

複数人のモデリングのための部分観測と力学的制約を伴う分散型方策学習

藤井 慶輔^{1,2}・武石 直也^{3,2}・河原 吉伸^{4,2}・武田 一哉¹

¹名古屋大学, ²理研AIP, ³西スイス応用科学大学, ⁴九州大学

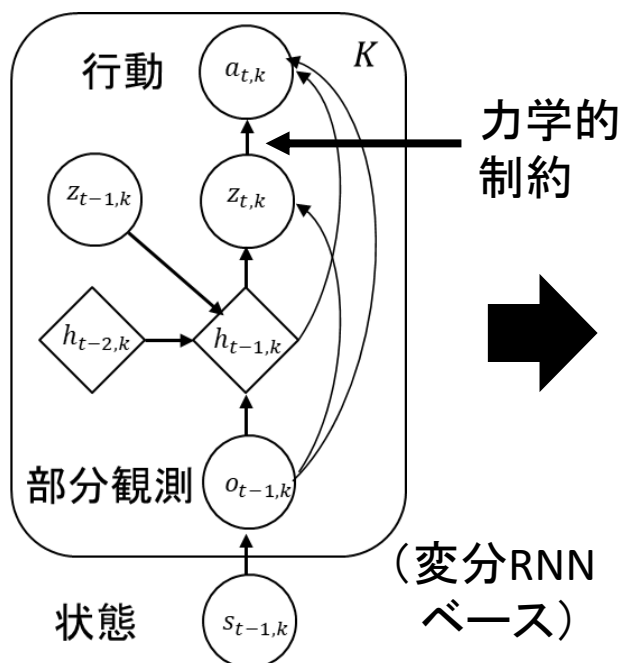
IBIS2020

発表番号: 55

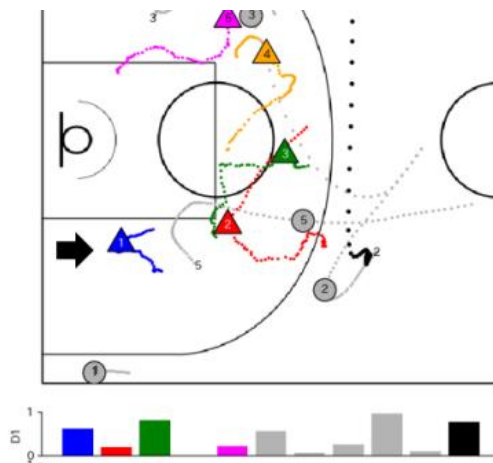
目的 生物学的制約を考慮した長期予測・操作可能な集団運動のモデル化

手法 部分観測過程と力学的制約を導入した分散型模倣学習モデルを提案

結果 集団スポーツにて、正確な長期予測と観測を操作した反事実的予測が可能

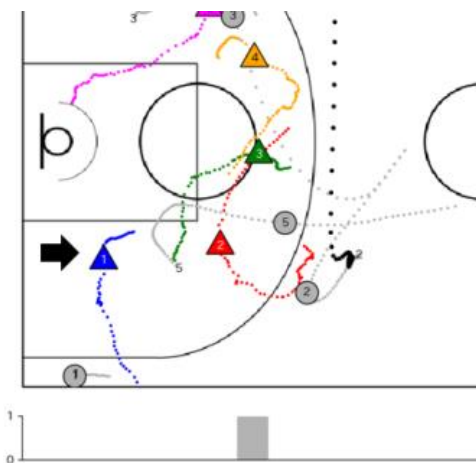


正確な軌道予測 (6s)



