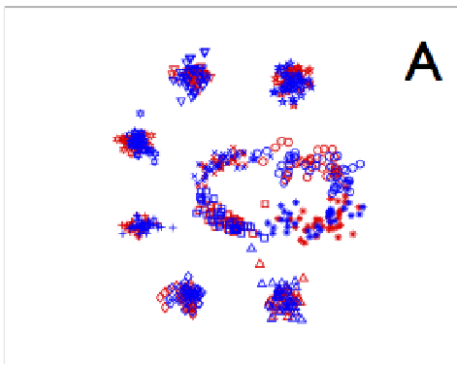
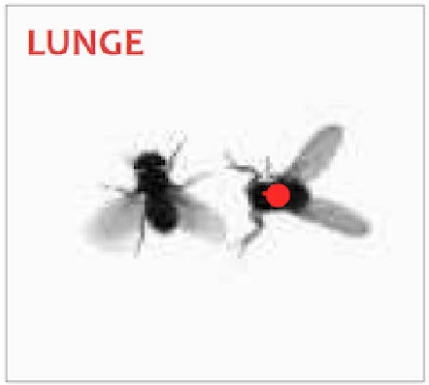
dolcit

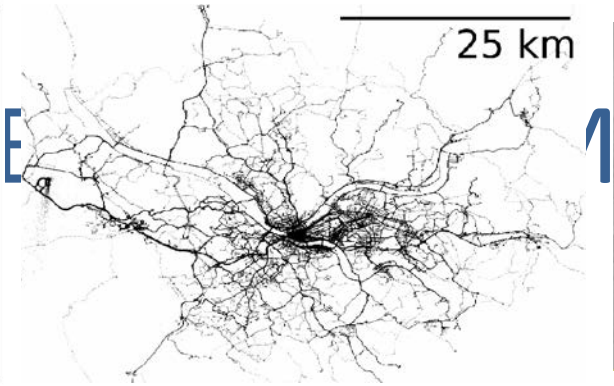
Caltech

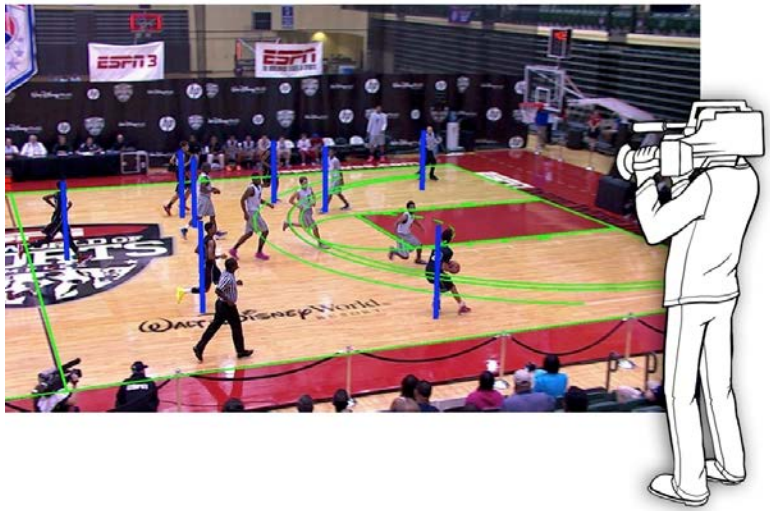
New Frontiers in Imitation Learning

Yisong Yue


A


LUNGE


25 km

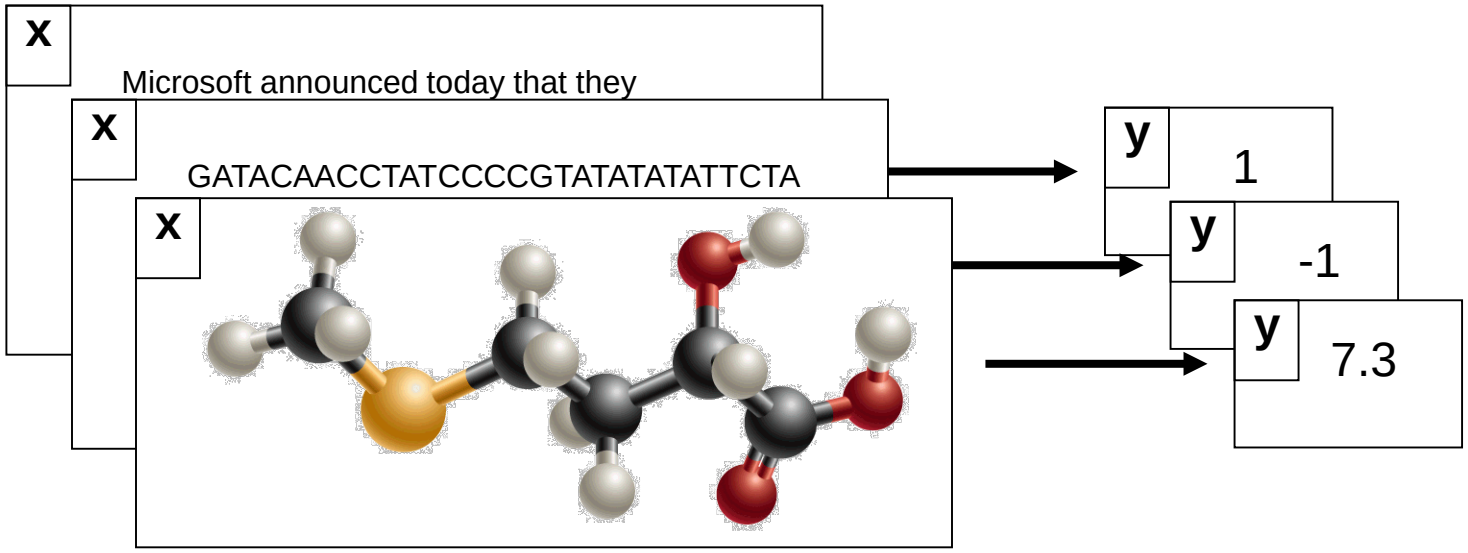

ESPN3
ESPN

Warm Up: Supervised Learning

- Find function from input space *X* to output space *Y*

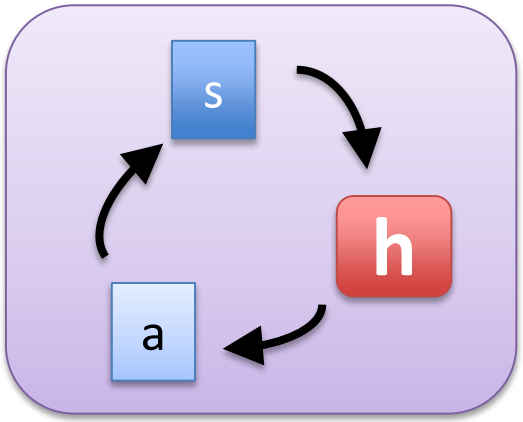
$$h : X \longrightarrow Y$$

such that the prediction error is low.



