

第3回全脳アーキテクチャ・ ハッカソン 結果発表

小山内 琢也

神経科学的妥当性評価：実装したものに✓印を入れてください。

		✓			✓
海馬内活動	リプレイ		脳領域構造	CA1	
	プリプレイ			CA2	
	場所細胞			CA3	
	グリッド細胞			歯状回	
	頭部方向細胞			嗅内皮質	
	シータ位相歳差			海馬支脚	
	スパース表現			Perirhinal Cortex	
	パターン補完			Postrhinal Cortex	
	細胞新生		その他	コネクトームの導入	
行動機能	自律的フェーズ変化			BiCAMONでの可視化	
	エピソード記憶			その他	
	場所の再認				
	記憶転送				
	ナビゲーション/空間認知				
	Path integration				

規定課題点評価：成功・失敗エピソード数を記入してください。

課題番号	成功エピソード数	失敗エピソード数	合計エピソード数（成功＋失敗）
1－1			
1－2			
1－3			
1－4			
1－5			
1－6			
1－7			
1－8			
2－1			
2－2			
3－1			
3－2			
3－3			

課題 1 – 1

課題 1 – 2

課題 1 – 3

課題 1 – 4

課題 1 – 5

課題 1 – 6

課題 1 – 7

課題 1 – 8

見てわかる通り、**実装が間に合いませんでした。**

なので、やりたかったことを発表します。

やりたかったこと

- ・ 強化学習DQNとVAE (Variational Autoencoder)を組み合わせよう。

↓

- ・ VAEによるエンコードで求められた潜在変数 z を基にし、行動決定や次の状態の予測、自己位置推定などをしよう。

VAE(Variational Autoencoder)とは

簡単に言うと、入力 X から潜在変数 z (潜在要素)を求められる。
また逆に、潜在変数 z から X を生成することができる。

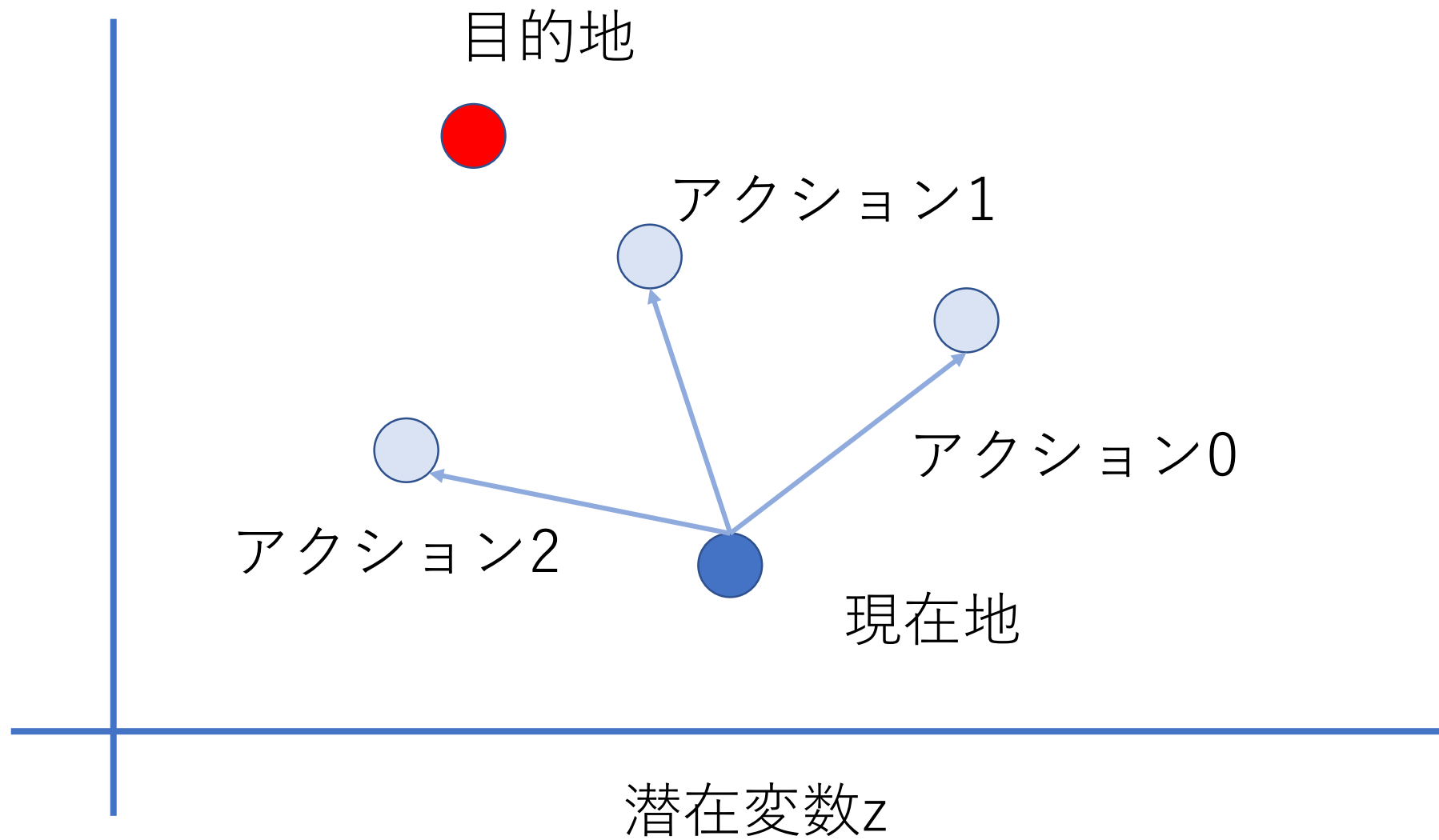


潜在変数 z から次の最適な行動を選択できるのではないか？

考え方

- ・ 迷路探索問題を潜在変数 z 間での移動と仮定する。
- ・ 各アクションごとの移動距離を計算しておき、次の状態の予測に使用する。
- ・ プラスの報酬が得られた時の潜在変数 z を記録しておき、それに近づく行動を選択する。

イメージ



実装方法

- ・現在の状態を潜在変数 z にエンコード。
- ・プラスの報酬を得ていた場合、その状態の潜在変数 z を記録。
- ・アクション(右を向く、左を向く、前進する)ごとの潜在変数 z の移動距離を記録。
- ・アクションごとの潜在変数 z の移動距離の平均から各アクション後の潜在変数 z の位置を計算。
- ・各アクション後の潜在変数 z の位置と、プラス報酬時の潜在変数 z の位置とで、一番近くなるアクションを選択する。