部分観測をモデル化

観測を1つにした反事実予測



観測した1人を追いかける

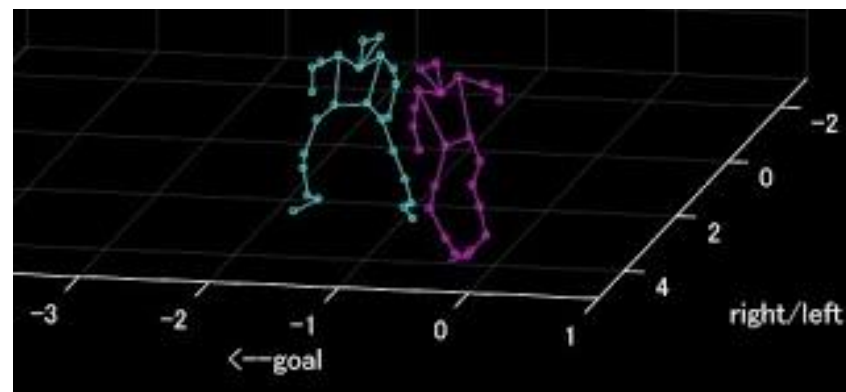
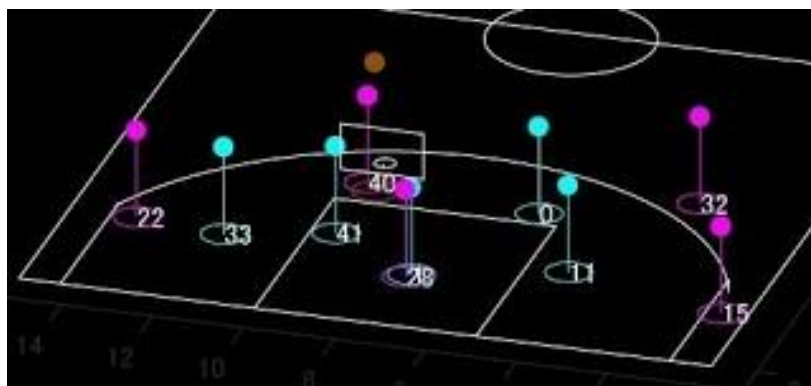
複数人のモデリングのための部分観測と 力学的制約を伴う分散型方策学習

藤井 慶輔^{1,2}・武石 直也^{3,2}・河原 吉伸^{4,2}・武田 一哉¹

¹名古屋大学, ²理研AIP, ³西スイス応用科学大学, ⁴九州大学

Paper : <https://arxiv.org/abs/2007.03155>

Codes: <https://github.com/keisuke198619/PO-MC-DHVRNN>



部分観測過程における生物集団のモデリング

生物の視野等に基づき、部分観測は古くからモデル化

(例: Aoki, 1982)

数学的には部分観測マルコフ決定過程(POMDP)として定式化

AR (angular range of interaction)

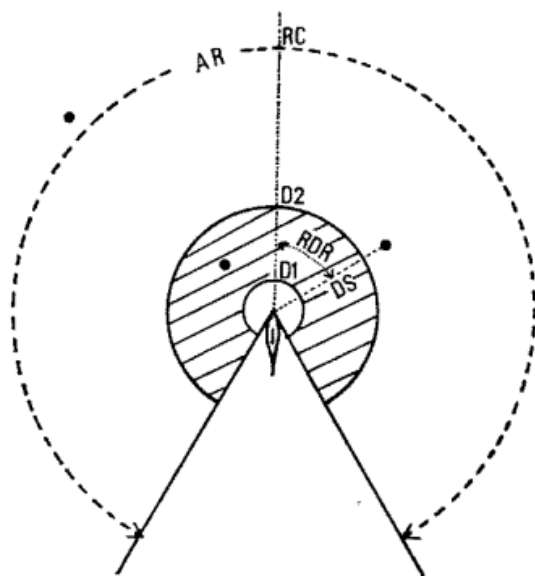
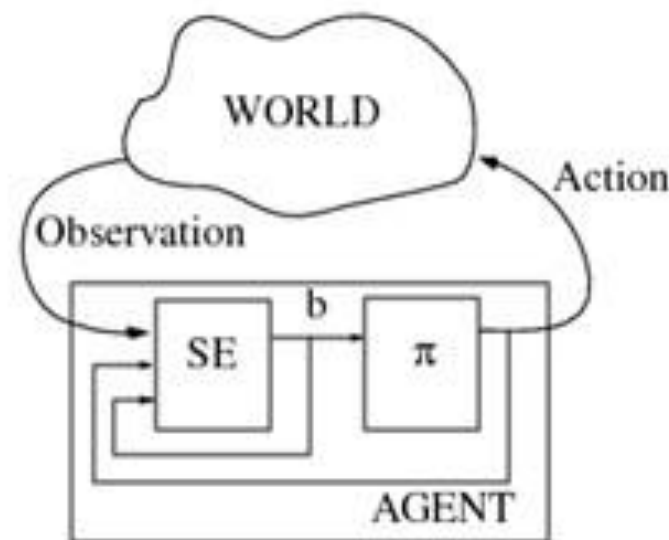