Imitation Learning

- Input:
  - Sequence of contexts/states:
- Predict:
  - Sequence of actions
- Learn Using:
  - Sequences of demonstrated actions

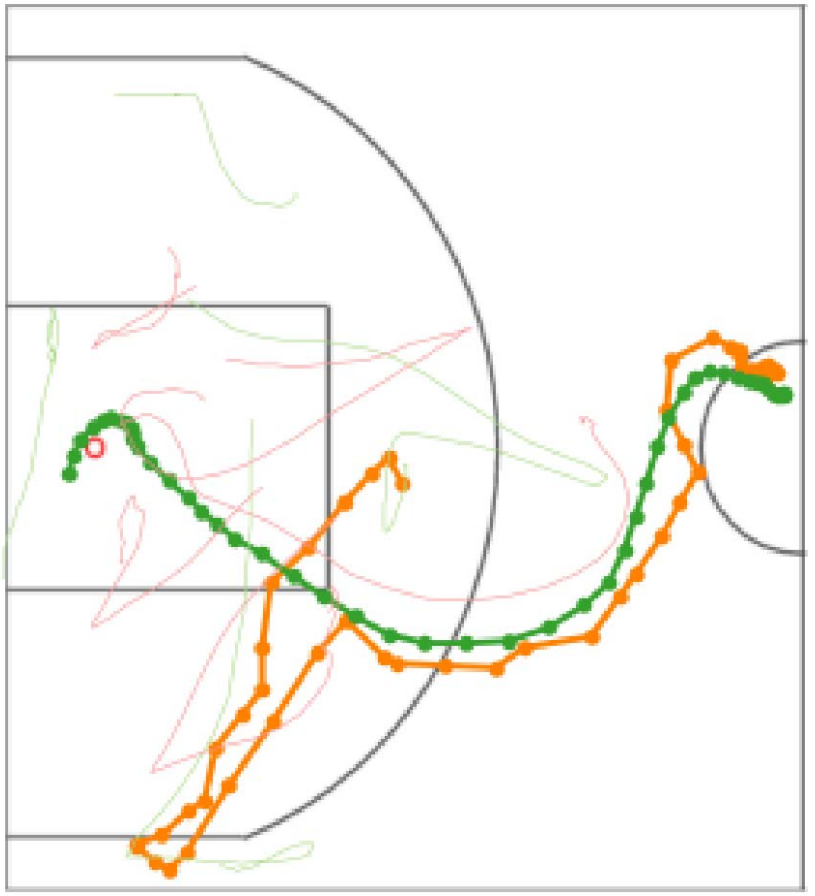


Example: Basketball Player Trajectories

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

- Training set: $D = \{(\vec{s}, \vec{a})\}$
  - $\vec{s}$ = sequence of s
  - $\vec{a}$ = sequence of a

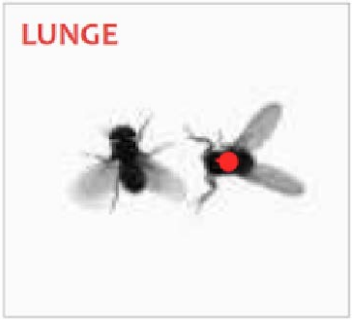
- **Goal:** learn $h(s) \rightarrow a$



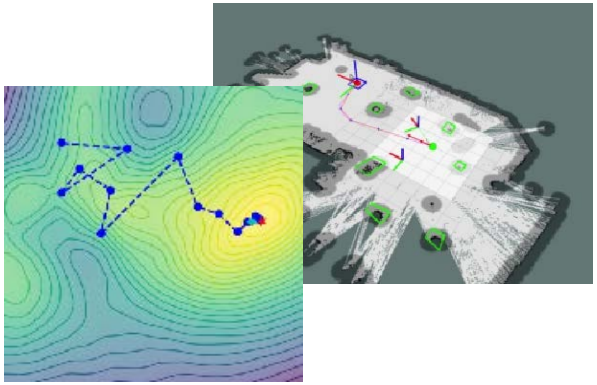
What to Imitate?