懸念点

- ・ プラス報酬時の潜在変数 z を記録し予測に利用するため、学習始めはどうしても安定しない。

そもそも、

- ・ プラス報酬時の潜在変数 z の位置は一定の範囲に収まる
- ・ 各アクション後の移動距離には一定の法則がある

は正しいのか？

→ 上記想定が否定される場合、この方法では目的地に近づくことができない。

学習始めに安定しないことに対する対策案

- ・ 学習始めは現在の状態と次の状態の距離に応じたプラスの内部報酬を与えるようにする。※離れているほどプラス
→様々な場所を探索し、経験を蓄積する。
- ・ ある程度のプラス報酬時の潜在変数 z が記録されたら、それに近づくようアクションを選択するようにする。

想定が違っていた場合に対する対策案

- ・各アクション後の位置の予測にもVAEを使用すればよいのではないか？

→ VAEの最適化には、元の入力 X_t と潜在変数 z から生成した予測 X_t' の誤差を利用しており、元の入力 X_t を次の入力 X_{t+1} にすれば、次の状態を予測できるのではないか？



上手くいくかどうかは、やってみないとわからない……

DQN+VAEによる今回の方法が上手くいった場合、

- ・自己位置推定→現在の潜在変数 z の位置
- ・行動選択→プラス報酬時の潜在変数 z への最短移動距離の計算
- ・次の状態を予測→潜在変数 z での移動

と置き換えることができる。

今後の展望

- ・ プログラムを完成させ、想定の正否を判断する。
- ・ VAEのチューニングや別理論(GAN)での実装も試してみる。
- ・ プラス報酬時の潜在変数 z だけではなく、極端なマイナス報酬時の潜在変数 z も記録しておき、マイナス報酬時の潜在変数 z から離れるよう行動選択するよう学習させる。
- ・ 潜在変数 z をRNNに入力して学習させてみたい。
- ・ 視覚からの入力画像に対し、ざっくりクラスタリングして、画像の中でメインとなるものとそうでないもので、CNNへの入力の重みを変更するよう、機械学習させてみたい。

感想

- ・実装が間に合わなかったのが悔しい・恥ずかしい。
→今回、Keras+TensorFlowでVAEを実装しようとしたが、
tf.Graphがスレッドセーフではないため、学習処理がエラーで落ちる。
→実装の経験や知識の不足が原因のため、さらなる勉強が必要。
- ・数式からプログラムに落とし込む作業が難しい。
→データの形・次元を意識しながらやるようにする。慣れるしかない。
- ・今回のプログラムを完成させること。

余談

今回使用したLISとBriCaのコードに対し、Chainer1.24、Python3対応をしました。

Chainer2化は、Chainer1からの変更点が大きかったため、今回はやりませんでした。(確かCoffeモデルpickleの読み込みに失敗したような)

→どうやって連携すればよいですか？

ご清聴ありがとうございました