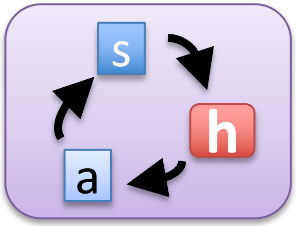
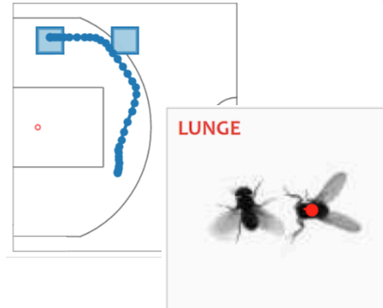
Inference + Imitation

Yisong Yue





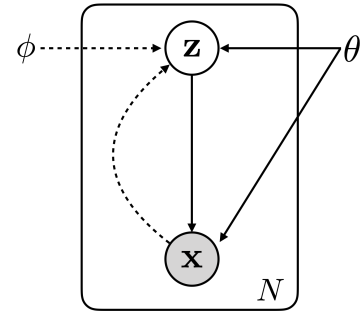
$$\mathrm{argmin}_{\boldsymbol{\theta}}\ \mathrm{E}_{s\sim P(s|\boldsymbol{\theta})} L(\pi^*(s), \pi_{\boldsymbol{\theta}}(s))$$

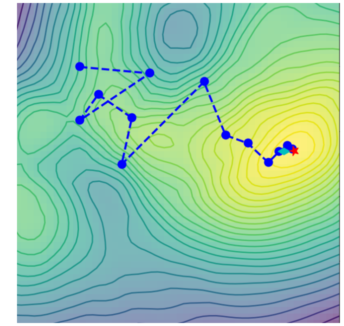
$$\max_{\pi\in\Pi} \min_{r\in\mathcal{R}} \mathbb{E}_{\pi}\left[r(s,a)\right] - \mathbb{E}_{\pi^*}\left[r(s,a)\right]$$

Imitation Learning

Optimize desired behavior

Learn from demonstrations





$$\mathbb{E}_q\left[\log p_\theta(\mathbf{x}|\mathbf{z})\right] - D_{KL}(q(\mathbf{z}|\mathbf{x};\lambda)||p_\theta(\mathbf{z}))$$

(Variational) Inference

Inference in probabilistic models

Phrased as optimization

← Probabilistic Imitation Learning

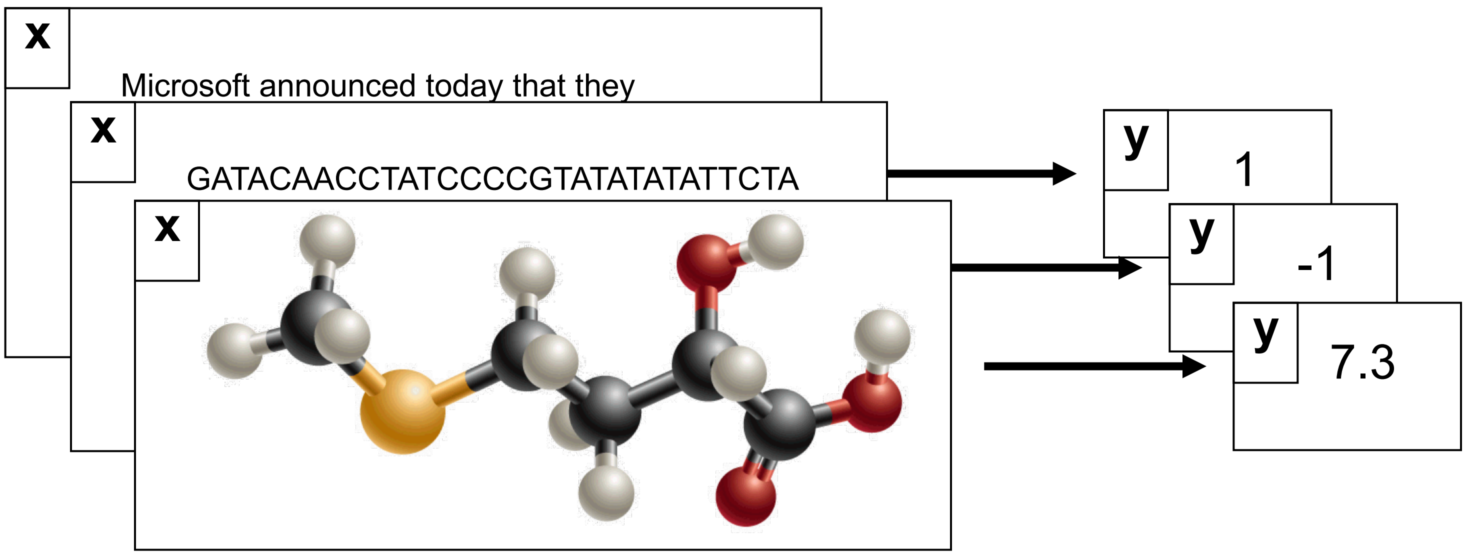
→ Learning to Infer

Warm Up: Supervised Learning

- Find function from input space X to output space Y

$$h : X \longrightarrow Y$$

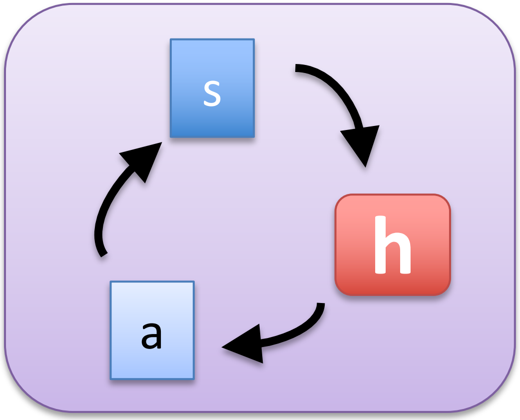
such that the prediction error is low **



** error can also be probabilistic (e.g., log likelihood)

Imitation Learning

- Input:
  - Sequence of contexts/states:

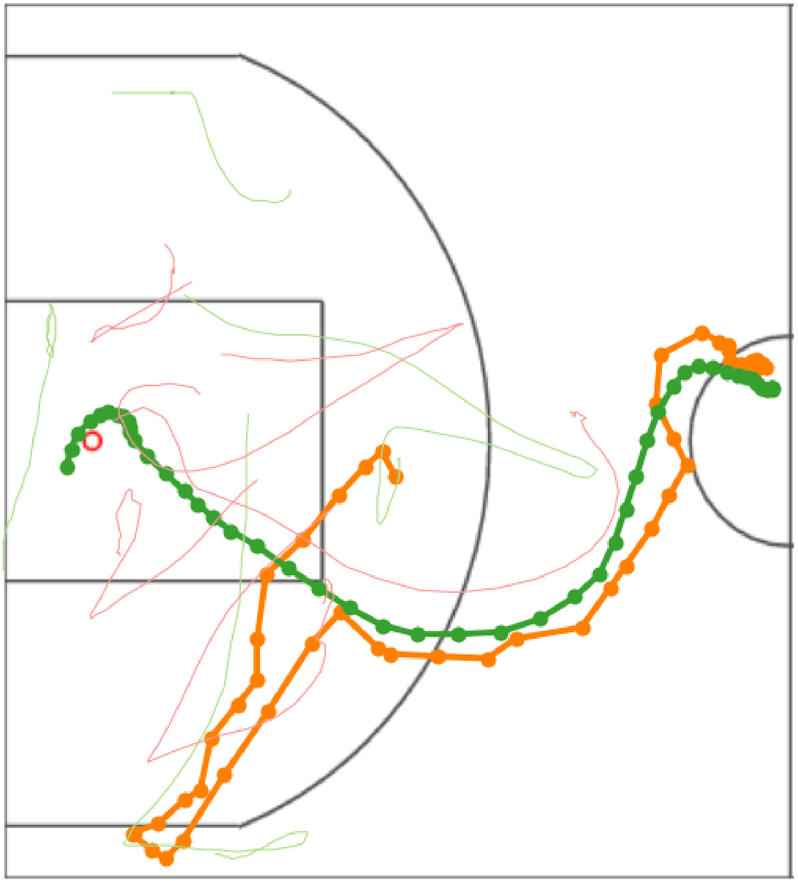


- Predict:
  - Sequence of actions
- Learn Using:
  - Sequences of demonstrated actions

Example: Basketball Player Trajectories

- s = location of players & ball
- a = next location of player

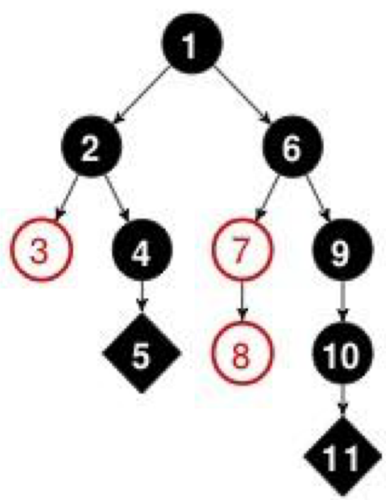
- **Goal:** learn $h(s) \rightarrow a$

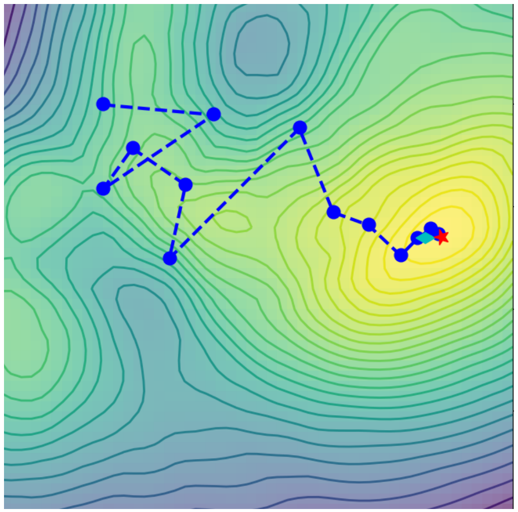


Example: Learning to Optimize

- s = optimization problem & current location
- a = next location

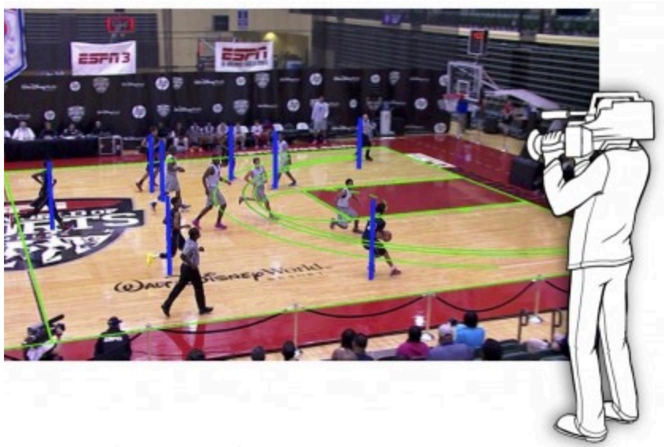
- **Goal:** learn $h(s) \rightarrow a$





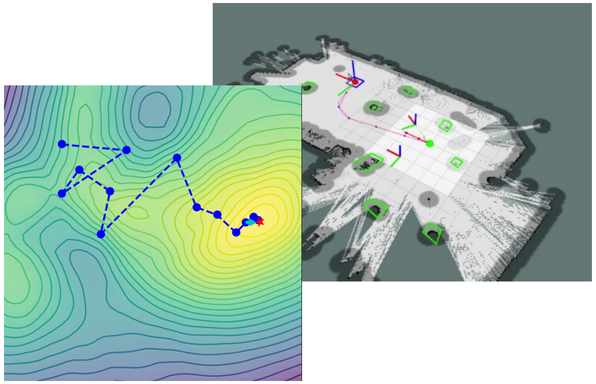
What to Imitate?

Human Demonstrations



Animal Demonstrations

Computational Oracle



Latent Variable Models

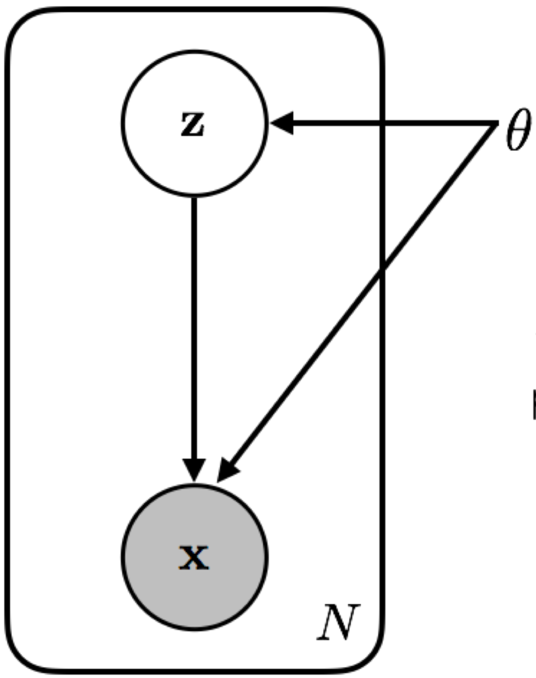
(Segue to Variational Inference)