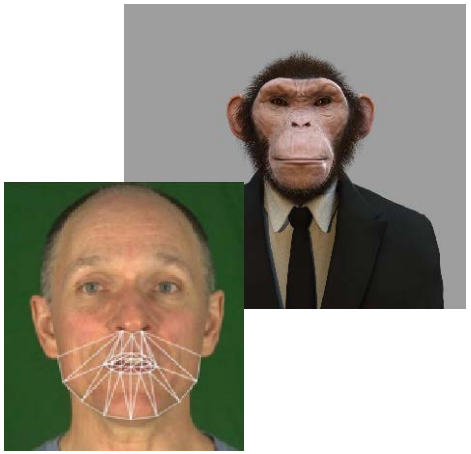
Human Demonstrations

Animal Demonstrations

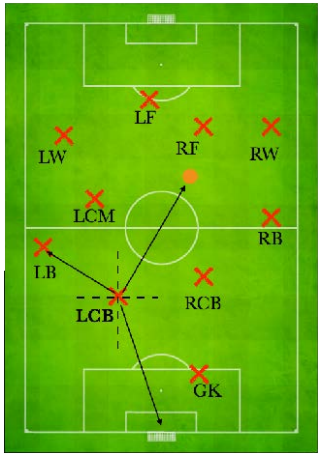


Computational Oracle

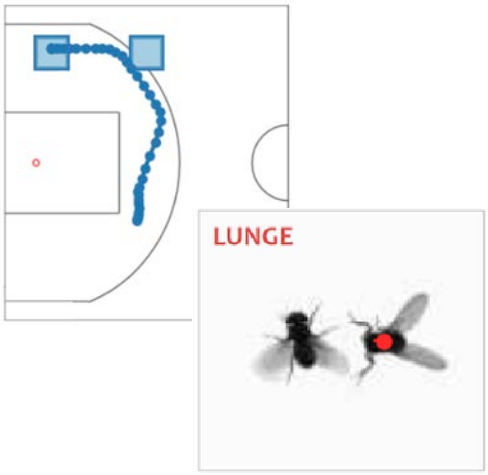




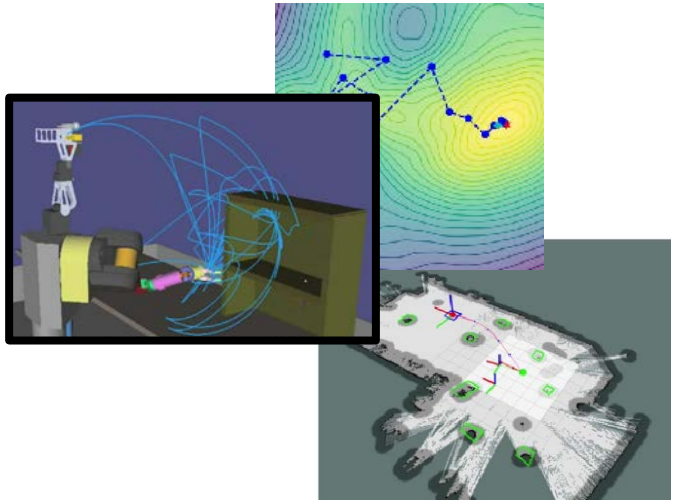
Speech Animation



Coordinated Learning

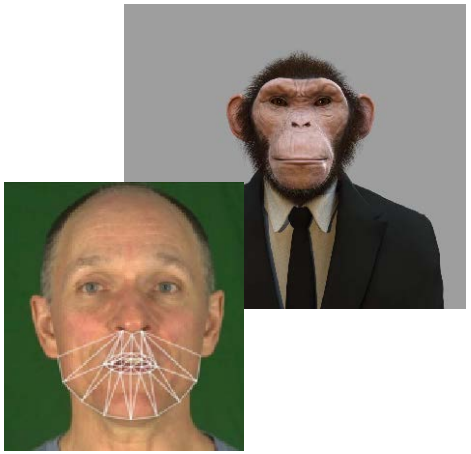


Hierarchical Behaviors (Generative)

