

preplayを 利用した探索の補助

石井・鈴ヶ嶺

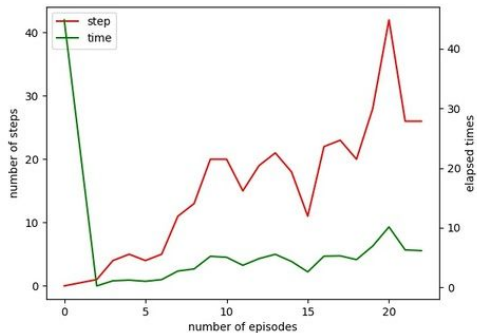
神経科学的妥当性評価：実装したものに✓印を入れてください。

		✓			✓
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				

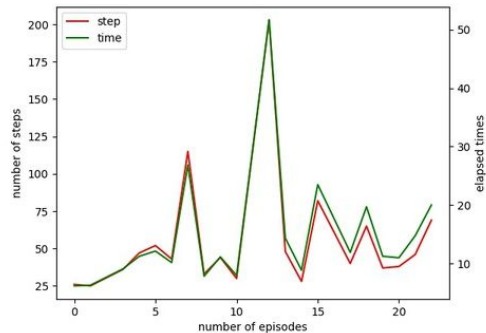
規定課題点評価：成功・失敗エピソード数を記入してください。

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	22	0	22
1－2	22	0	22
1－3	762	1029	1791
1－4			
1－5			
1－6			
1－7			
1－8			
2－1			
2－2			
3－1			
3－2			
3－3			

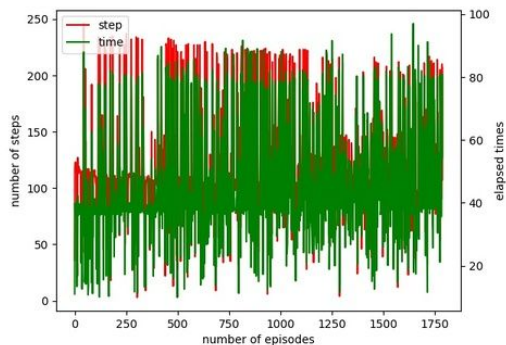
課題1-1



課題1-2



課題1-3



課題1-4

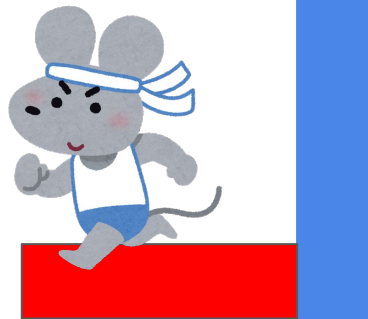
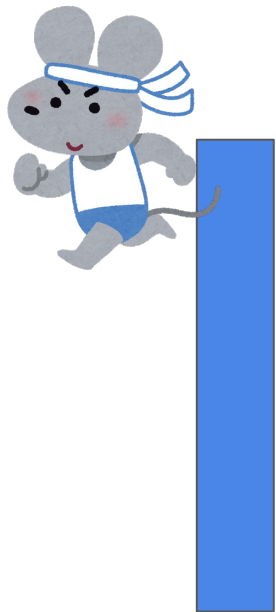
* episode_result.logのplotを張り付けてください。

preplayとは

George Dragoi, Susumu Tonegawa. Preplay of future place cell sequence by hippocampal cellular assemblies. Nature. 2011;469:397-401

休息中や睡眠中に、経験に先立って、
場所細胞の発火が起こる現象

マウスを使った実験



マウスを使った実験

	青い経路を 走っているときに発 火した細胞	青い経路上で 休んでいるときに発 火した細胞
赤い経路を 走行中に発火 している細胞	相関 なし	相関 大

preplayによって、迷路探索の際の
情報収集の基盤ができるのではないか

実装

preplayの実現においては、
事前に予測するための情報を
いかに取得するかがポイントだと考えた

処理プロセス

- experience replay
- preplay
 - experience replayと同様、ランダムサンプリングで更新する

preplayの更新

$$Q(s_t, a) \leftarrow Q(s_t, a) + \alpha \left[\underline{r_{t+1}} + \gamma \max_{a'} (\overline{s_{t+1}}, a') - Q(s_t, a) \right]$$