$p_\theta(\mathbf{z})$

prior

$p_\theta(\mathbf{x}|\mathbf{z})$

conditional likelihood



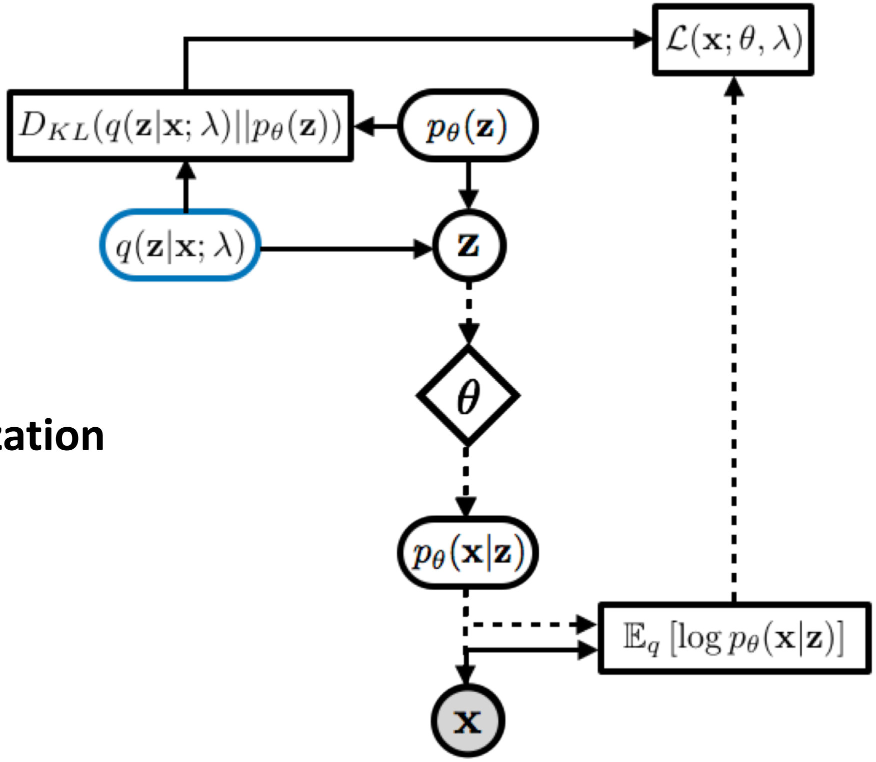
$$p_\theta(\mathbf{z}|\mathbf{x}) = \frac{p_\theta(\mathbf{x}|\mathbf{z})p_\theta(\mathbf{z})}{p_\theta(\mathbf{x})}$$

posterior

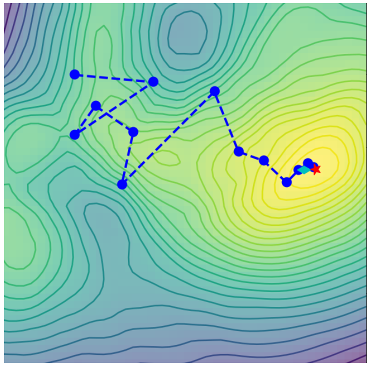
Variational Inference

approximate posterior $q(\mathbf{z}|\mathbf{x};\lambda)$

ELBO $\mathcal{L}(\mathbf{x};\theta,\lambda) = \mathbb{E}_q\left[\log p_\theta(\mathbf{x}|\mathbf{z})\right] - D_{KL}(q(\mathbf{z}|\mathbf{x};\lambda)||p_\theta(\mathbf{z}))$



Inference = Optimization

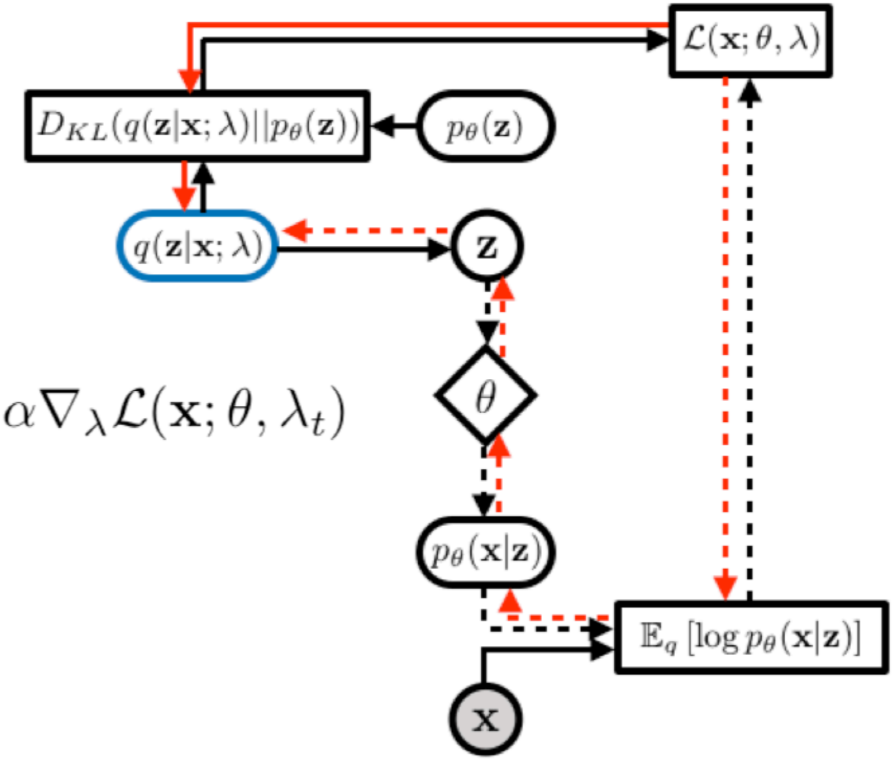


Stochastic Variational Inference

approximate posterior $q(\mathbf{z}|\mathbf{x};\lambda)$

ELBO $\mathcal{L}(\mathbf{x};\theta,\lambda) = \mathbb{E}_q\left[\log p_\theta(\mathbf{x}|\mathbf{z})\right] - D_{KL}(q(\mathbf{z}|\mathbf{x};\lambda)||p_\theta(\mathbf{z}))$

$\lambda_{t+1} \leftarrow \lambda_t + \alpha\nabla_\lambda\mathcal{L}(\mathbf{x};\theta,\lambda_t)$

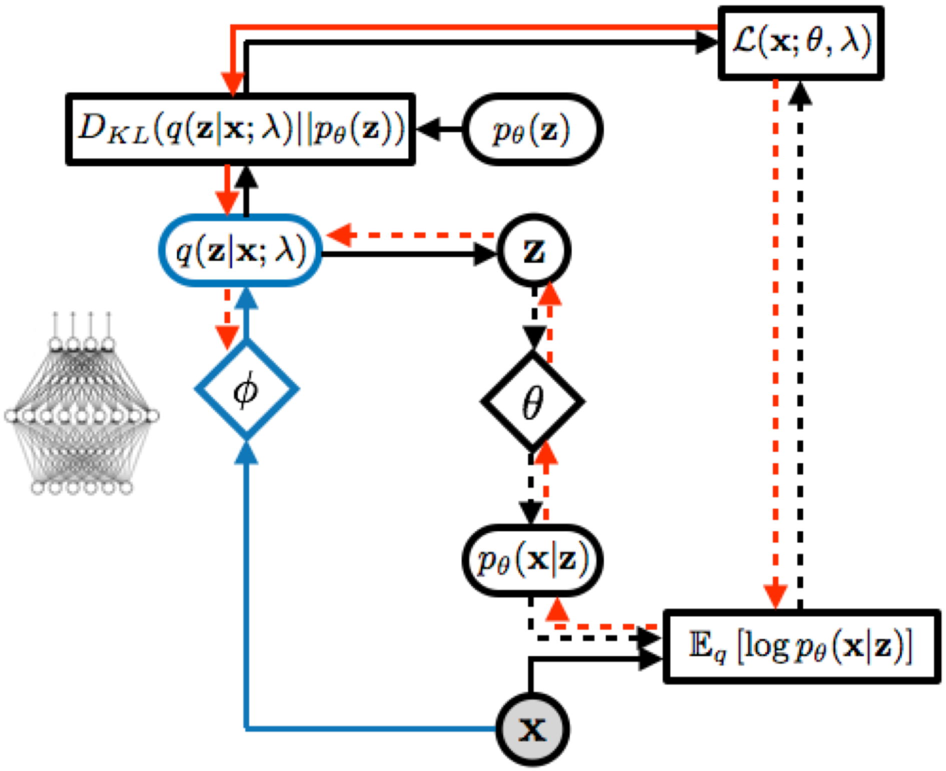


E.g., Hoffman et al., 2013

Amortized Variational Inference

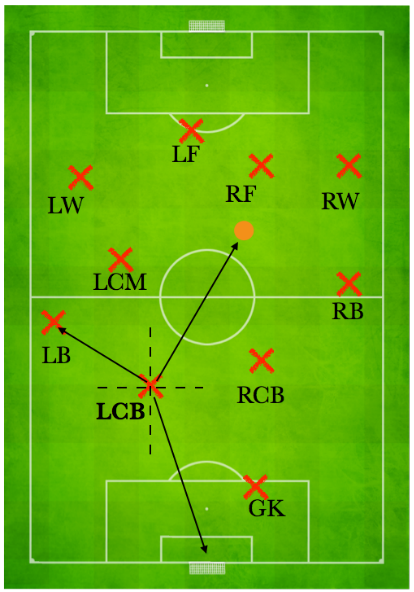
approximate posterior $q(\mathbf{z}|\mathbf{x};\lambda)$

ELBO $\mathcal{L}(\mathbf{x};\theta,\lambda) = \mathbb{E}_q\left[\log p_\theta(\mathbf{x}|\mathbf{z})\right] - D_{KL}(q(\mathbf{z}|\mathbf{x};\lambda)||p_\theta(\mathbf{z}))$



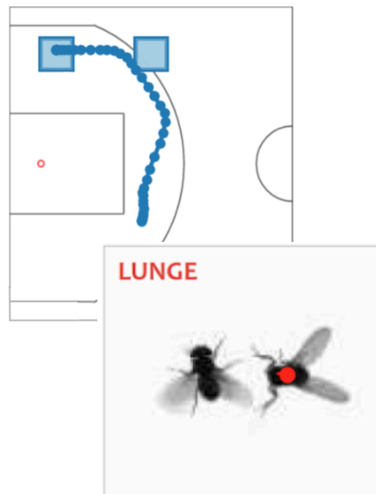
E.g., VAEs [Rezende et al., 2014] [Kingma & Welling, 2014]