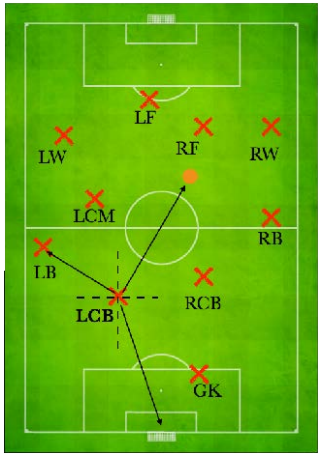


Learning to Optimize

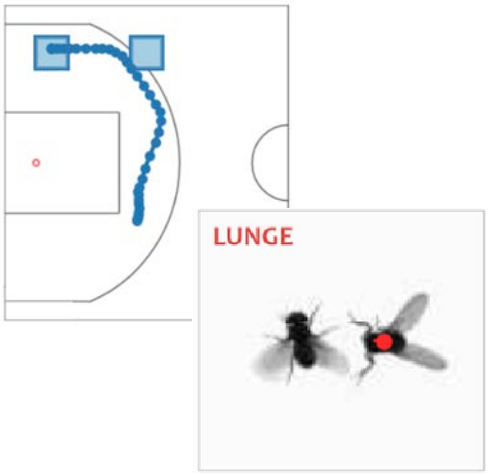
Smooth Imitation Learning



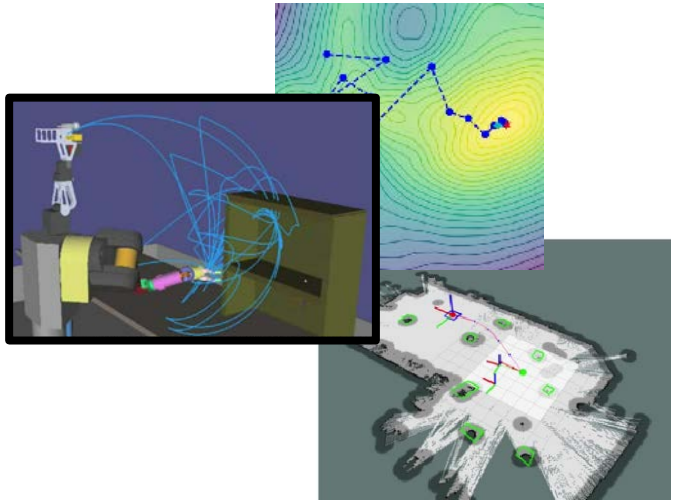
Speech Animation



Coordinated Learning

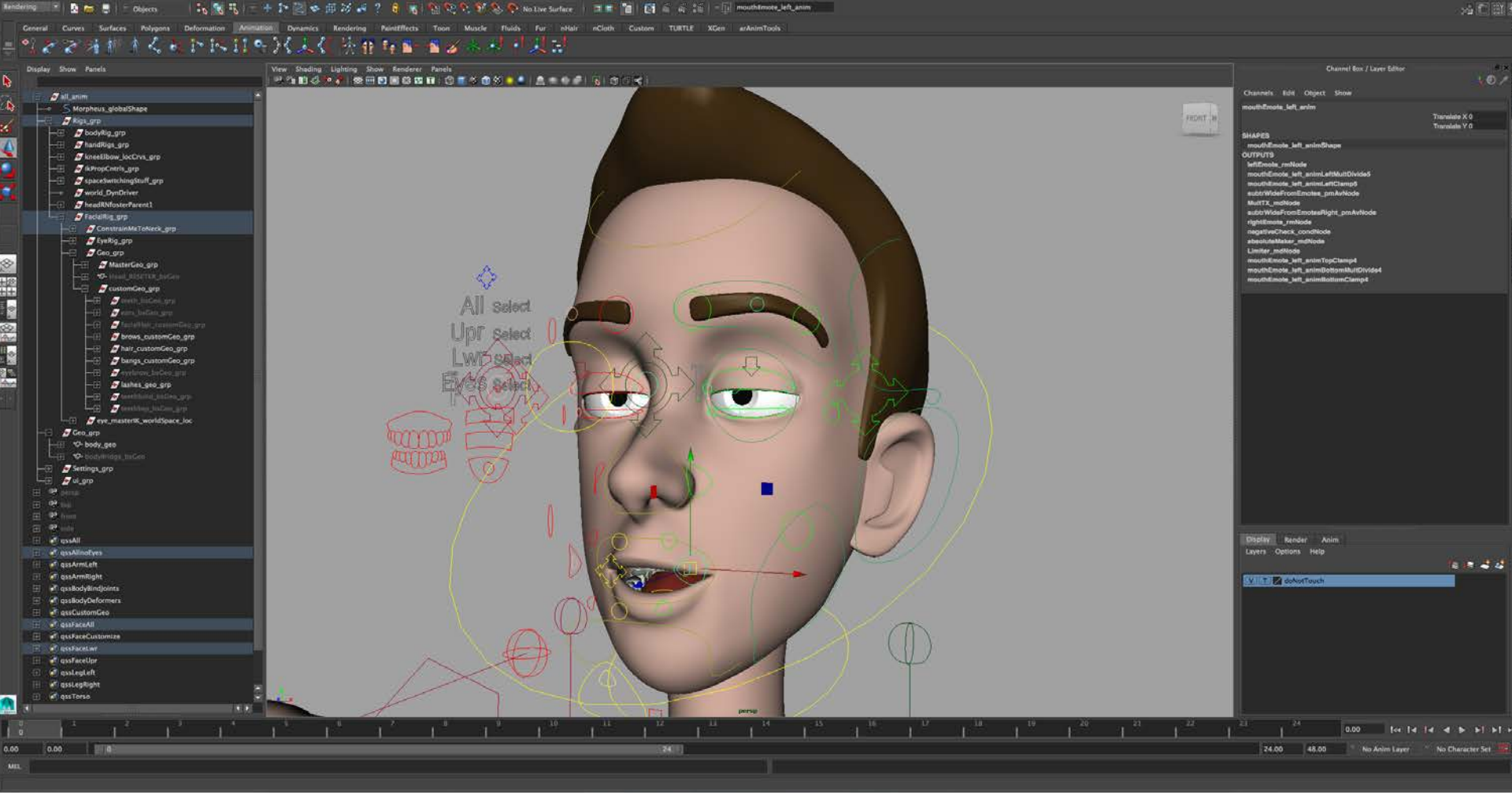


Hierarchical Behaviors (Generative)



Learning to Optimize

Smooth Imitation Learning



- **Animation artists spend ≥50% time on face**
  - Mostly eyes & mouth
  - Very tedious

We'll focus on mouth & speech.

Sarah Taylor

Prediction Task

Taehwan Kim

Input sequence $X =< x_1, x_2, \ldots, x_{|x|} >$

Output sequence $Y =< y_1, y_2, \ldots, y_{|y|} > \ , y_t \in R^D$

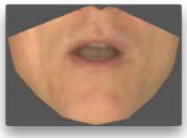
Goal: learn predictor $h : X \rightarrow Y$

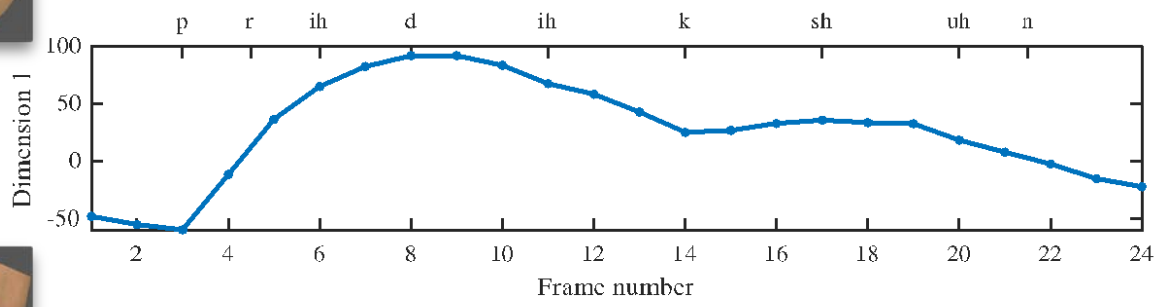
X

Frame	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Token	-	p	p	r	ih	ih	d	d	ih	ih	ih	ih	k	k	sh	sh	sh	sh	uh	uh	n	-

Phoneme sequence

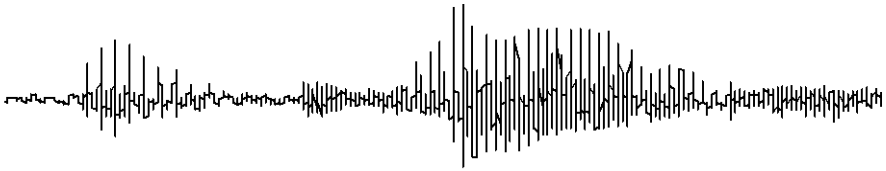
h





Y

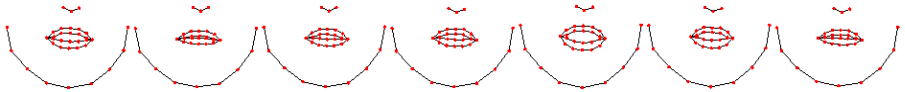
Sequence of face configurations



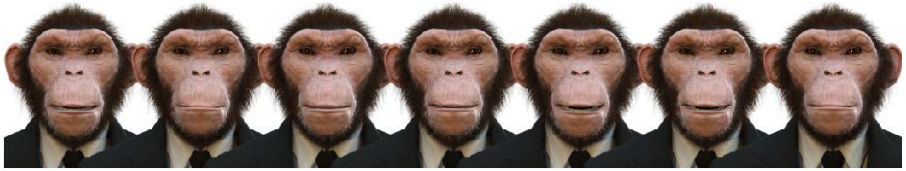
Input Audio

s s s s s ih ih ih g g r r ae ae ae ae f f f

Speech Recognition



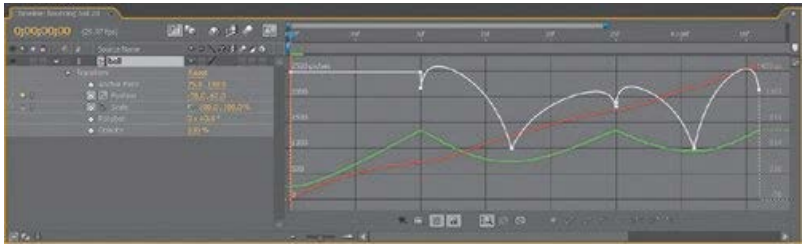
Speech Animation



(chimp rig courtesy of Hao Li)