Fig. 1. Geometrical drawing to illustrate parameters specifying interactions. [Aoki, 1982]

POMDP



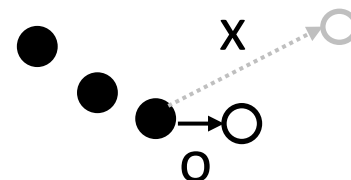
SE: State Estimator
b: belief state

[Kaelbling+, 1998]

これまでの生物集団のデータ駆動モデリング(いずれか)

1. 外部環境の情報をすべて用いる or ルールベースで観測
2. 中央制御的・最適なコミュニケーションを仮定[Zhan+19; Yeh+19など]
3. 生物の移動に関わる力学的制約を無視

(詳細は論文<https://arxiv.org/abs/2007.03155>を参照)



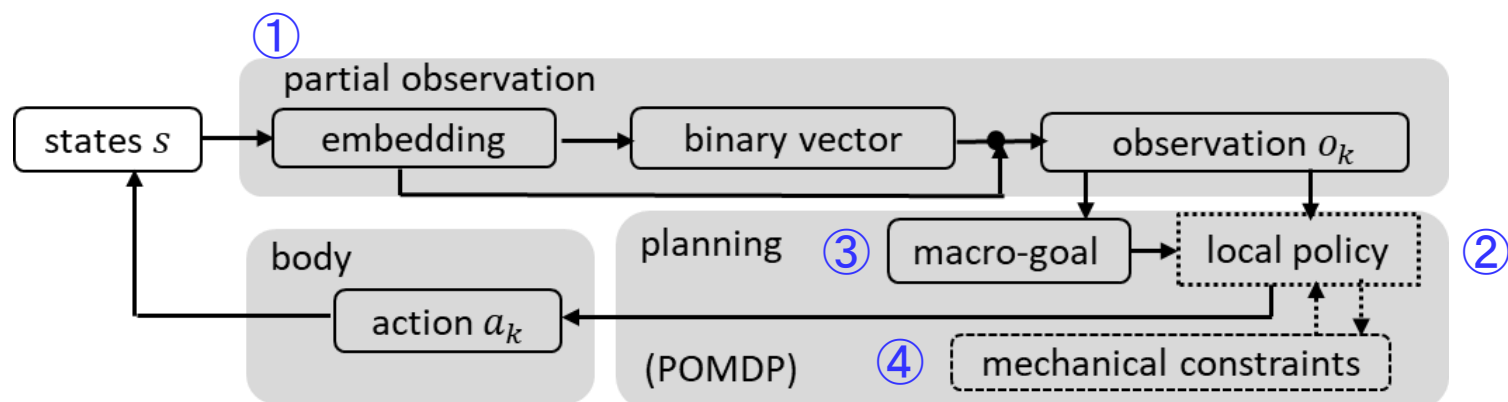
生物集団の特徴:

1. 限られた認知資源の中で、**情報(観測)を柔軟に取捨選択**
2. 不均一な特性の個体が**自律分散的に**振舞う
3. **物理則・運動制御則**から外れない(例: 加速度は滑らか)

目的: 上記の要件を満たすデータ駆動モデルを構築したい

部分観測と力学的制約を考慮した分散型模倣学習モデル[Fujii+20]

全体像:



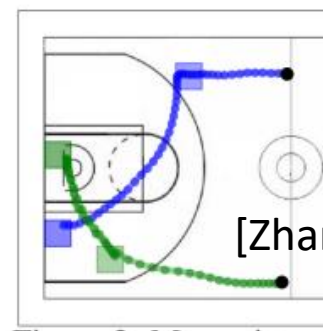
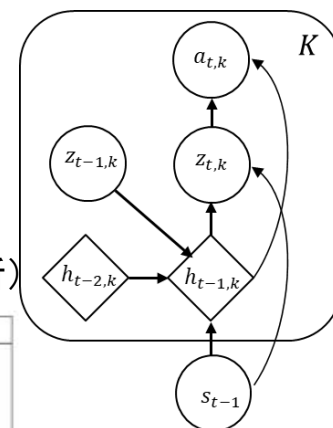
構成要素:

- ① 部分観測モデル: multi-hot表現を学習
- ② 局所的政策: 変分RNN [Chung+15]
- ③ 大域的目標: 部分観測・分散型の弱教師情報(例: 止まりそうな場所)
- ④ 力学的制約
 - a. 位置・速度・加速度の整合性(物理)に関する罰則
 - b. 加速度の滑らかさ(生体力学)に関する罰則

Multi-hot 表現



変分RNN



[Zhan+19]

モデルの詳細と実験結果の前に(前提: 実際の問題)

データセット:

- サッカー: ヨーロッパリーグ45試合、10 Hz
 - <https://www.statsperform.com/>に連絡を取って入手
- バスケット: NBA100試合、25 Hz (→10 Hzにダウンサプリング)
 - <https://github.com/rajshah4/BasketballData>から

誰をどこまで予測?

- 守備チーム全員(10人/5人)を予測
- 2秒間の軌道を使って6秒間の軌道を予測

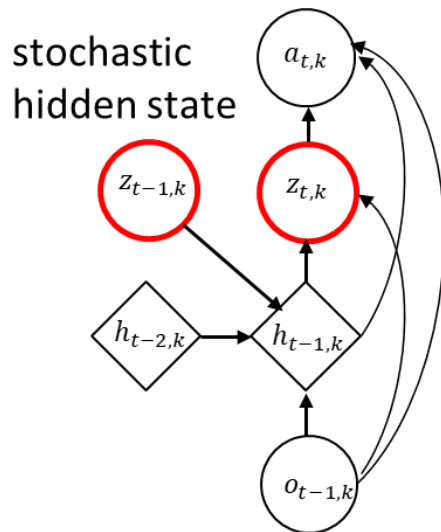
何から何を予測? ※各選手に1つのモデル(例: RNN)を割り当てる

- 入力: 全選手の位置・速度・加速度
- 出力: 各選手の速度と加速度(検証は[Fujii+20])

モデルの詳細と実験結果②③ 分散型方策

変分 RNN

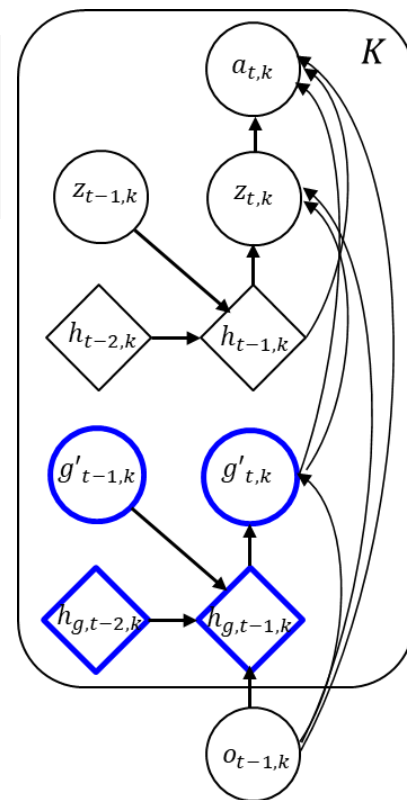
[Chung+15, NIPS]