Outline For Today



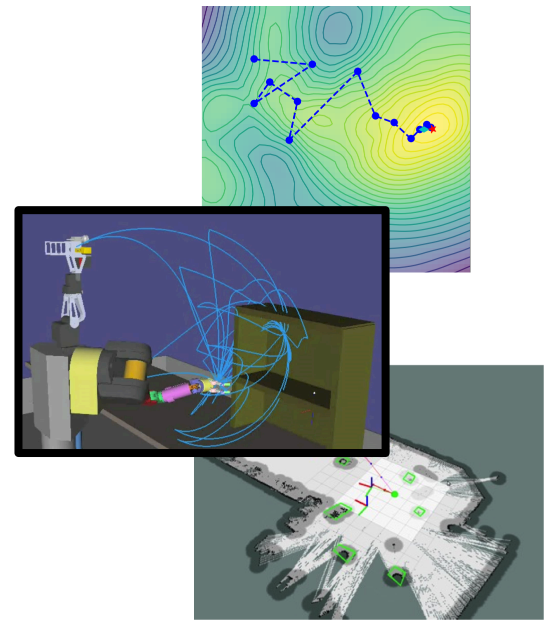
Coordinated Learning

Infer Latent Roles



Hierarchical Behaviors

Generative Behavior



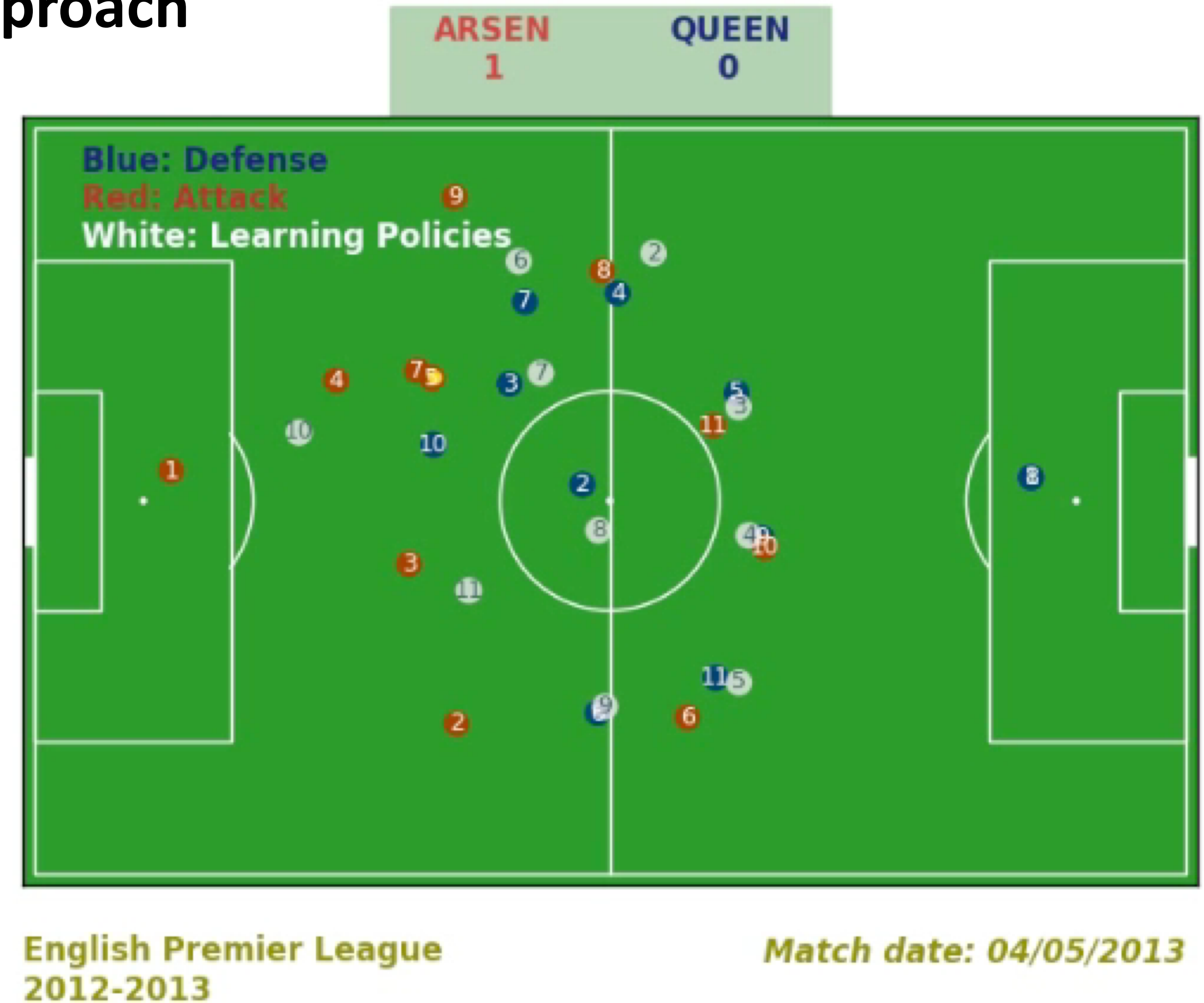
Learning to Optimize

Learn to Infer

← **Probabilistic Imitation Learning**

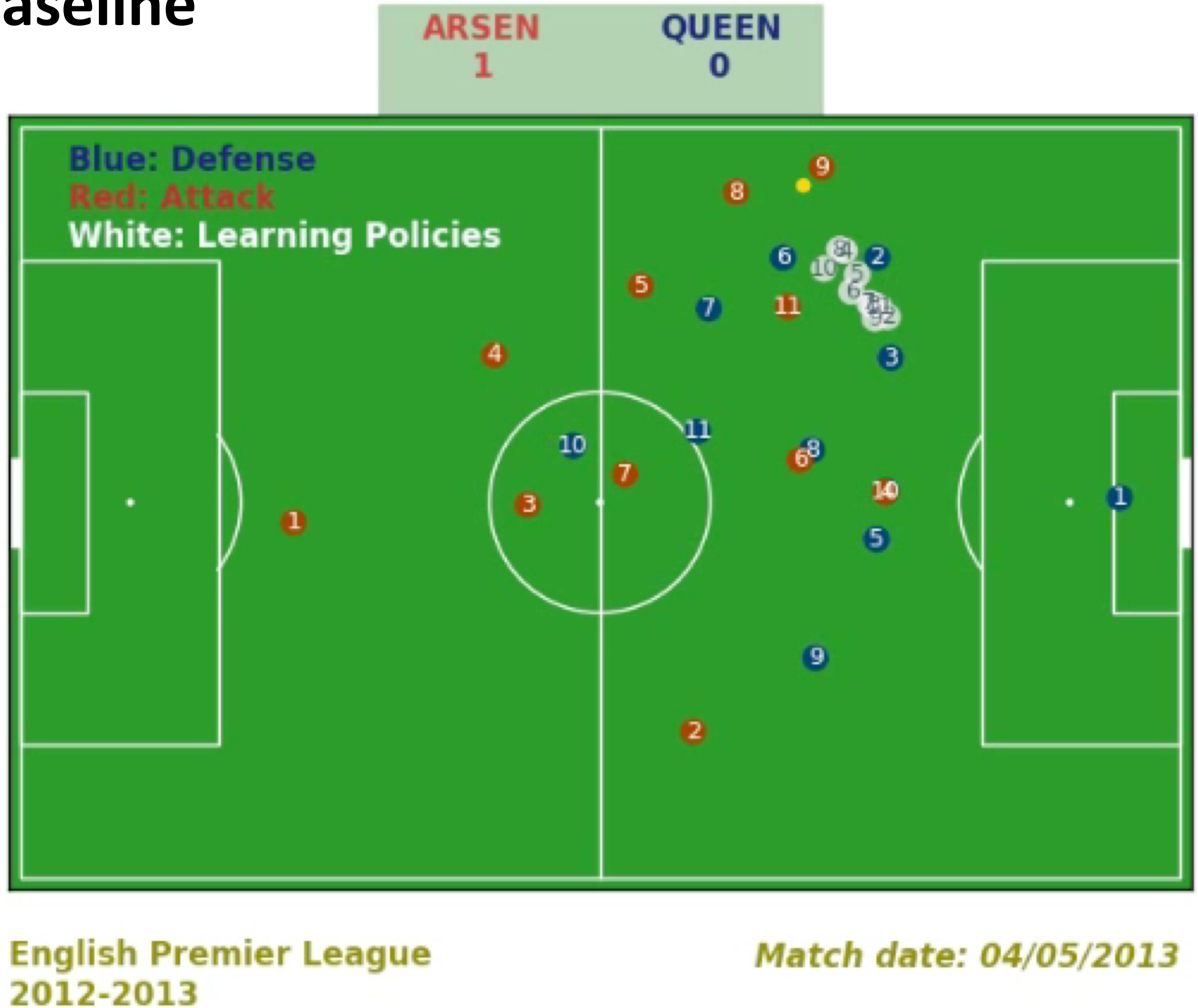
→ **Learning to Infer**

Our Approach

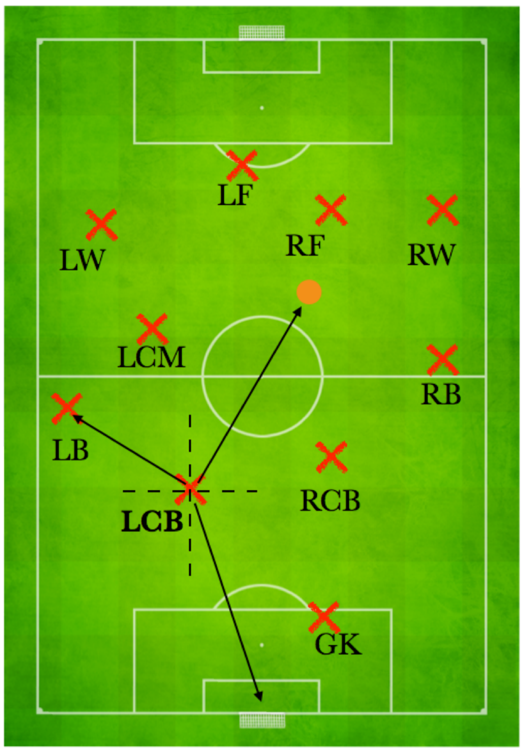


Data-Driven Ghosting using Deep Imitation Learning
Hoang Le, Peter Carr, Yisong Yue, Patrick Lucey. SSAC 2017

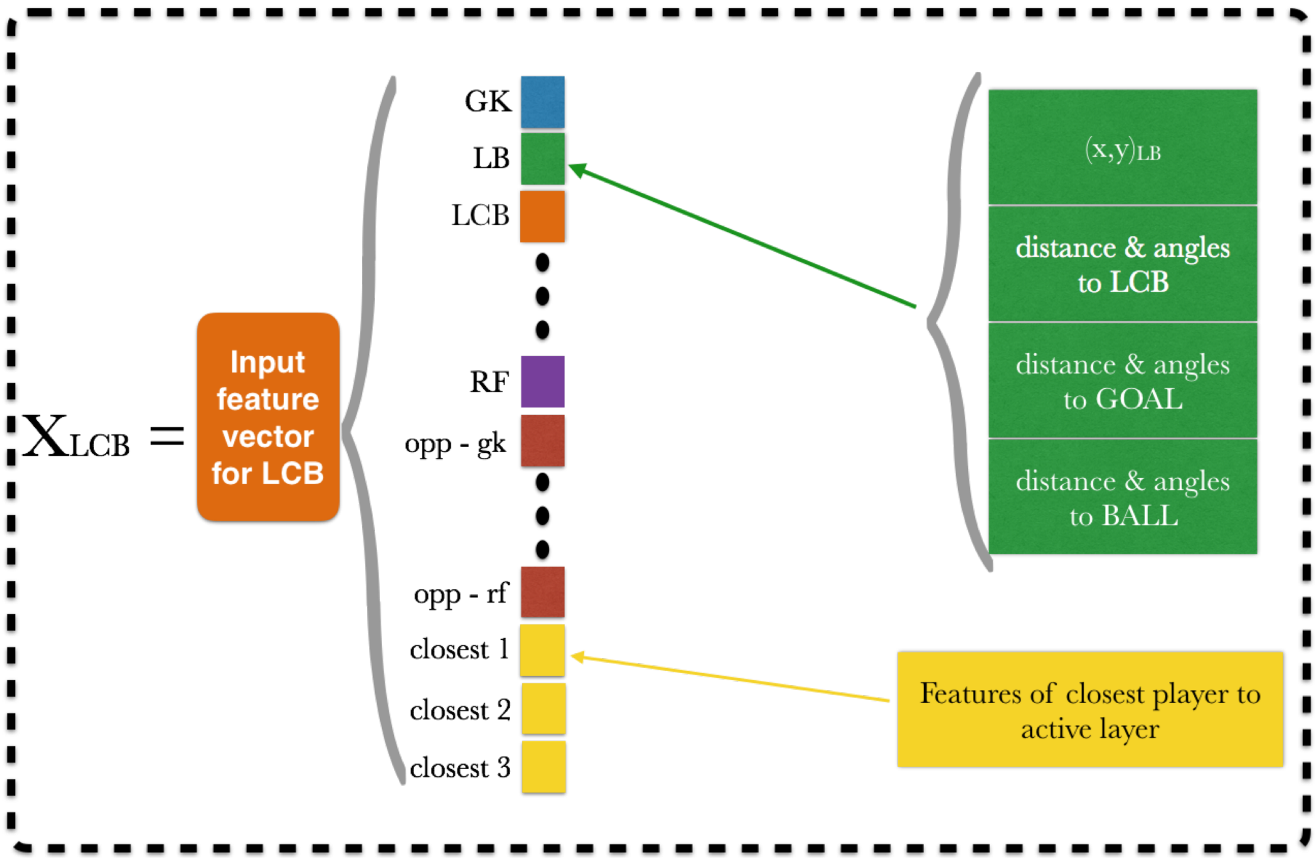
Naïve Baseline



State Representation



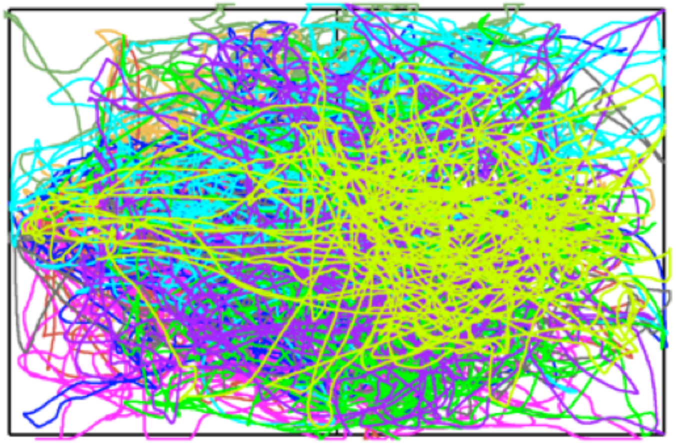
Geometric features computed



Data-Driven Ghosting using Deep Imitation Learning
Hoang Le, Peter Carr, Yisong Yue, Patrick Lucey. SSAC 2017

But Who Plays Which Role?

- All we get are trajectories!
  - Don’t know which belongs to which role.

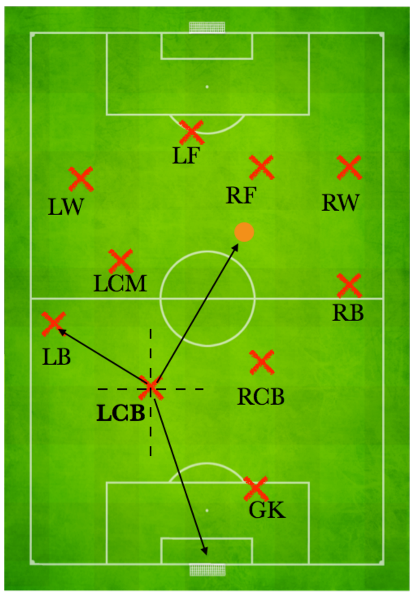


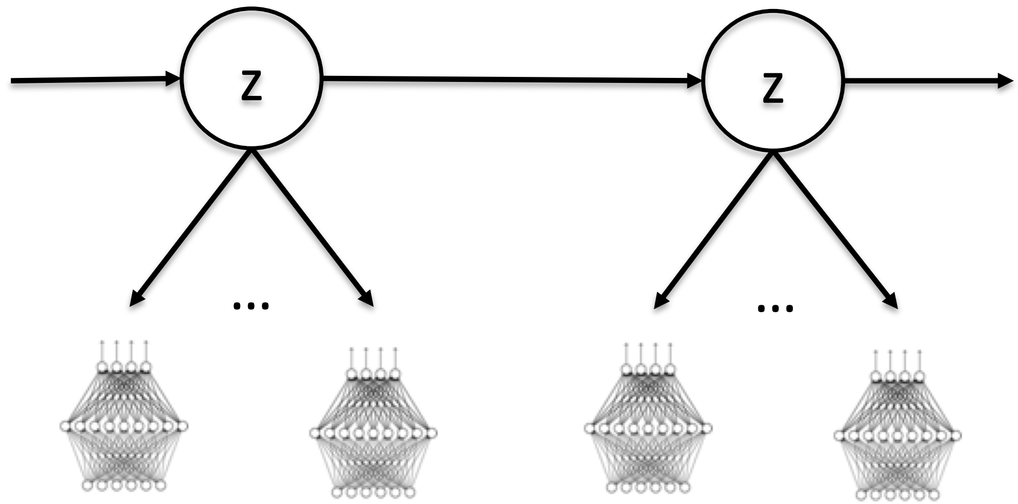
- Need to solve a permutation problem
  - **Naïve baseline ignores this!**

Hoang
Le

Coordination Model

Mixture of Gaussians HMM





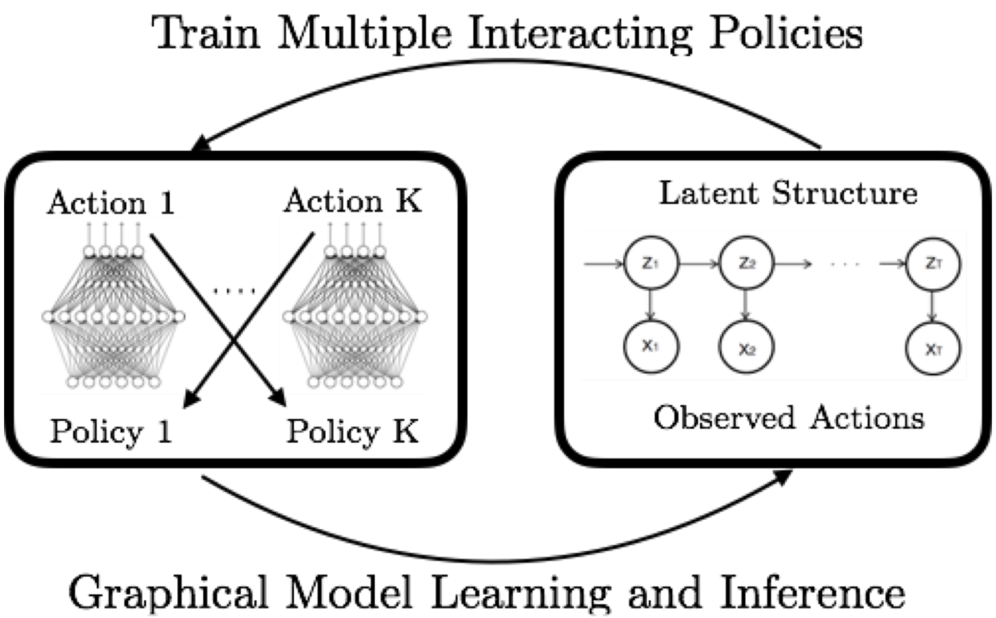
Single-Agent Policies

Coordinated Multi-Agent Imitation Learning
Hoang Le, Yisong Yue, Peter Carr, Patrick Lucey. ICML 2017