Retargeting

E.g., [Sumner & Popovic 2004]



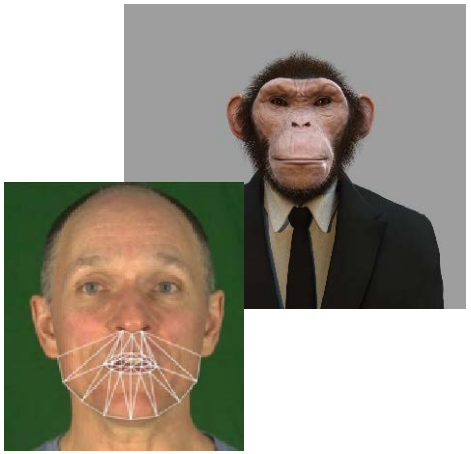
Editing

Sarah Taylor

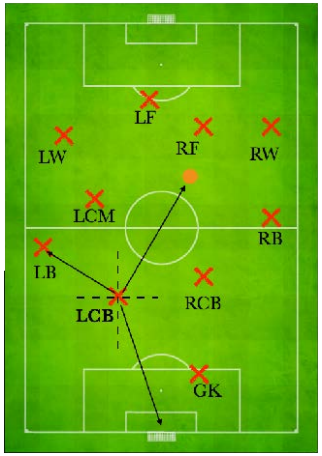
Taehwan Kim

A Decision Tree Framework for Spatiotemporal Sequence Prediction
Taehwan Kim, Yisong Yue, Sarah Taylor, Iain Matthews. KDD 2015
A Deep Learning Approach for Generalized Speech Animation
Sarah Taylor, Taehwan Kim, Yisong Yue, et al. SIGGRAPH 2017

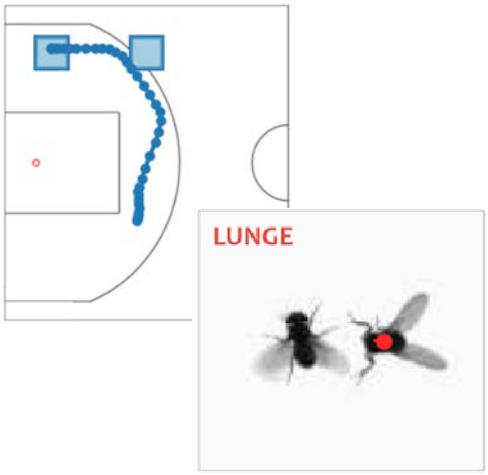
Behind the Scenes of Pandora - The World of Avatar https://youtu.be/URSOqWtLix4



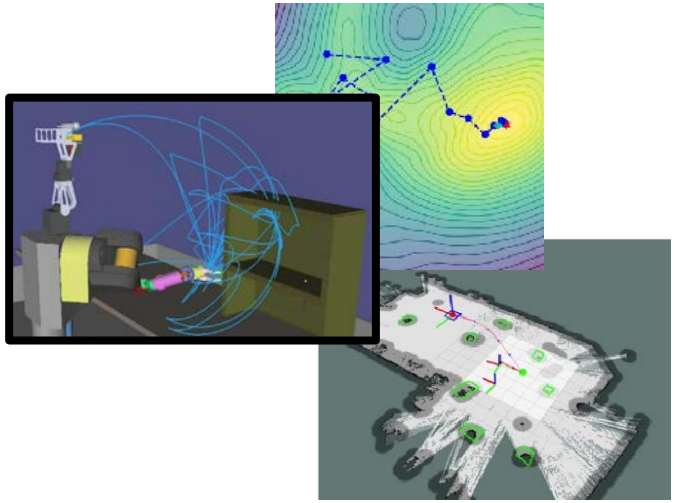
Speech Animation



Coordinated Learning



Hierarchical Behaviors (Generative)



