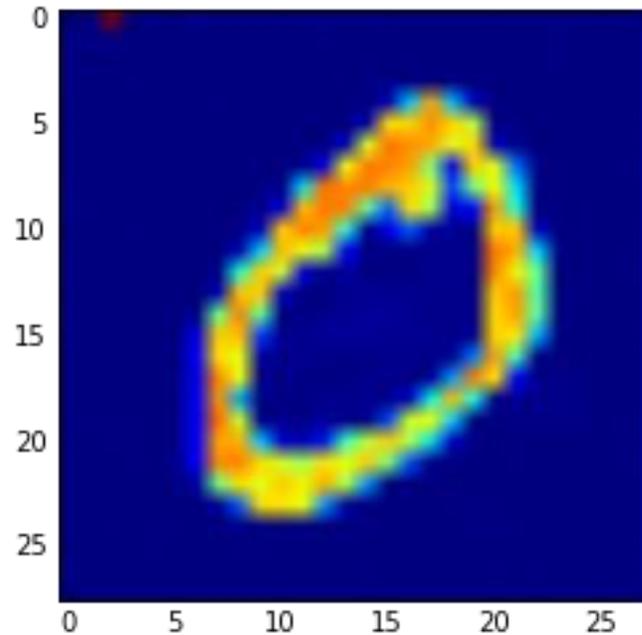
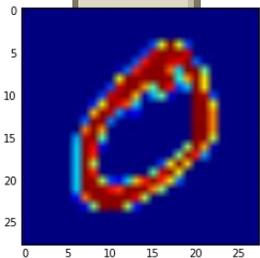
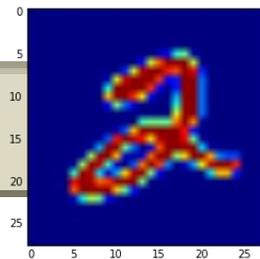
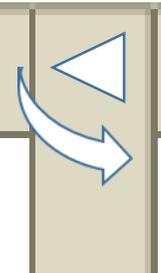
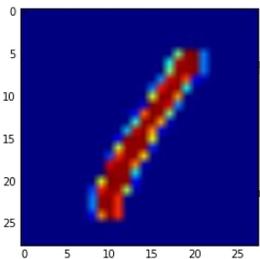
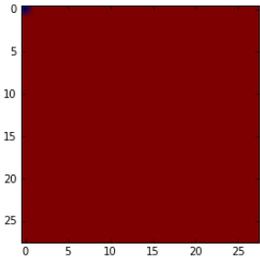


# Epistemic Value 2: 左向きから下向き

$$D[Q(s_t|o_t, a_{t-1}, o_{t-1}) || Q(s_t|a_{t-1}, o_{t-1})] = 3.319$$

$Q(o_t|s_t)$

$Q(o_t|s_t)$

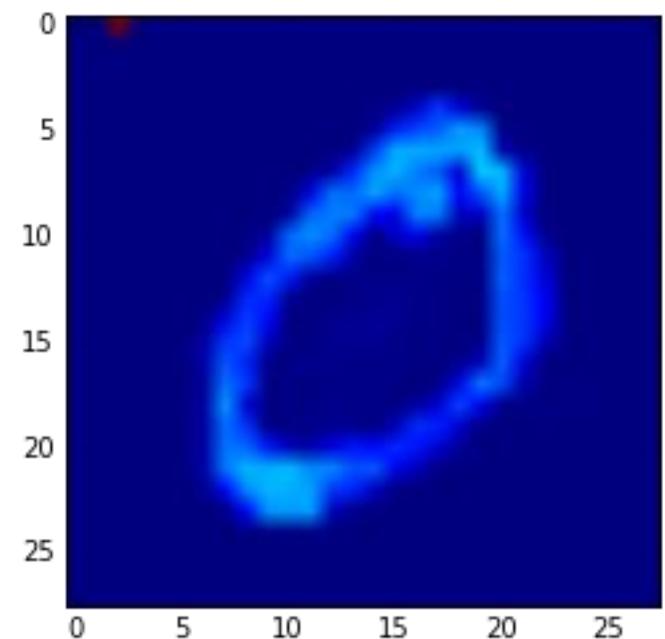
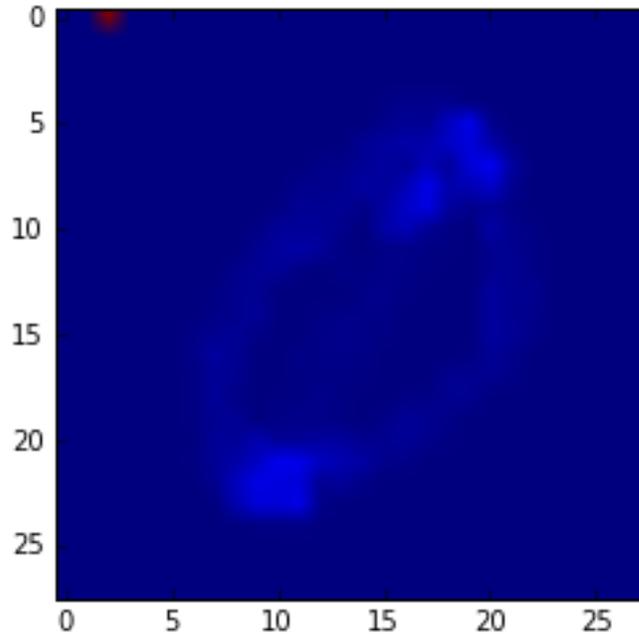
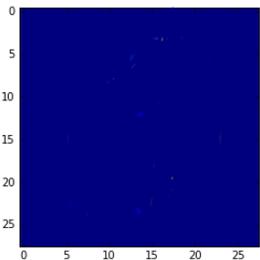
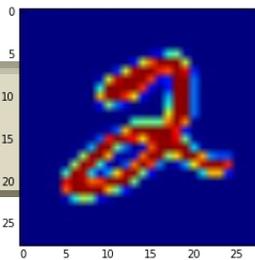
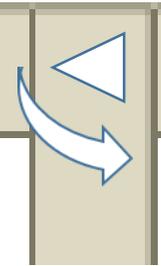
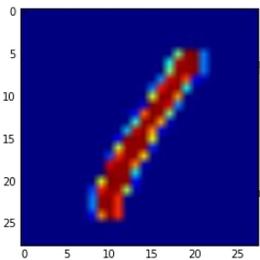
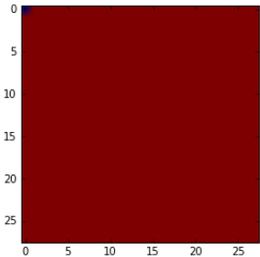


# Epistemic Value 3: 左向きから下向き

$$D[Q(s_t|o_t, a_{t-1}, o_{t-1}) || Q(s_t|a_{t-1}, o_{t-1})] = 3.319$$

$Q(o_t|s_t)$

$Q(o_t|s_t)$



# まとめ

- Active InferenceをANNで実装（ハッカソン）
- Epistemic ValueをANNで実装（追加実験）
  - 環境の中で観測する価値の高い場所で高い値を取ることを確認
- 今後の発展
  - Extrinsic Valueの実装
  - 記憶を持ったモデルによるEpistemic Valueの計算

質問用

# 学習の工夫

- アクションは3次元のone hotベクトル。そのまま使うと観測値(28\*28次元)と次元の差がありすぎて無視される  
→100倍にカサ増し(同じベクトルの反復)して使用