Hoang Le

Learning Algorithm

Standard Imitation Learning

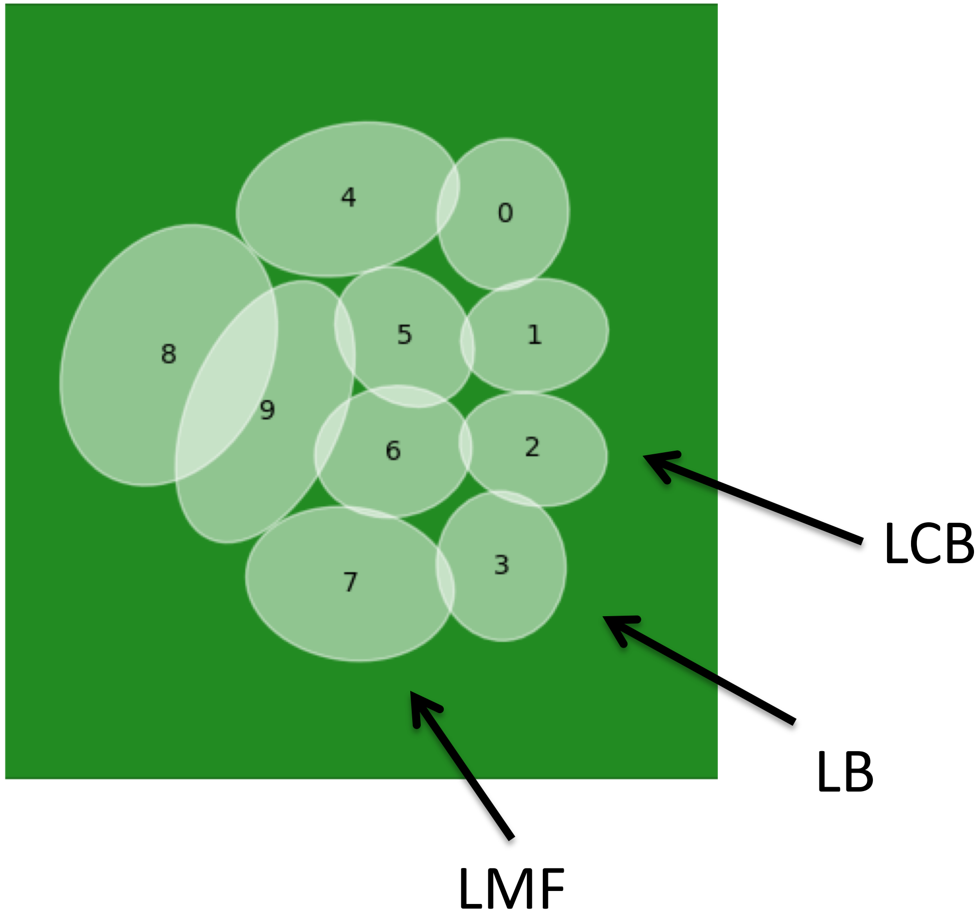


Stochastic Variational Inference

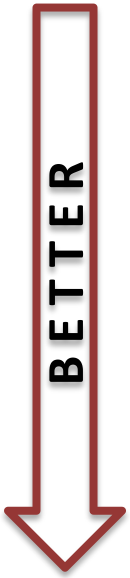
Coordinated Multi-Agent Imitation Learning

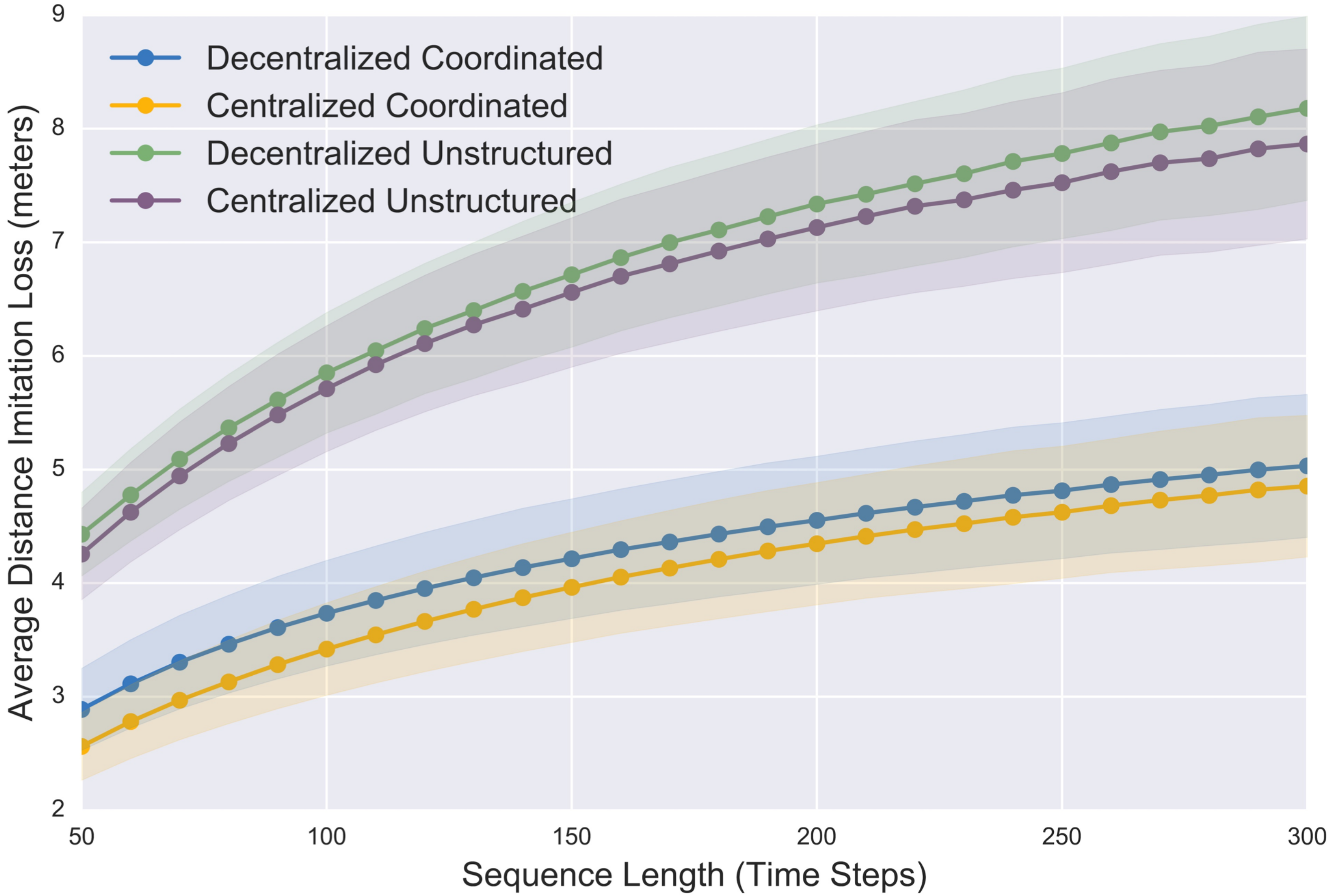
Hoang Le, Yisong Yue, Peter Carr, Patrick Lucey. ICML 2017

Learned Roles



Coordinated vs Uncoordinated

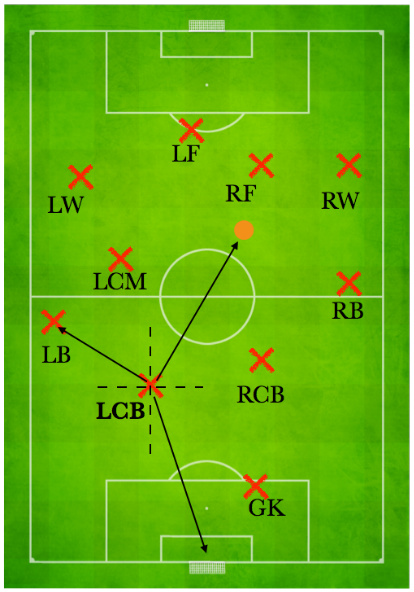




Coordinated Multi-Agent Imitation Learning

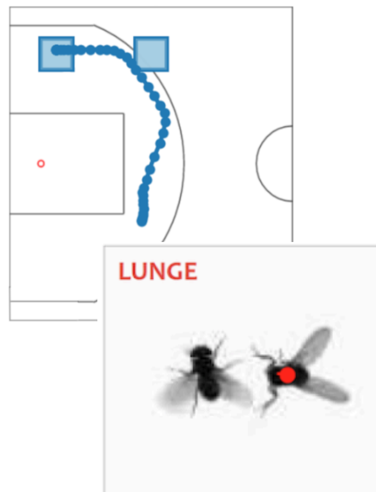
Hoang Le, Yisong Yue, Peter Carr, Patrick Lucey. ICML 2017

Outline For Today



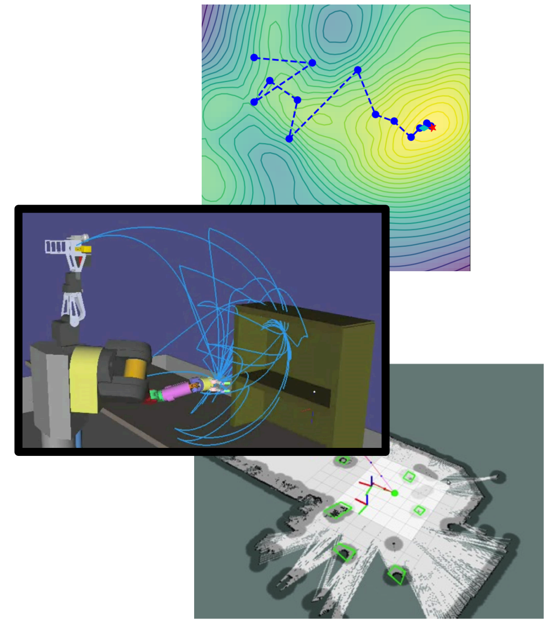
Coordinated Learning

Infer Latent Roles



Hierarchical Behaviors

Generative Behavior



Learning to Optimize

Learn to Infer

← **Probabilistic Imitation Learning**

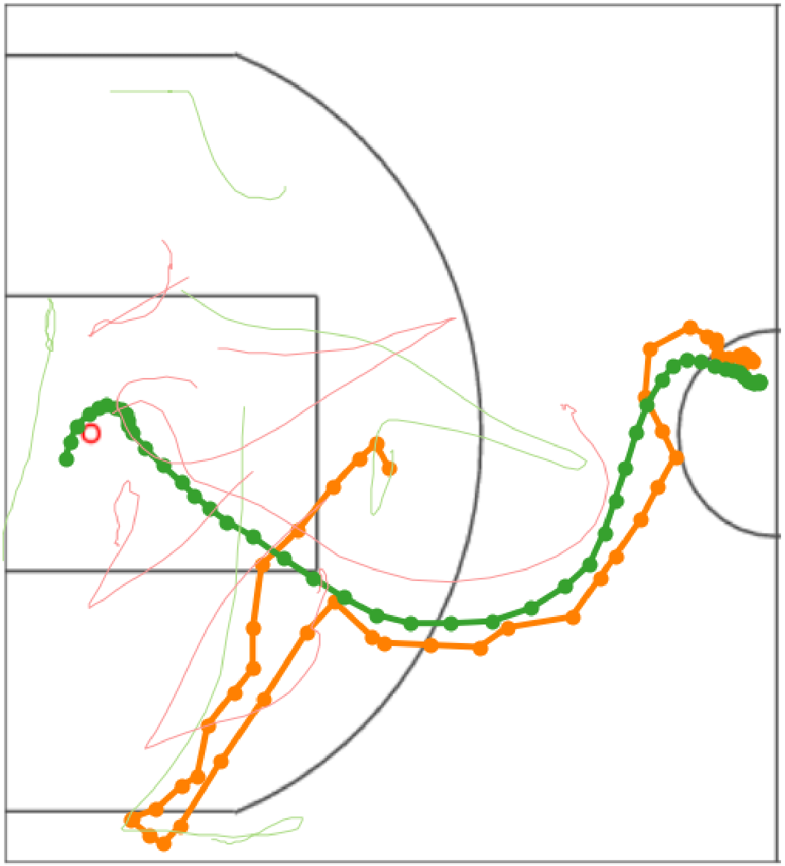
→ **Learning to Infer**

Strategy vs Tactics

- Long-term Goal:
  - Curl around basket
- Tactics
  - Drive left w/ ball
  - Pass ball
  - Cut towards basket

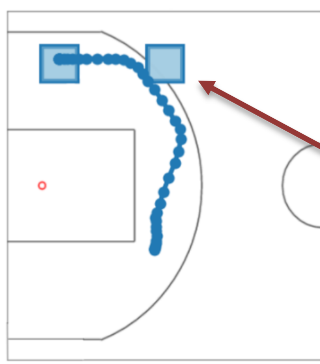
Stephan Zheng

Eric Zhan

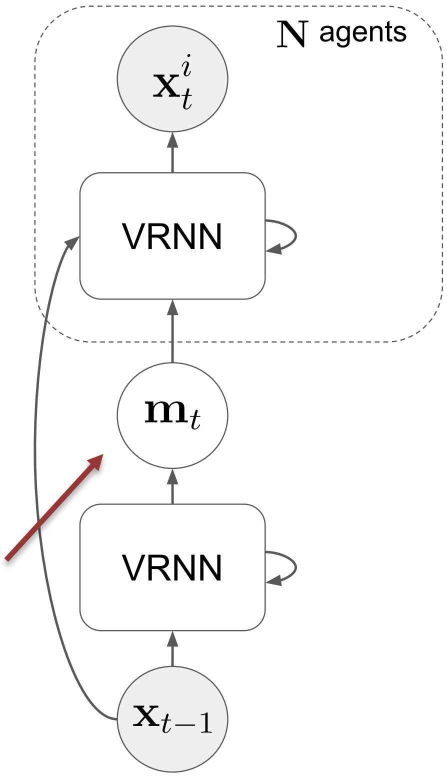


Generative + Hierarchical Imitation Learning

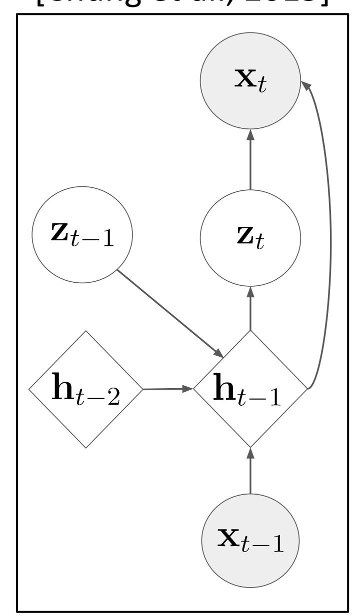
- **Generative Imitation Learning**
  - No single "correct" action
- **Hierarchical**
  - Predictions at multiple resolutions