Learning to Optimize

Smooth Imitation Learning

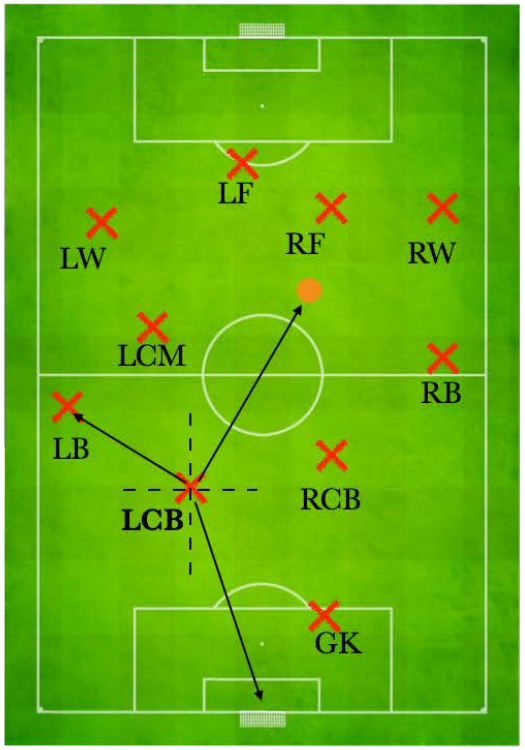
C

Data-Driven Ghosting using Deep Imitation Learning

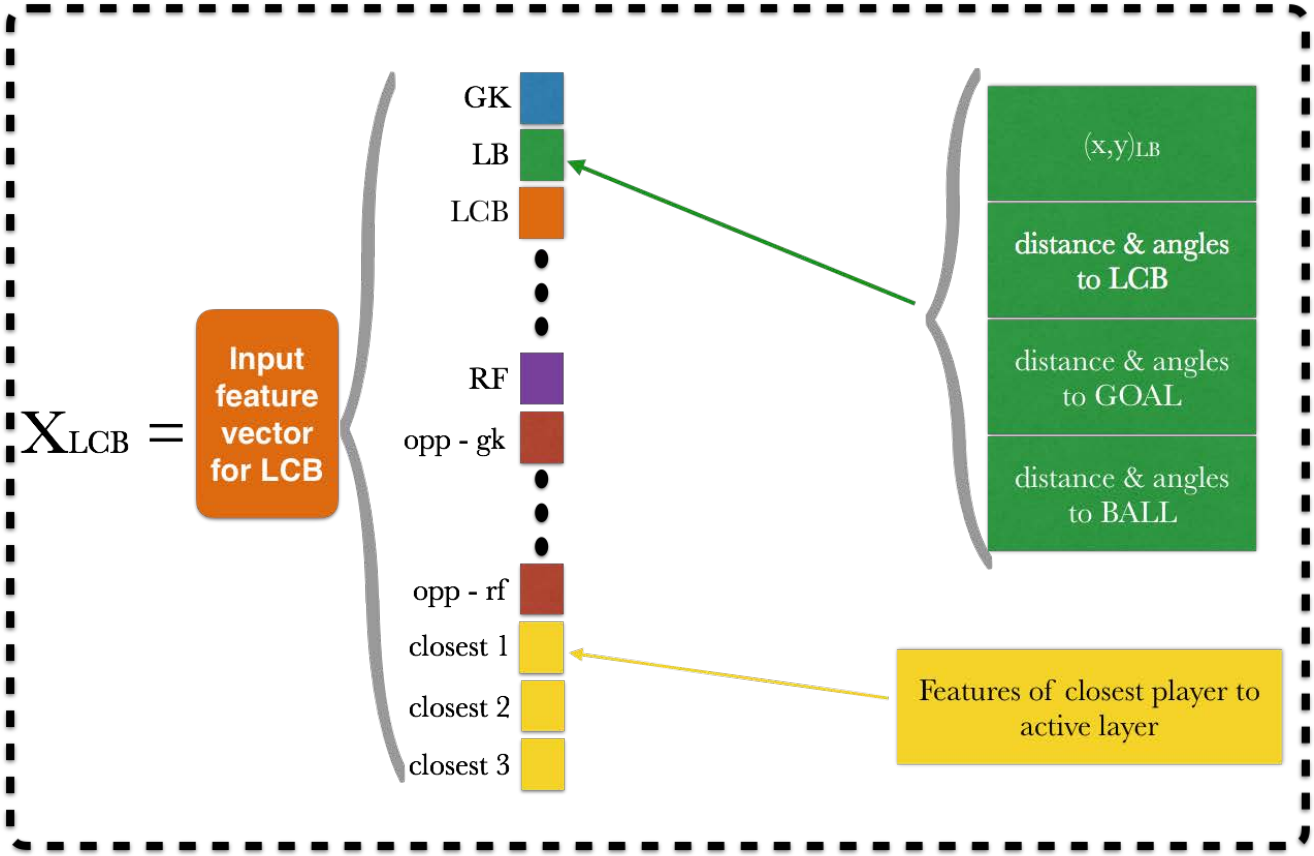
Hoang Le, Peter Carr, Yisong Yue, Patrick Lucey. SSAC 2017

N

State Representation



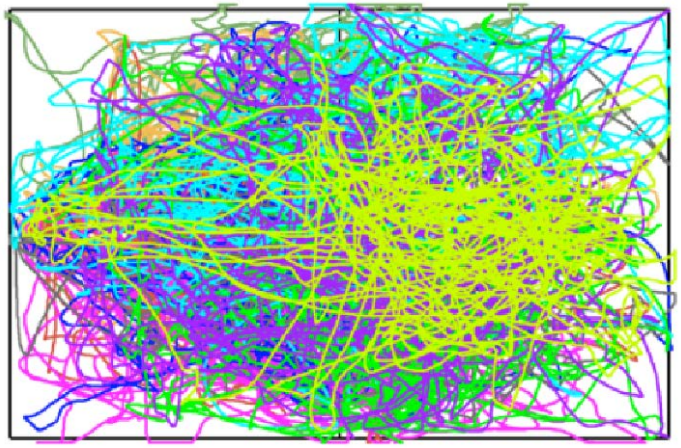
Geometric features computed



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

But Who Plays Which Role?

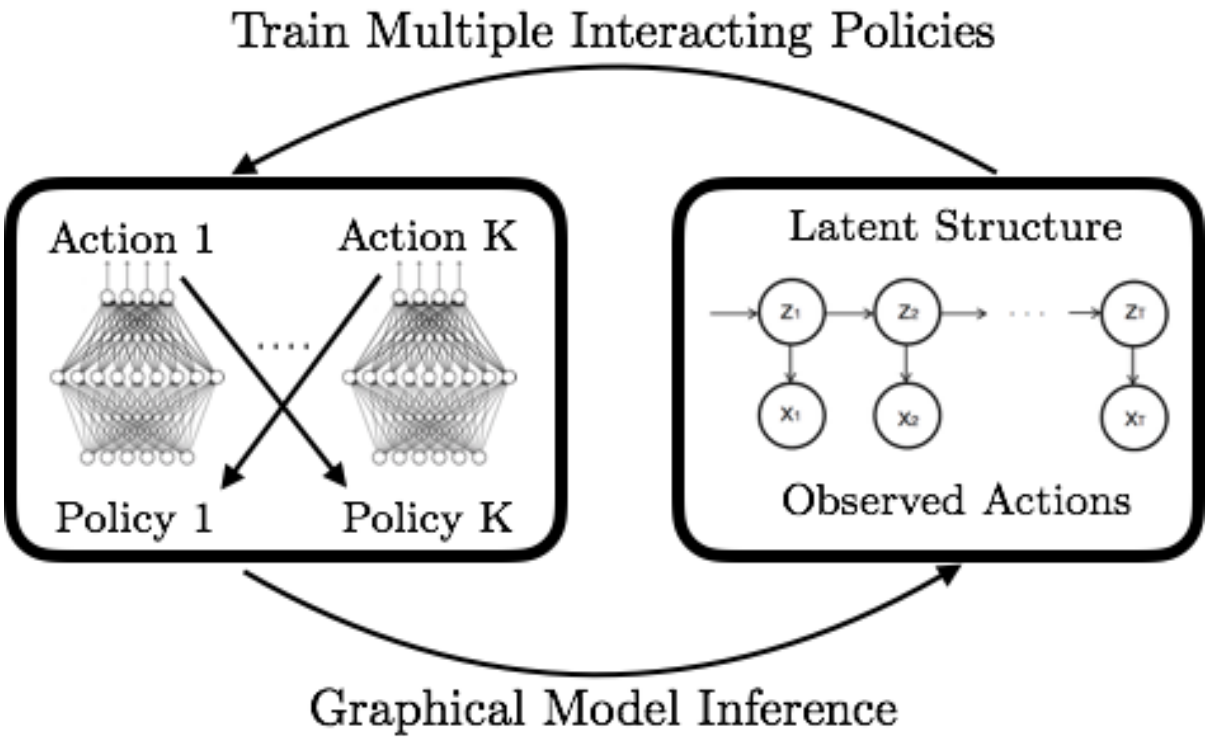
- All we get are trajectories!
  - Don’t know which belongs to which role.