大域的目標を用いた 階層型変分RNN

Decentralized version
of [Zhan+19, ICLR]

macro-goals



Root mean squared error (6-s prediction)

	Basketball data			Soccer data		
	position	velocity	acceleration	position	velocity	acceleration
Velocity	1.41 ± 0.34	1.08 ± 0.21	10.90 ± 2.09	4.83 ± 1.27	2.72 ± 0.46	27.22 ± 4.56
RNN-Gauss	1.31 ± 0.32	1.05 ± 0.13	1.88 ± 0.30	2.97 ± 0.89	1.65 ± 0.23	2.22 ± 0.33
VRNN	0.71 ± 0.17	0.68 ± 0.10	1.43 ± 0.20	1.23 ± 0.35	1.06 ± 0.21	1.57 ± 0.26
VRNN-macro	0.71 ± 0.17	0.68 ± 0.10	1.43 ± 0.20	1.37 ± 0.40	1.11 ± 0.21	1.57 ± 0.24
VRNN-Mech	0.69 ± 0.17	0.68 ± 0.10	1.37 ± 0.20	1.22 ± 0.34	1.05 ± 0.20	1.59 ± 0.26
VRNN-Bi	0.72 ± 0.19	0.66 ± 0.10	1.36 ± 0.19	1.20 ± 0.37	1.02 ± 0.20	1.50 ± 0.24
VRNN-macro-Bi-Mech	0.73 ± 0.18	0.68 ± 0.10	1.34 ± 0.19	1.31 ± 0.42	1.06 ± 0.21	1.47 ± 0.23

モデルの詳細と実験結果④ 力学的制約

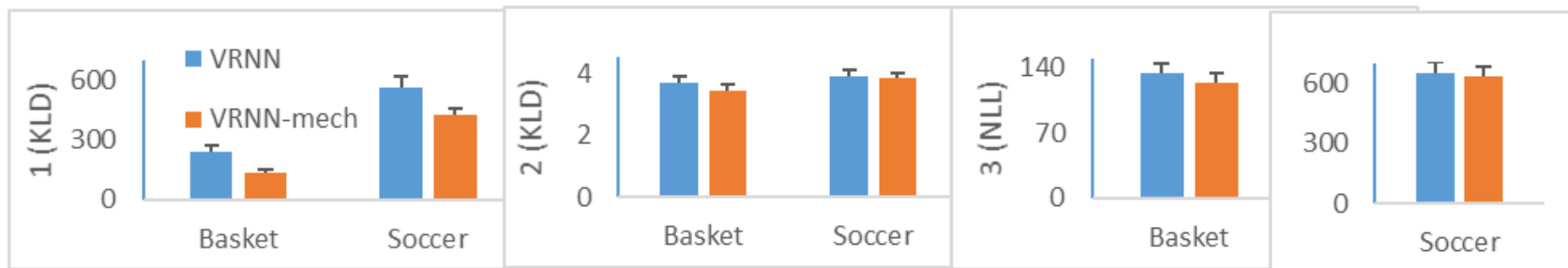
NNの学習時に、制約損失(罰則)を加えた

- 提案モデルは、速度と加速度を直接予測し、位置は間接的に予測 [Fujii+20]

$$\mathcal{L}_{mech} = \mathbb{E}_{\theta} \sum_{t=2}^T \left[\underbrace{\lambda_{acc} D_{KL} (p_{\theta}(\hat{a}_{acc,t} \mid z_{\leq t}, o_{< t}) \parallel p_{\theta}(\tilde{a}_{acc,t} \mid z_{\leq t}, o_{< t}))}_1 - \underbrace{\lambda_{pos} \log p_{\theta}(a_{pos,t} \mid z_{\leq t}, o_{< t})}_2 - \underbrace{\lambda_{jrk} \log p_{\theta}(a_{acc,t+1} \mid z_{\leq t}, o_{< t})}_3 \right].$$