Macro-goals

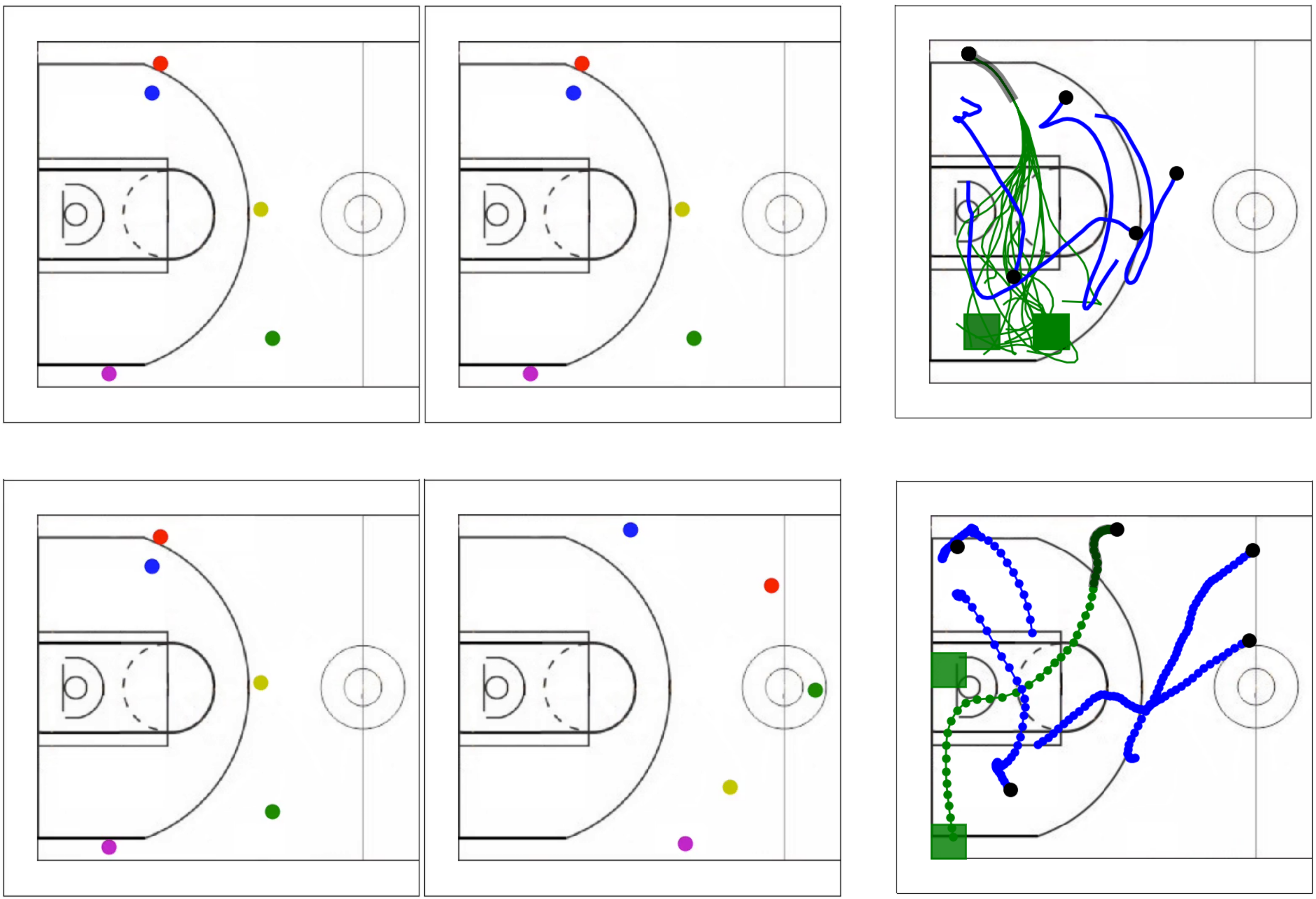


VRNN
[Chung et al., 2015]



Generating Long-term Trajectories using Deep Hierarchical Networks
Stephan Zheng, Yisong Yue, Patrick Lucey. NIPS 2016

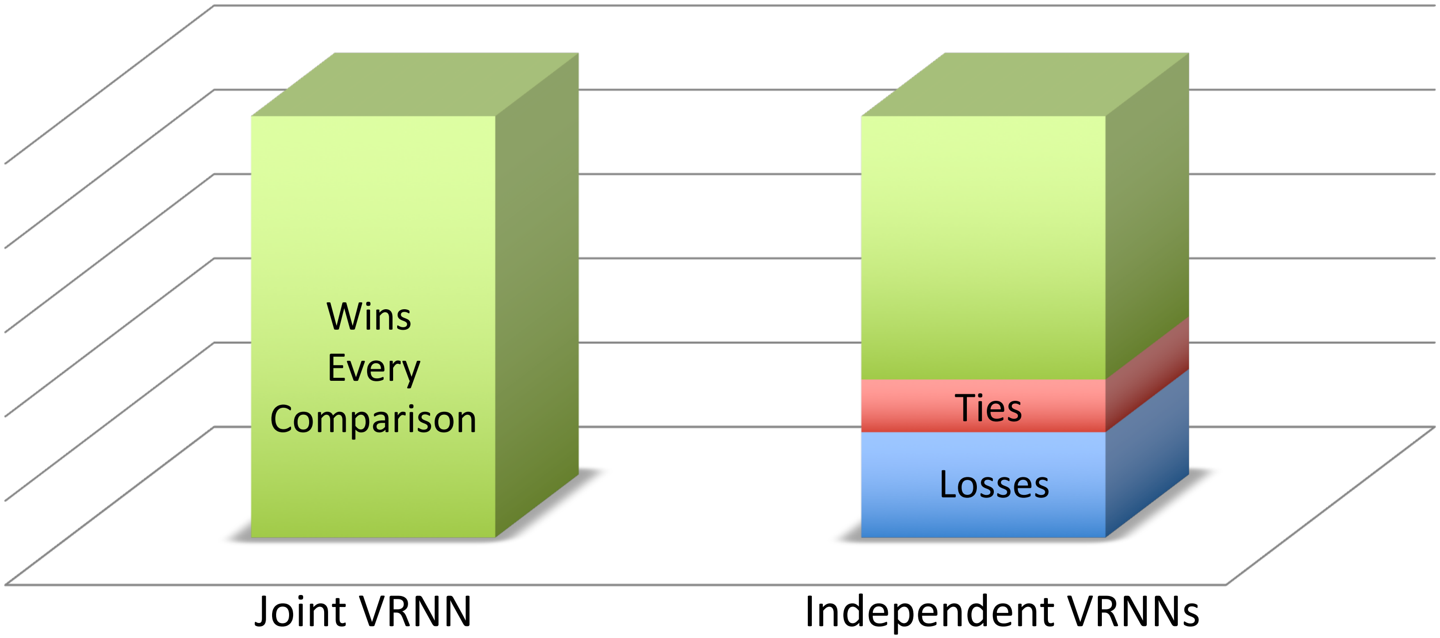
Generative Multi-Agent Behavioral Cloning
Eric Zhan, Stephan Zheng, Yisong Yue, Long Sha, Patrick Lucey. (under review)



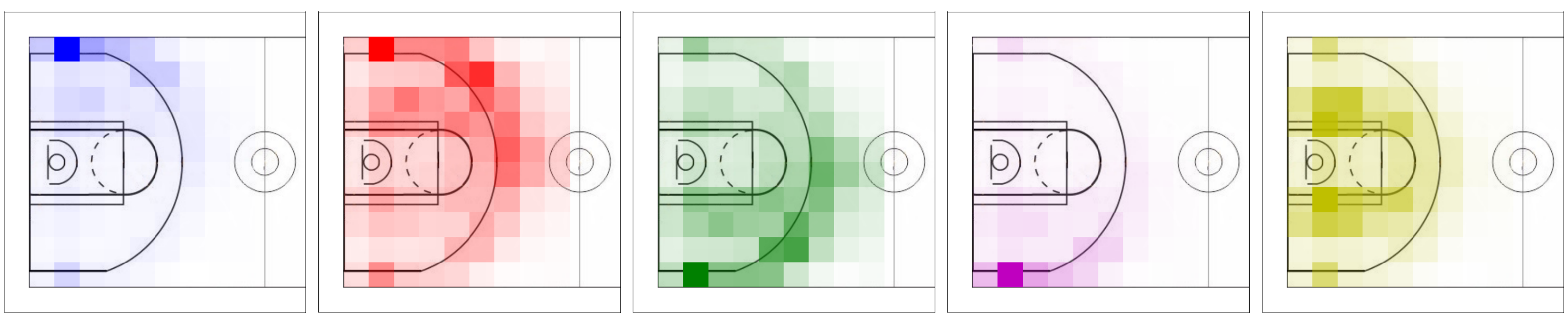
http://basketball-ai.com/

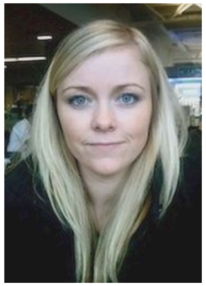
User Study

(14 Professional Sports Analysts, 25 scenarios)



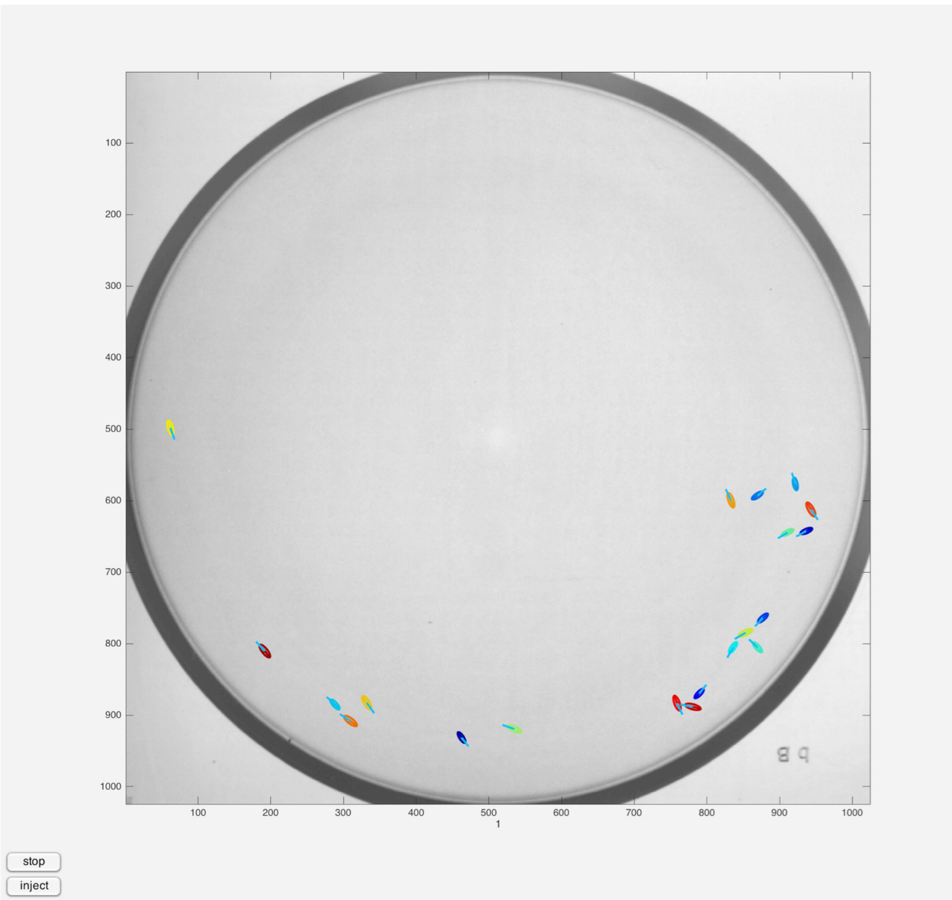
Heatmap of Macro-Goals per Role:

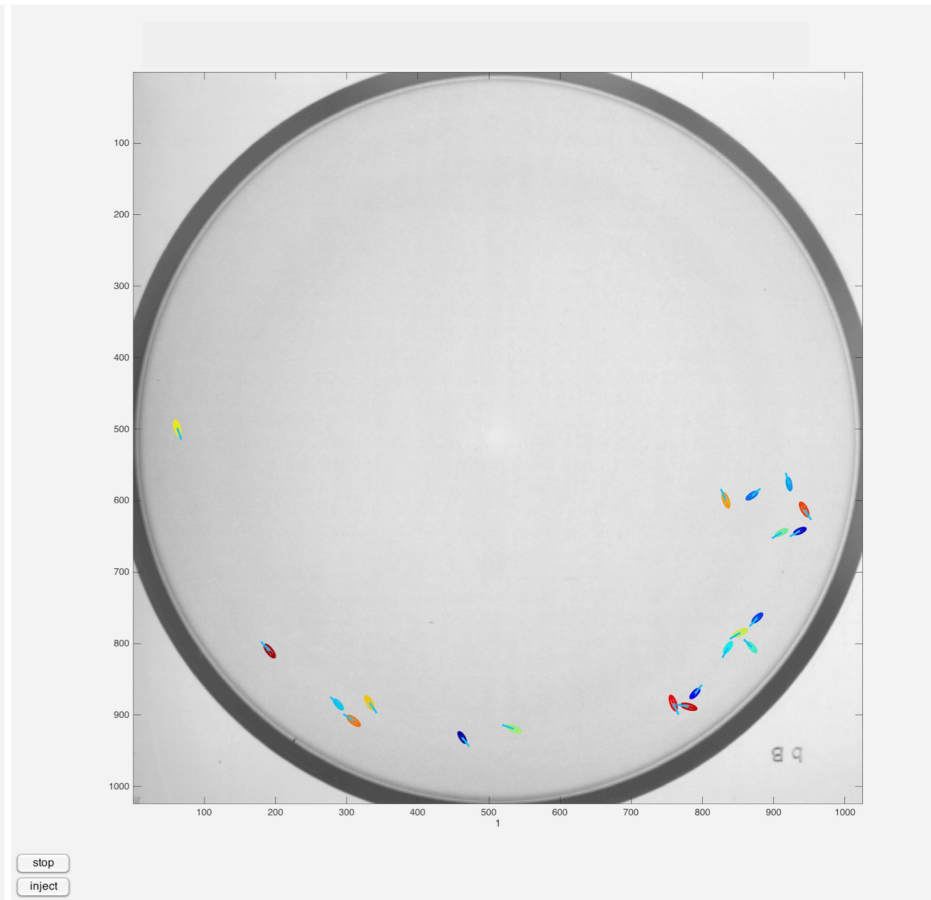




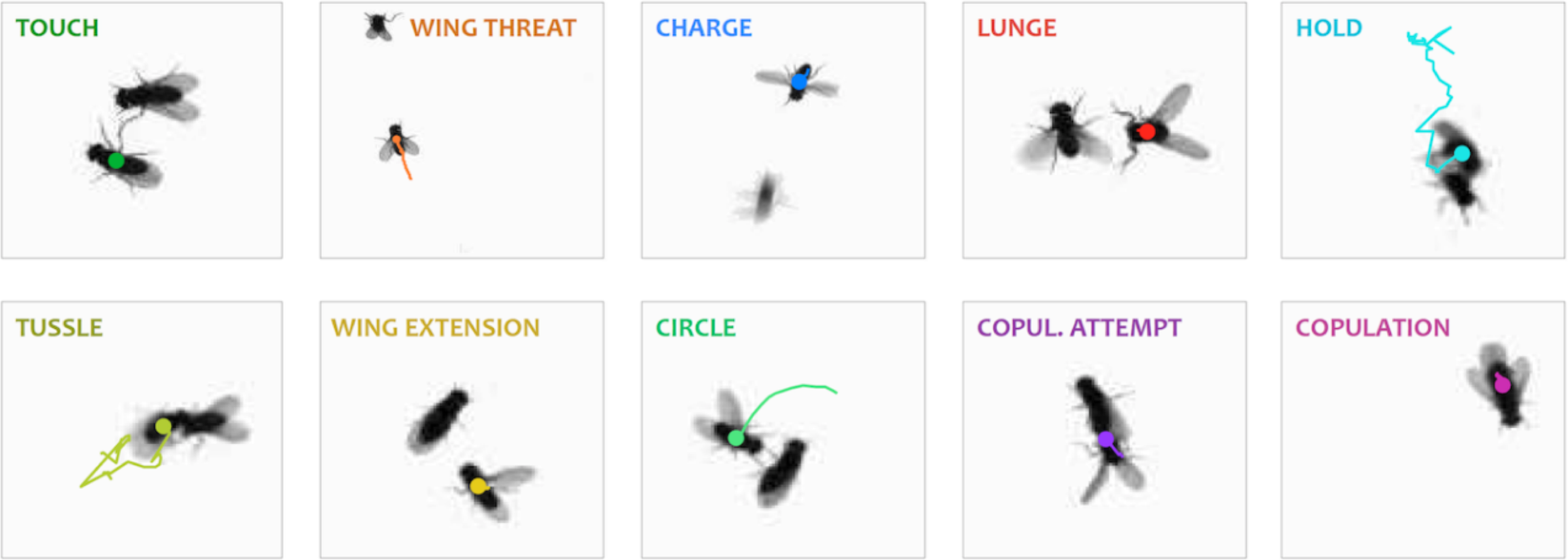
Eyrun
Eyolfsdottir

Aside: Animal Behavior



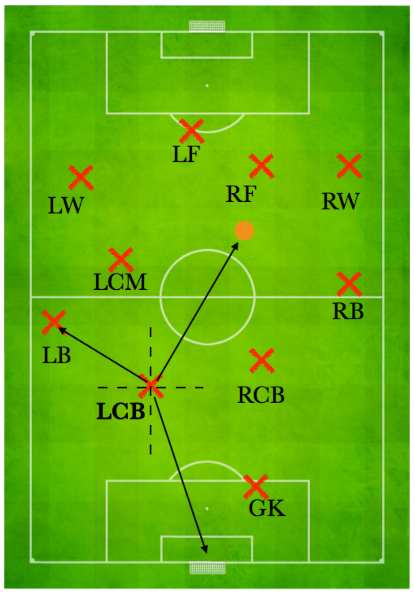


Activity Labels



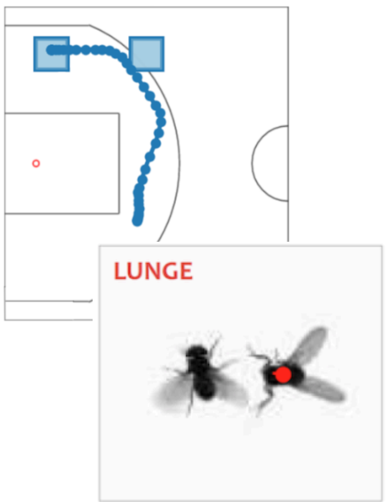
Learning recurrent representations for hierarchical behavior modeling
Eyrun Eyolfsdottir, Kristin Branson, Yisong Yue, Pietro Perona, ICLR 2017

Outline For Today



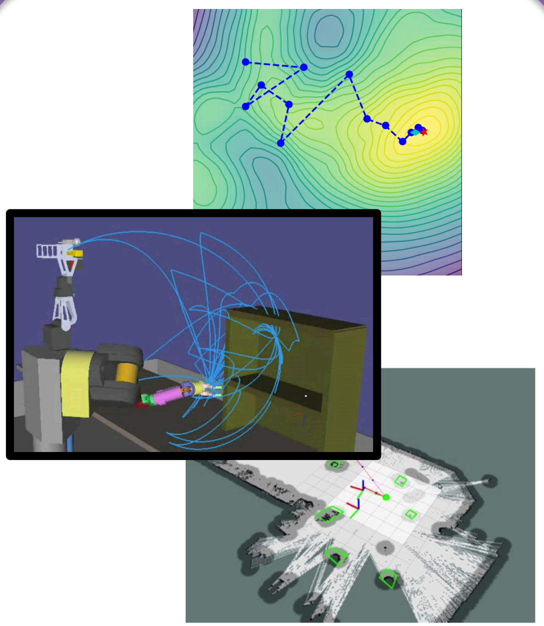
Coordinated Learning

Infer Latent Roles



Hierarchical Behaviors

Generative Behavior



Learning to Optimize

Learn to Infer

← **Probabilistic Imitation Learning**

→ **Learning to Infer**

Optimization as Sequential Decision Making

- Many solvers are sequential:
  - Greedy
  - Search heuristics
  - Gradient Descent

- Can view as solver as "agent"
  - State = intermediate solution
  - Find a state with high reward (solution)

Optimization as Sequential Decision Making

Contextual Submodular Maximization

- Training set: (x, F_x)
- Greedily maximize F_x using only x
- **Learning Policies for Contextual Submodular Prediction** [ICML 2013]

Stephane Ross

Learning to Search

- Training set: $(x{=}\mathrm{MILP}, y{=}\mathrm{solution/search{-}trace})$
- Find y (or better solution)
- **Learning to Search via Retrospective Imitation** [under review]

Jialin Song

Learning to Infer

- Training set: $(x{=}\mathrm{data/model}, L{=}\mathrm{likelihood}\ )$
- Iteratively optimize L (generalizes VAEs)
- **Iterative Amortized Inference** [ICML 2018]

Joe Marino

Optimization as Sequential Decision Making

Contextual Submodular Maximization

- Training set: (x, F_x)
- Greedily maximize F_x using only x
- **Learning Policies for Contextual Submodular Prediction** [ICML 2013]

Stephane Ross

Learning to Search

- Training set: $(x\text{=MILP}, y\text{=solution/search−trace})$
- Find y (or better solution)
- **Learning to Search via Retrospective Imitation** [under review]

Jialin Song

Learning to Infer

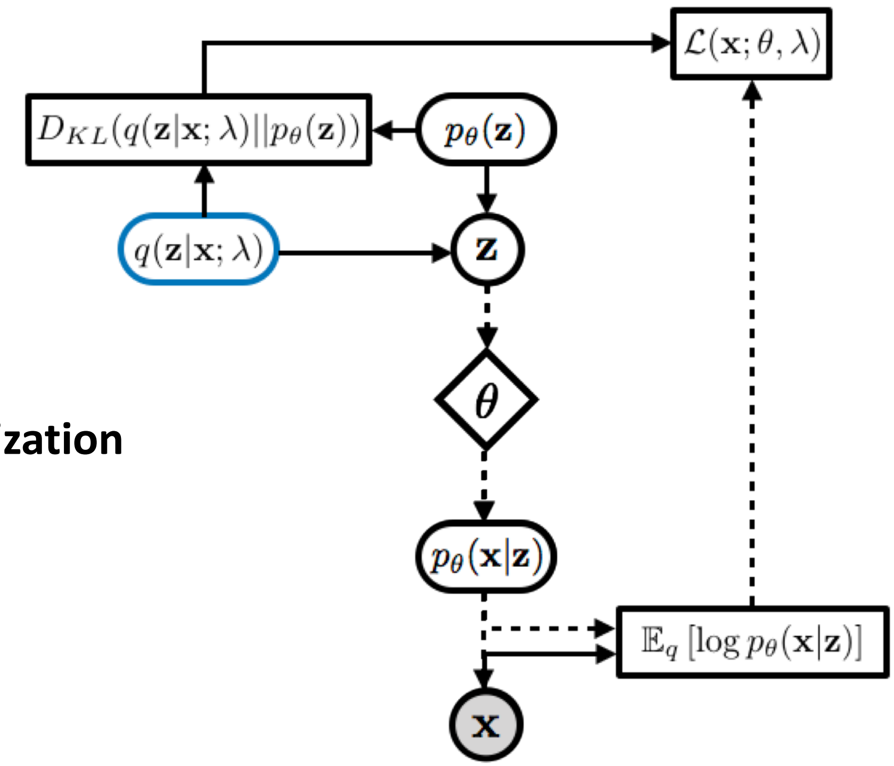
- Training set: $(x\text{=data/model}, L\text{=likelihood })$
- Iteratively optimize L (generalizes VAEs)
- **Iterative Amortized Inference** [ICML 2018]

Joe Marino

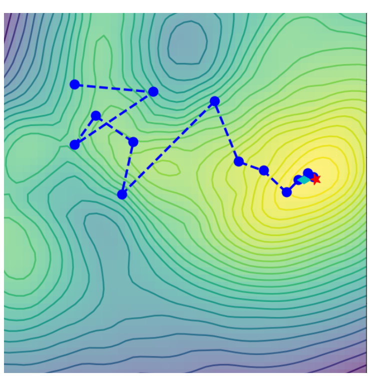
Variational Inference

approximate posterior $q(\mathbf{z}|\mathbf{x};\lambda)$

ELBO $\mathcal{L}(\mathbf{x};\theta,\lambda) = \mathbb{E}_q\left[\log p_\theta(\mathbf{x}|\mathbf{z})\right] - D_{KL}(q(\mathbf{z}|\mathbf{x};\lambda)||p_\theta(\mathbf{z}))$



Inference = Optimization



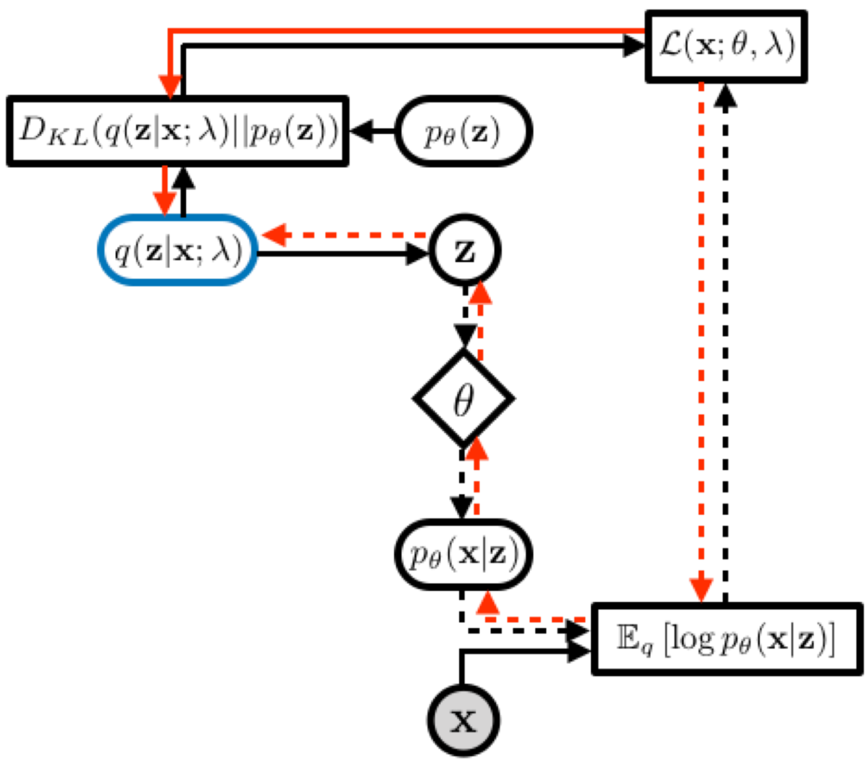
How do we solve $\lambda \leftarrow \text{argmax}_\lambda \mathcal{L}(\mathbf{x}; \theta, \lambda)$?