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

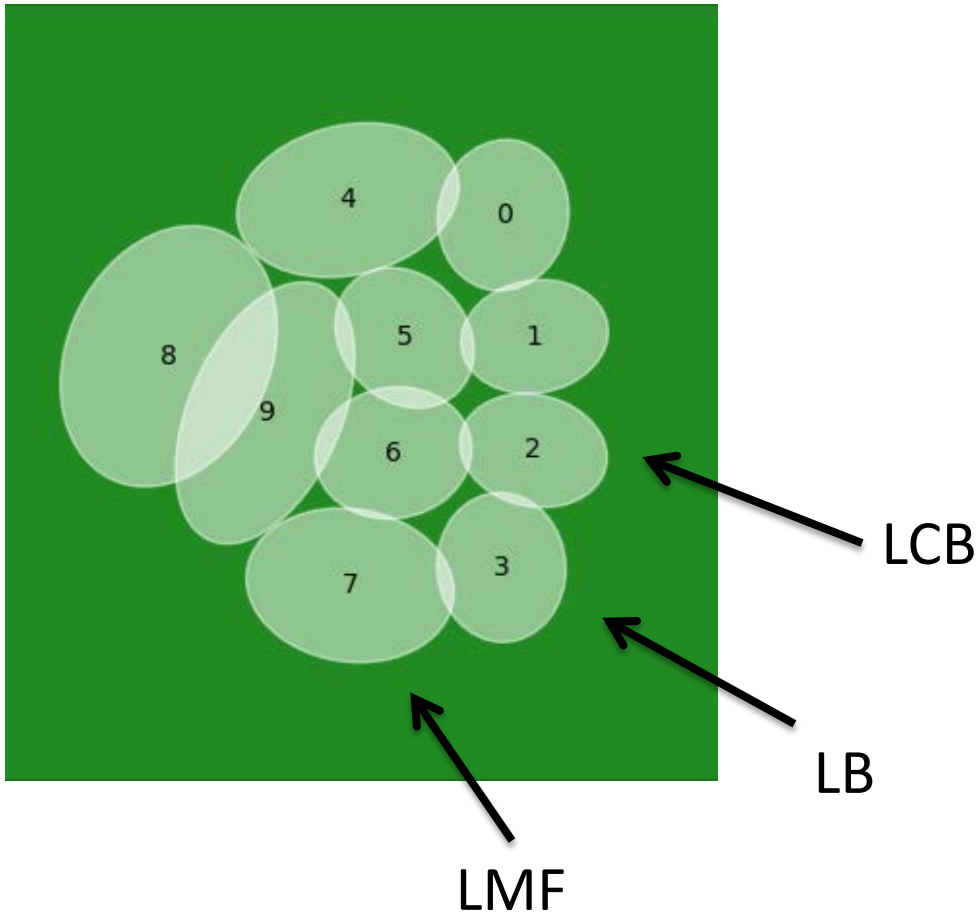
Hoang Le

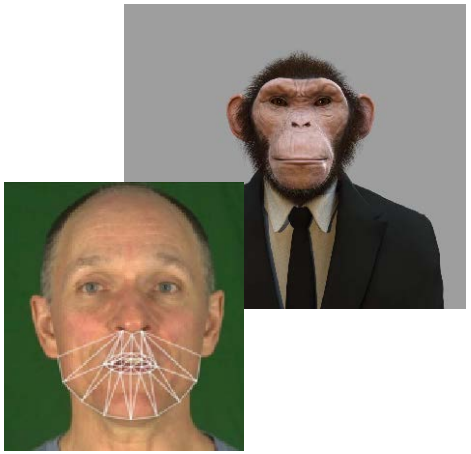
Coordination Model



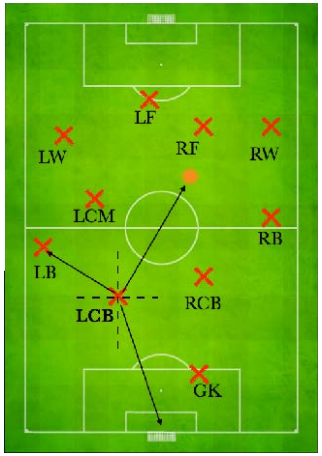
Coordinated Multi-Agent Imitation Learning
Hoang Le, Yisong Yue, Peter Carr, Patrick Lucey. ICML 2017

Learned Roles

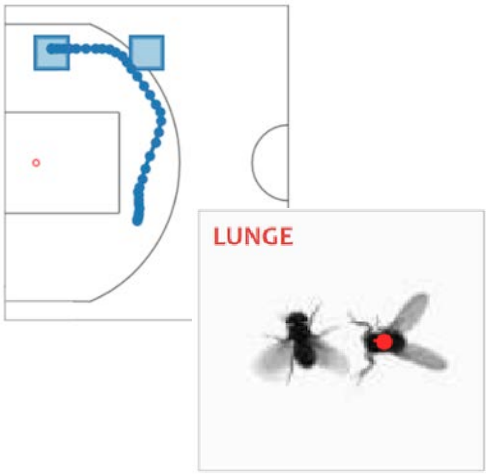




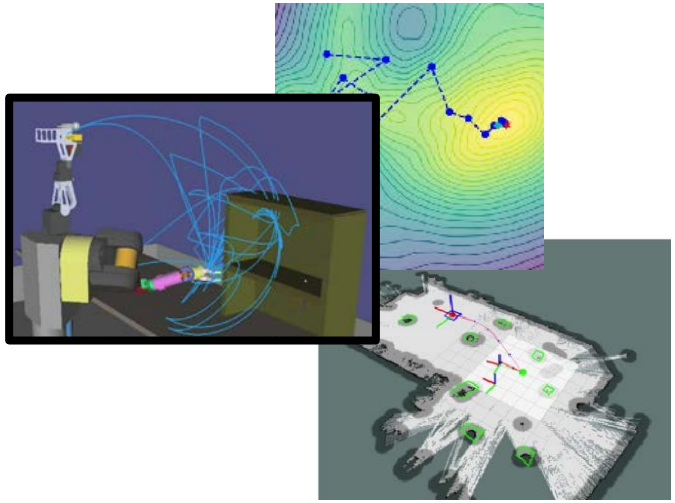
Speech Animation



Coordinated Learning



Hierarchical Behaviors (Generative)