- 時刻t+1におけるrewardに代わって、時刻tにおけるrewardを用いている
- t+1における状態sは、予測した情報を利用した

方針1 過去の経験から類似ベクトルを選ぶ

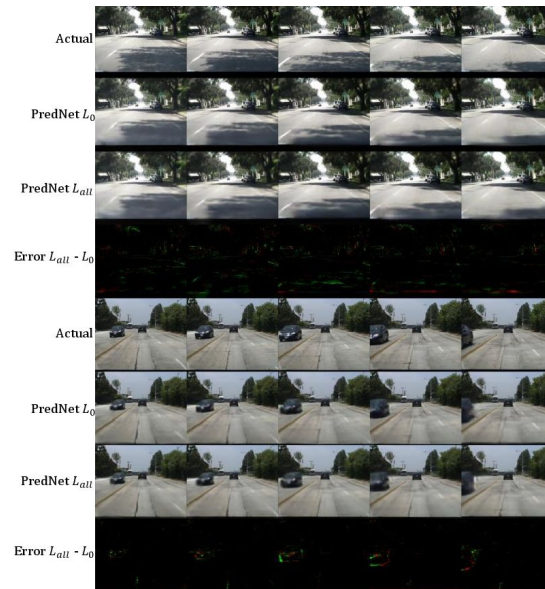
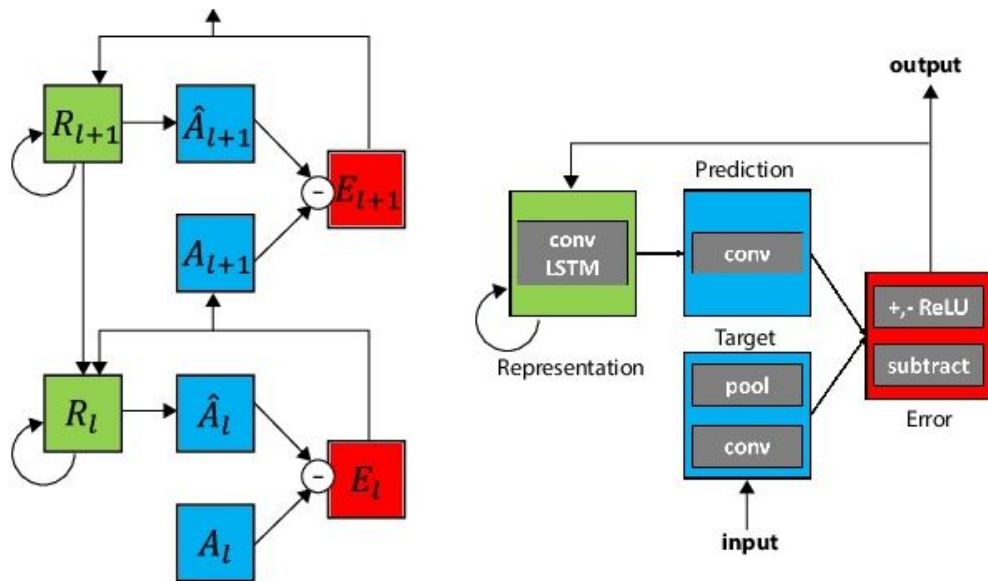
- ベクトル距離の近いもの
- コサイン類似度による判定

方針1 過去の経験から類似ベクトルを選ぶ

$$Q(s_t, \underline{a}) \leftarrow Q(s_t, \underline{a}) + \alpha \left[r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, \underline{a}) \right]$$

- a : 以前に選んだアクション以外のものからランダムに選択

方針2 PredNetを利用する



William Lotter, Gabriel Kreiman & David Cox. DEEP PREDICTIVE CODING NETWORKS FOR VIDEO PREDICTION AND UNSUPERVISED LEARNING. Harvard University. 2017

方針2 PredNetを利用する

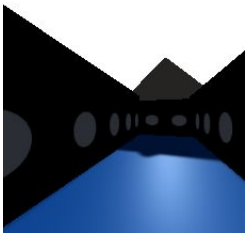

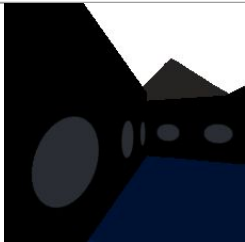
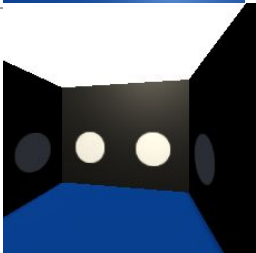
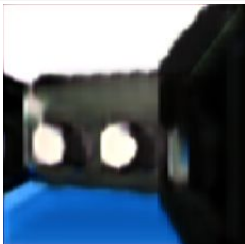
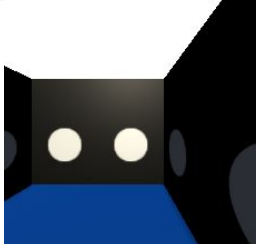
- RGBとdepthのそれぞれの画像について、前進、右回転、左回転の計6種類の予測を行う必要がある
- 事前に、学習を行ったものを使用した。

結果

方針1 過去の経験から類似ベクトルを選ぶ

- あまり良い結果は得られなかった
 - 得られた類似したものに対する行動をランダムに選択していたため
 - ランダムな選択は、予測とは程遠い

方針2 PredNetを利用する

	入力	予測	真の結果
前進			
右回転			

方針2 PredNetを利用する

- いずれのアクションについても、画像の予測は難しかった
 - 進むときの歩幅(画像の変化量)が大きく、
画像の予測がより難しくなったため
- 正面に進むときの画像の予測については、特にうまくいかなかった
 - 正面が壁の場合と空間が開けている場合とで、画像の予測が変わってくるため
 - 正面が壁なのか空間なのかの情報が必要だった。
画像のみでは、突き当たるまでの距離の情報が無く、
その判定は難しい

方針2 PredNetを利用する

- 改善するには？
 - エージェントの移動量をより小さく設定する
 - 得られる画像は、より連続的なものになる
(車載カメラの例により近づく)
 - より細かな移動量で得られた画像を用いて
PredNetの予測を繰り返せば、より精度が上がったのではないか

まとめ

- 我々はpreplayのモデル化に関する検討を行った
- preplayという現象が探索において基盤となる活動になりうると推測する