conventional optimization techniques (e.g. SGD)

update using an estimate of the gradient

$$\lambda_{t+1} \leftarrow \lambda_t + \alpha \nabla_\lambda \mathcal{L}(\mathbf{x}; \theta, \lambda_t)$$

e.g. Hoffman *et al.*, 2013



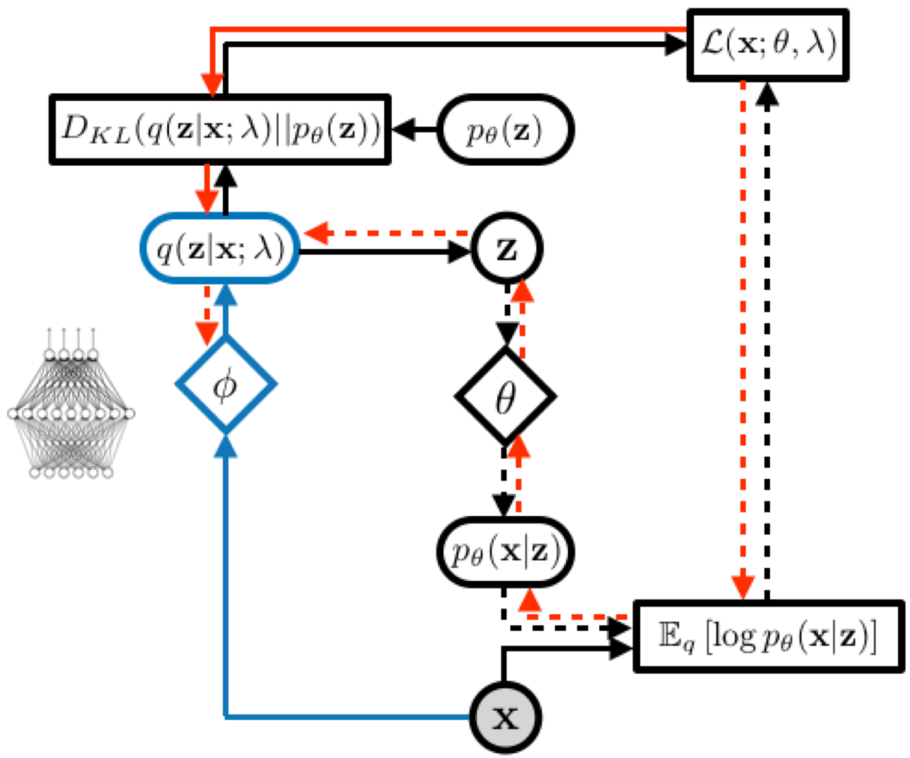
How do we solve $\lambda \leftarrow \text{argmax}_\lambda \mathcal{L}(\mathbf{x}; \theta, \lambda)$?

amortized inference

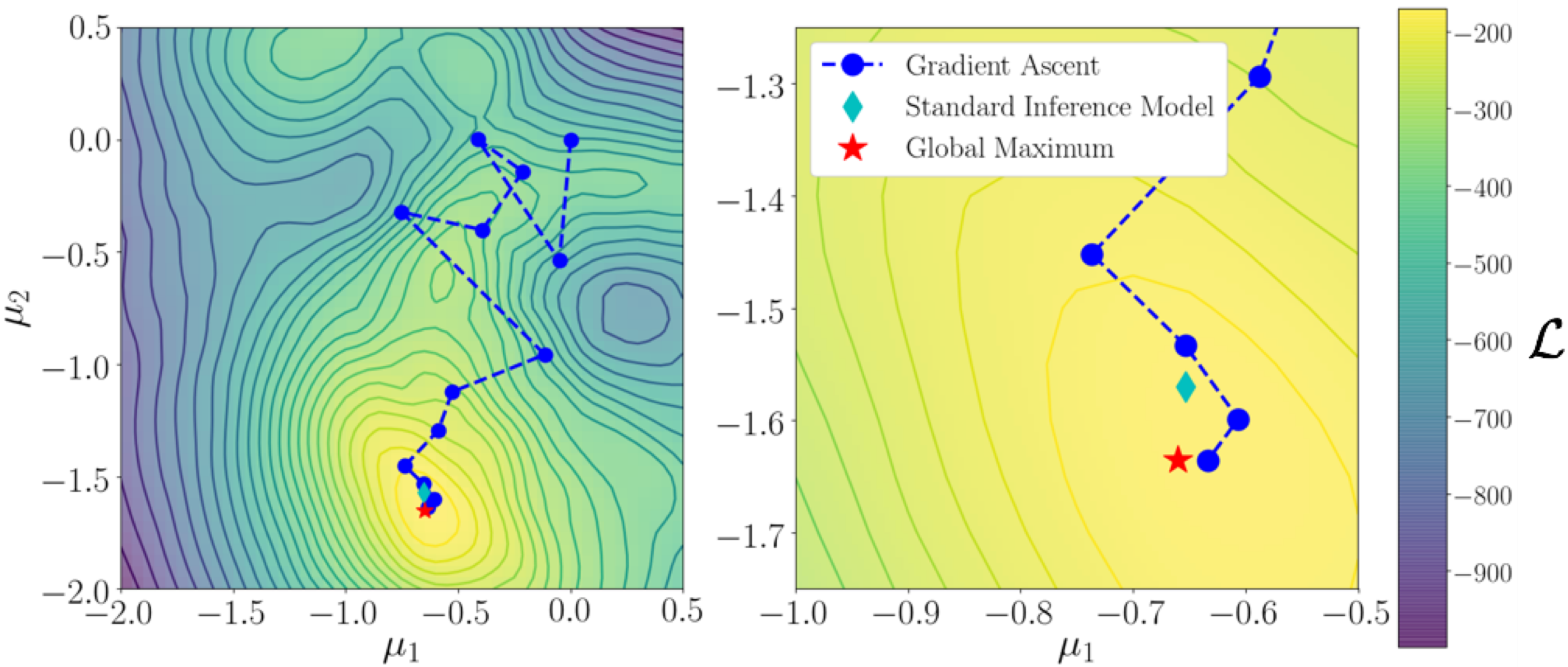
learn a model to solve inference optimization

$$\lambda \leftarrow f(\mathbf{x}; \phi)$$

e.g. Dayan *et al.*, 1995, Rezende *et al.*, 2014



Amortization Gap



2D Model, MNIST

No Explicit Prior Information

inference optimization depends on the <u>*prior*</u>

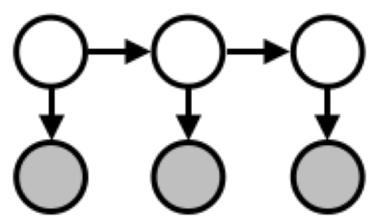
$$\mathcal{L}(\mathbf{x};\theta,\lambda) = \mathbb{E}_q\left[\log p_\theta(\mathbf{x}|\mathbf{z})\right] - D_{KL}(q(\mathbf{z}|\mathbf{x};\lambda)||p_\theta(\mathbf{z}))$$

standard inference models only condition on the <u>*data*</u>, and must therefore *implicitly* account for the prior

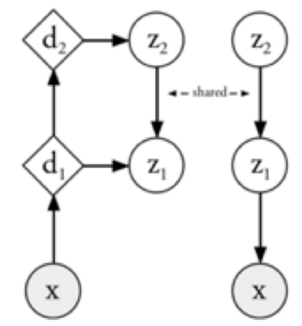
problematic in models with varying priors

hierarchical

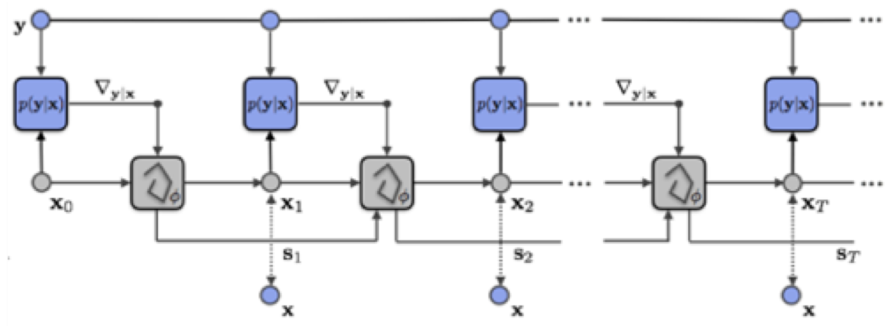
dynamical



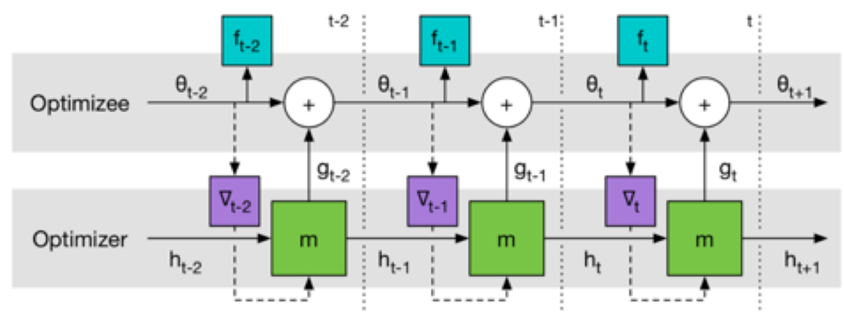
Related Work



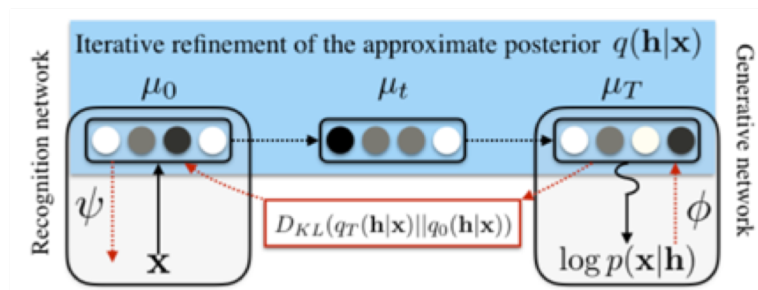
Ladder VAE
Sønderby *et al.*, 2016



Recurrent Inference Machines
Putzky & Welling, 2017

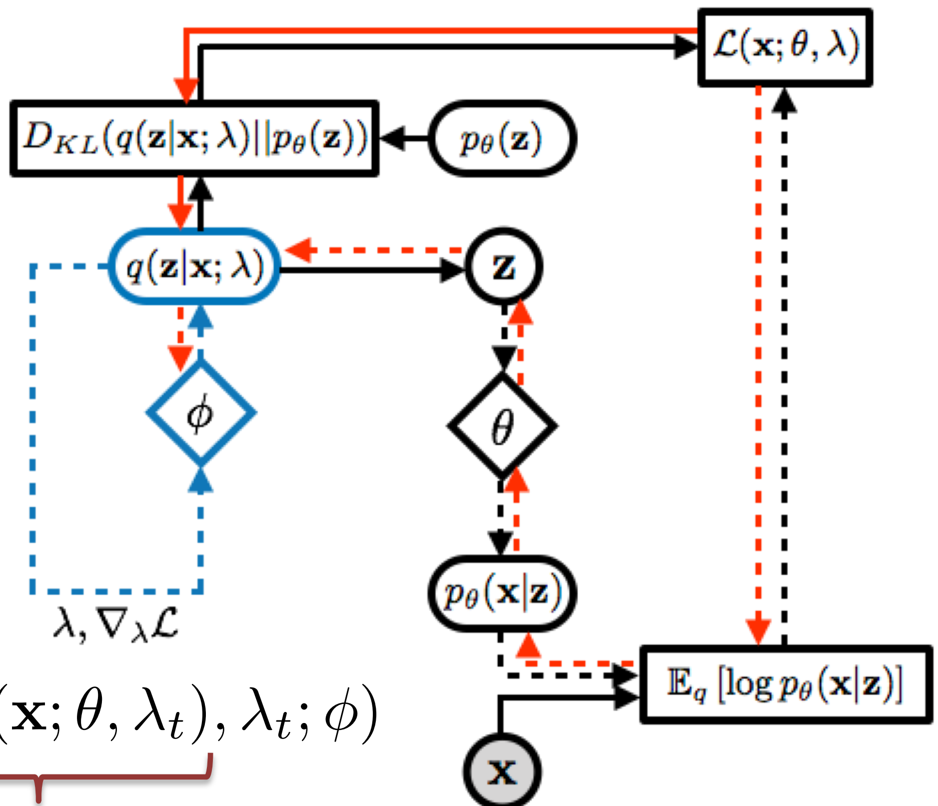


Learning to Learn GD by GD
Andychowicz *et al.*, 2016

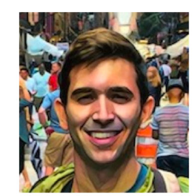


Initial Encoding, Iterative Refinement
Krishnan *et al.*, 2018
Hjelm *et al.*, 2016

Iterative Inference Networks



$$\lambda_{t+1} \leftarrow f(\underbrace{\nabla_\lambda \mathcal{L}(\mathbf{x};\theta,\lambda_t)}_{\text{Can encode errors from data-dependent priors}}, \lambda_t; \phi)$$