- 直接予測した分布と間接的に予測した分布の差 (KLD: Kullback–Leibler divergence)
(例: $\hat{a}_{acc,t}$ と $\tilde{a}_{acc,t} = (\hat{a}_{vel,t} - \hat{a}_{vel,t-1})/\Delta t$)
- 間接的に予測した位置の分布 $\tilde{a}_{pos,t}$ と計測された $a_{pos,t}$ の差 (NLL: 負の対数尤度)
- 直接予測した加速度の分布 $\hat{a}_{acc,t}$ と次の時刻の加速度 $a_{acc,t+1}$ の差 (NLL)
(これは滑らかな運動を達成する躍度最小原理 [Flash+85,JNS] を参考)

better



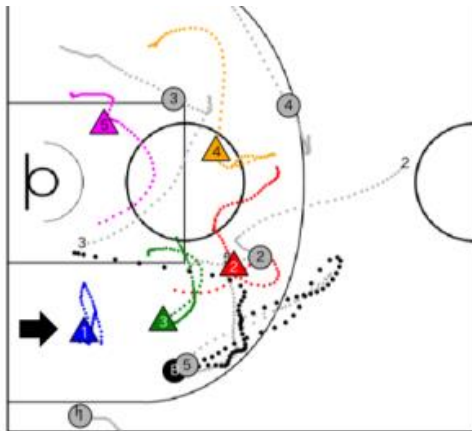
実験結果① 部分観測

観測の係数(平均の観測数) :

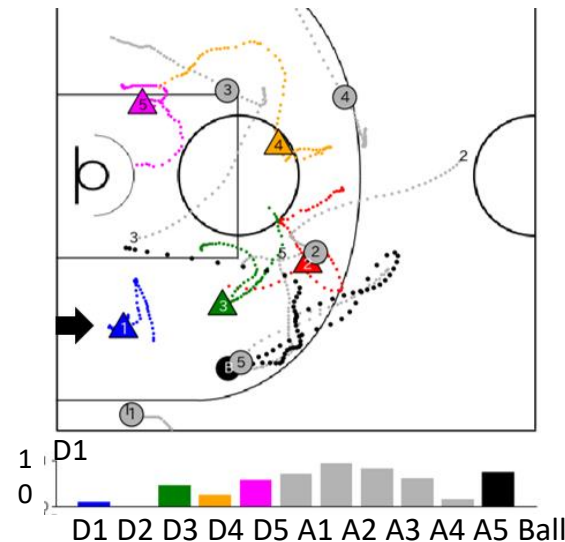
- ✓ Basketball: 4.42 ± 0.31 for each def. (max.: 11)
- ✓ Soccer: 8.04 ± 1.54 for each def. (max.: 23)

典型例と反事実的な予測:

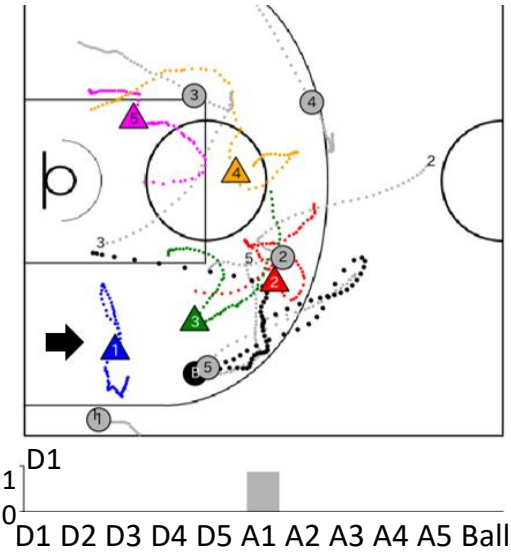
Ground truth



Accurate prediction (6 s)



Counterfactual prediction



結論

- 部分観測と力学的制約を考慮した分散型模倣学習モデルを提案した
- 生物らしい軌道の長期予測の検証と、観測の可視化や操作を行い、反事実予測による検証を行った

謝辞:

JSPS KAKENHI (Grant Numbers 19H04941 and 20H04075)

Paper : <https://arxiv.org/abs/2007.03155>

Codes: <https://github.com/keisuke198619/PO-MC-DHVRNN>