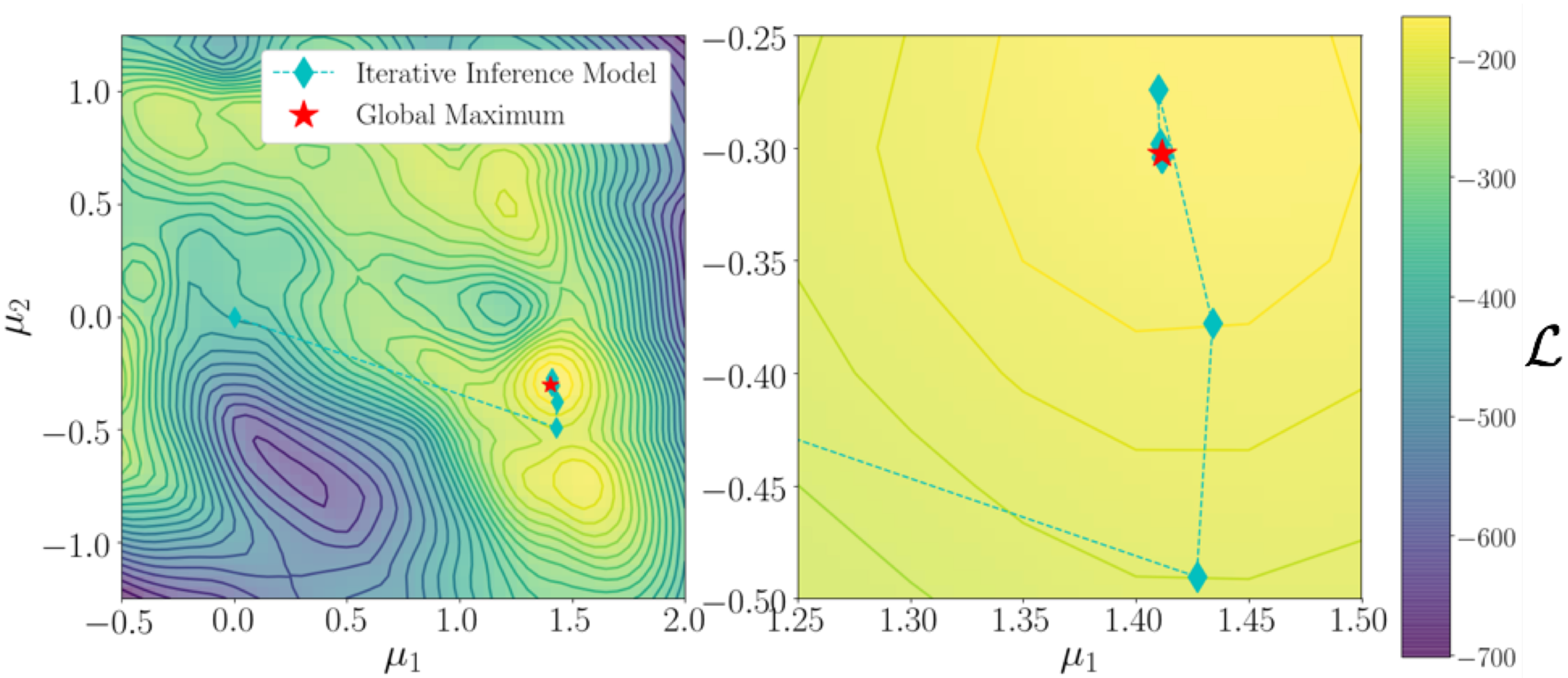


Joe Marino

Iterative Amortized Inference

Joe Marino, Yisong Yue, Stephan Mandt. ICML 2018

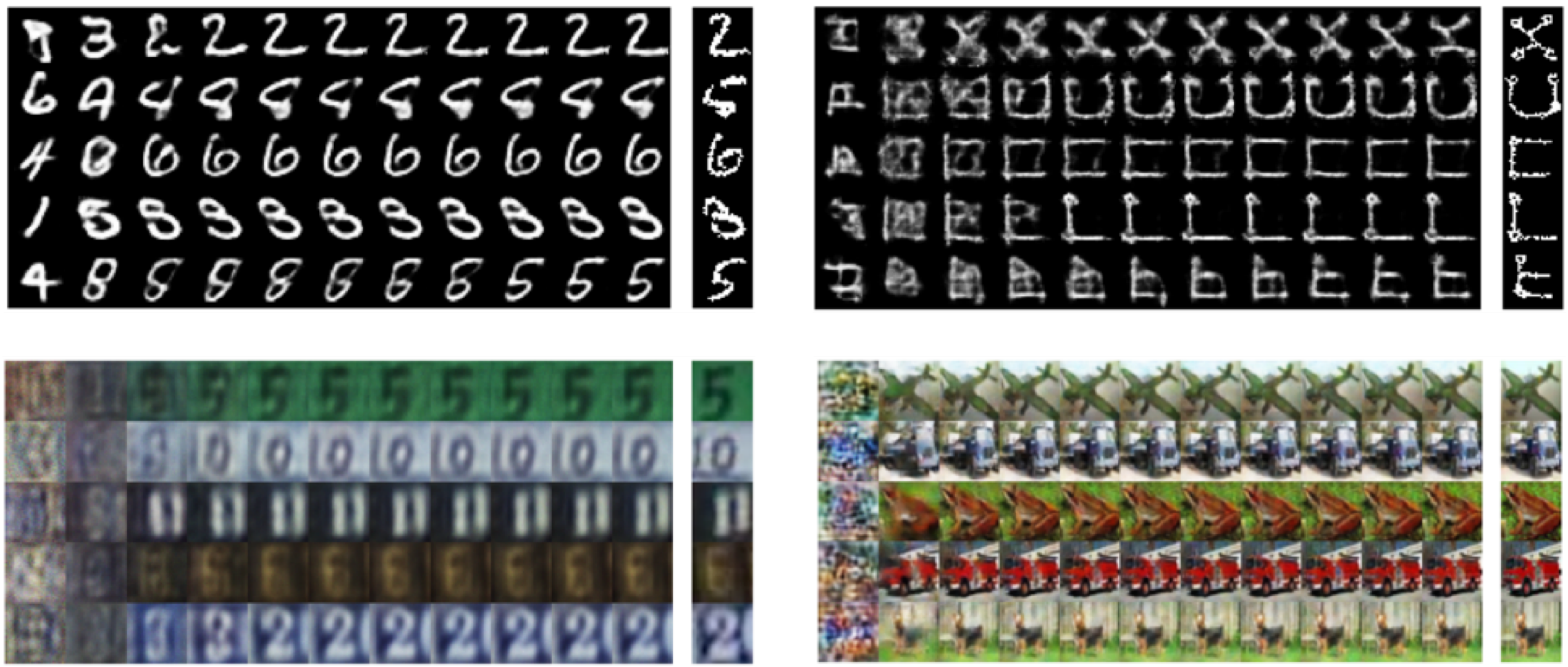
Inference Optimization



Iterative Amortized Inference

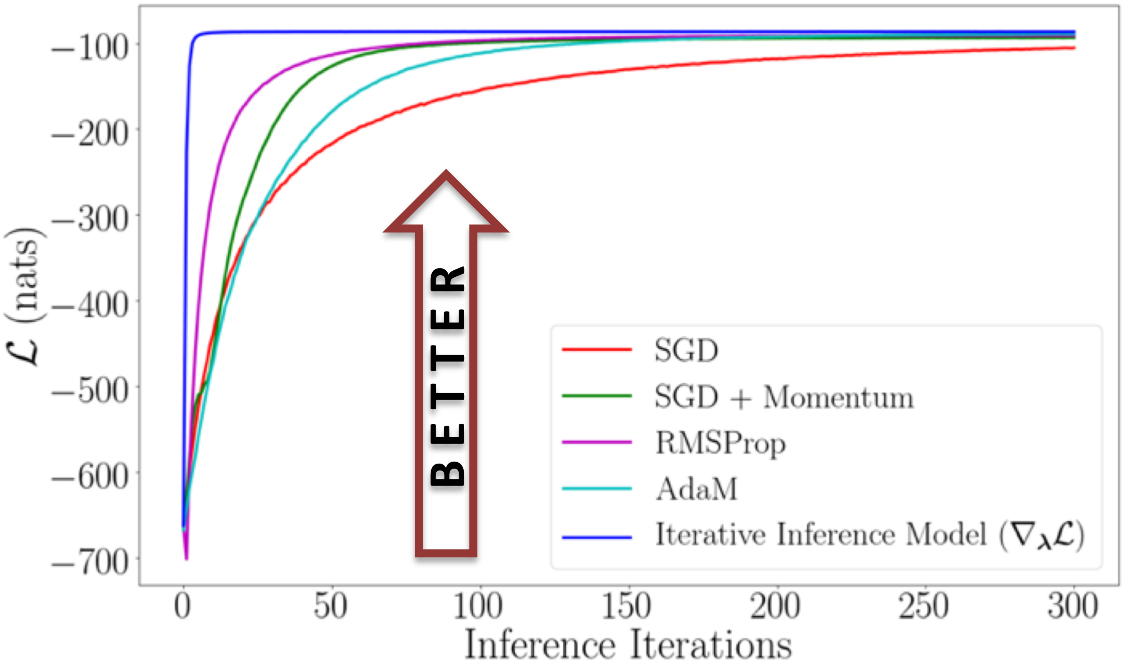
Joe Marino, Yisong Yue, Stephan Mandt. ICML 2018

Inference Optimization



Iterative Amortized Inference
Joe Marino, Yisong Yue, Stephan Mandt. ICML 2018

Inference Optimization



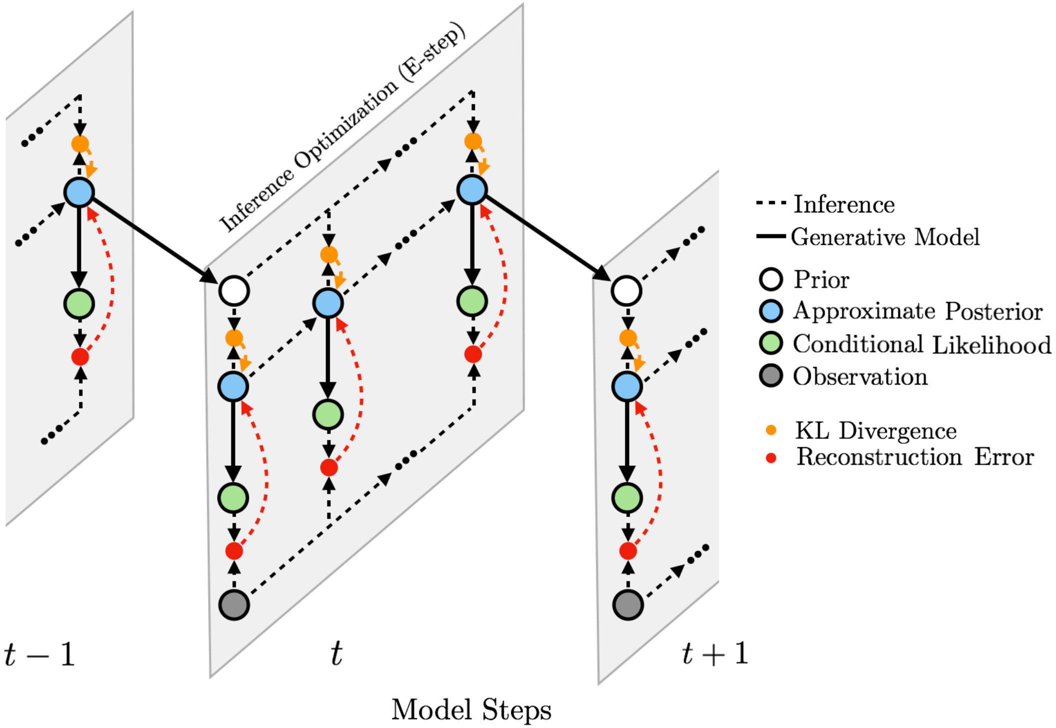
	$-\log p(\mathbf{x})$
MNIST	
Single-Level	
Standard	84.14 ± 0.02
Iterative	$\mathbf{83.84 \pm 0.05}$
Hierarchical	
Standard	82.63 ± 0.01
Iterative	$\mathbf{82.457 \pm 0.001}$
CIFAR-10	
Single-Level	
Standard	5.823 ± 0.001
Iterative	$\mathbf{5.64 \pm 0.03}$
Hierarchical	
Standard	5.565 ± 0.002
Iterative	$\mathbf{5.456 \pm 0.005}$

Iterative Amortized Inference

Joe Marino, Yisong Yue, Stephan Mandt. ICML 2018

Ongoing Work

Joe Marino

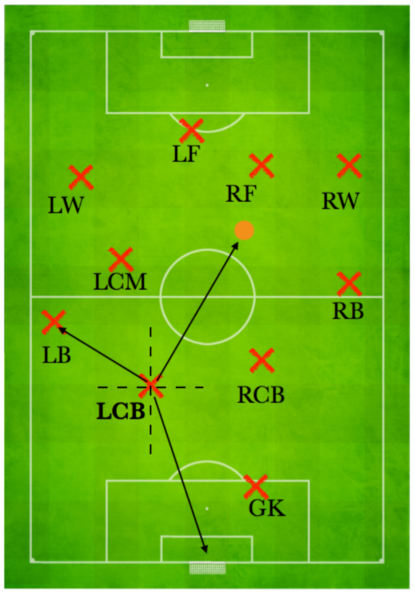


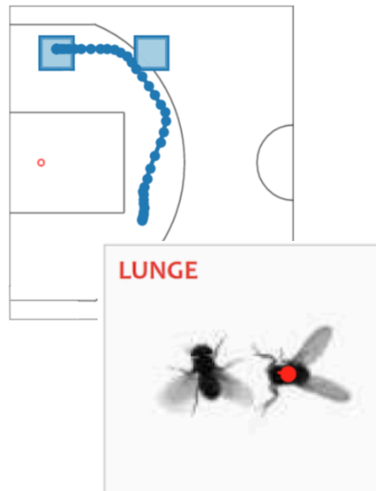
A General Framework for Amortizing Variational Filtering

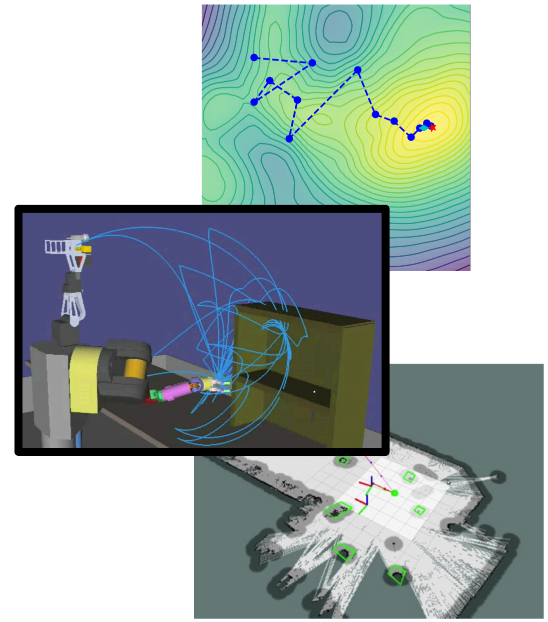
Theoretical Foundations & Applications of Deep Generative Models Workshop @ ICML

(Sat. & Sun.)

Inference + Imitation







Coordinated Learning

Infer Latent Roles

Hierarchical Behaviors

Generative Behavior

Learning to Optimize

Learn to Infer

← **Probabilistic Imitation Learning**

→ **Learning to Infer**

References

Data-Driven Ghosting using Deep Imitation Learning
Hoang Le, Peter Carr, Yisong Yue, Patrick Lucey. SSAC 2017 *(Best Paper Runner Up)*

Coordinated Multi-agent Imitation Learning
Hoang Le, Yisong Yue, Peter Carr, Patrick Lucey. ICML 2017

Generative Multi-Agent Behavioral Cloning
Eric Zhan, Stephan Zheng, Yisong Yue, Long Sha, Patrick Lucey. arXiv 2018

Generating Long-term Trajectories using Deep Hierarchical Networks
Stephan Zheng, Yisong Yue, Patrick Lucey. NIPS 2016

Learning recurrent representations for hierarchical behavior modeling
Eyrun Eyolfsdottir, Kristin Branson, Yisong Yue, Pietro Perona. ICLR 2017

Iterative Amortized Inference
Joseph Marino, Yisong Yue, Stephan Mandt. ICML 2018

Learning to Search via Retrospective Imitation
Jialin Song, Ravi Lanka, Albert Zhao, Yisong Yue, Masahiro Ono. arXiv 2018

Learning Policies for Contextual Submodular Prediction
Stephane Ross, Robin Zhou, Yisong Yue, Debadeepta Dey, Yisong Yue. ICML 2013

http://www.yisongyue.com