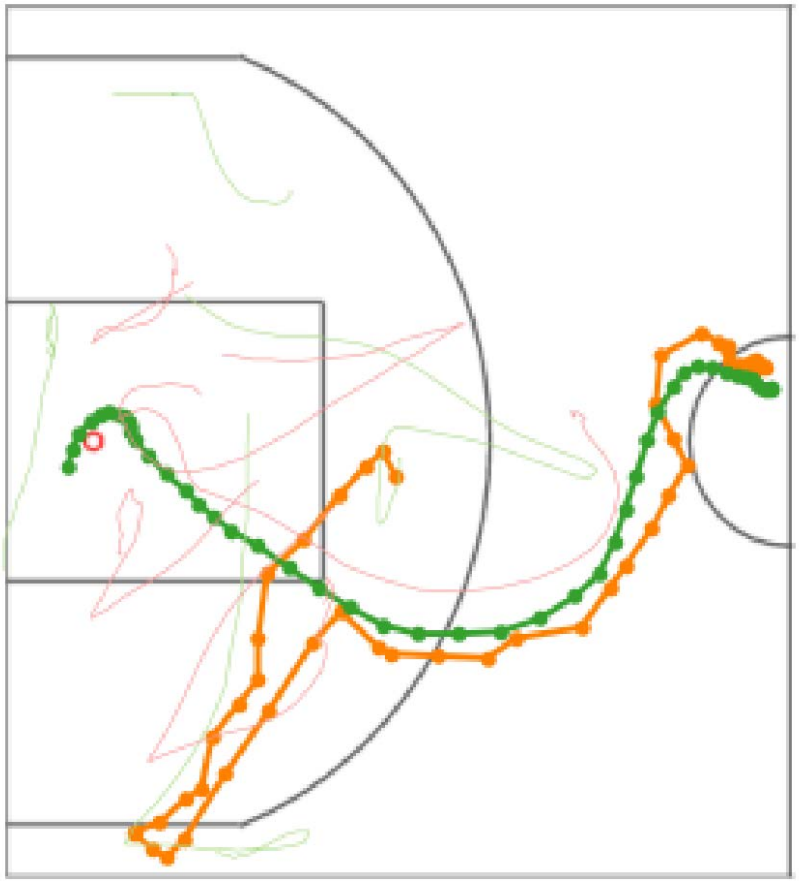


Learning to Optimize

Smooth Imitation Learning

Strategy vs Tactics

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



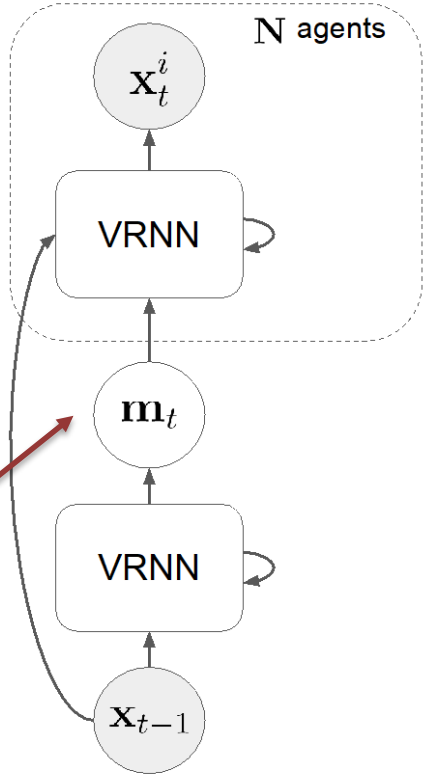
Stephan Zheng

Eric Zhan

Generative + Hierarchical Imitation Learning

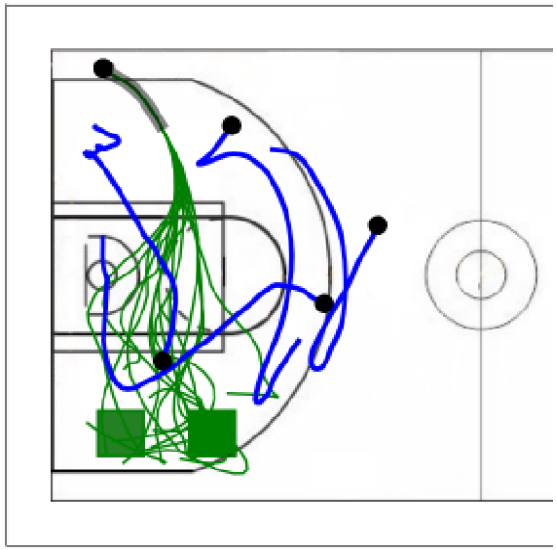
- **Generative Imitation Learning**
  - No single "correct" action
- **Hierarchical**
  - Make predictions at multiple resolutions

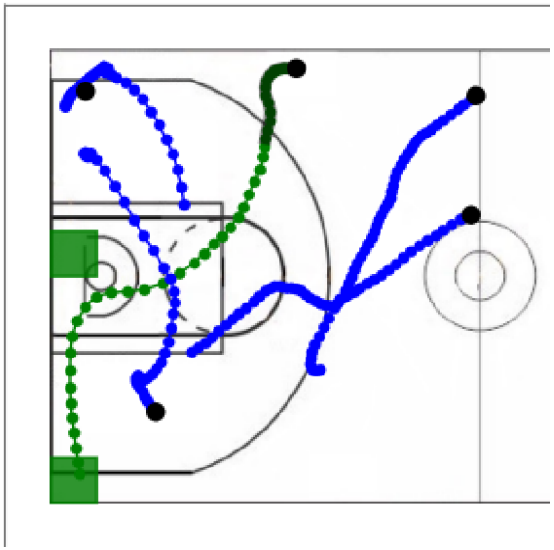
Macro-goals



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, Patrick Lucey. (under review)

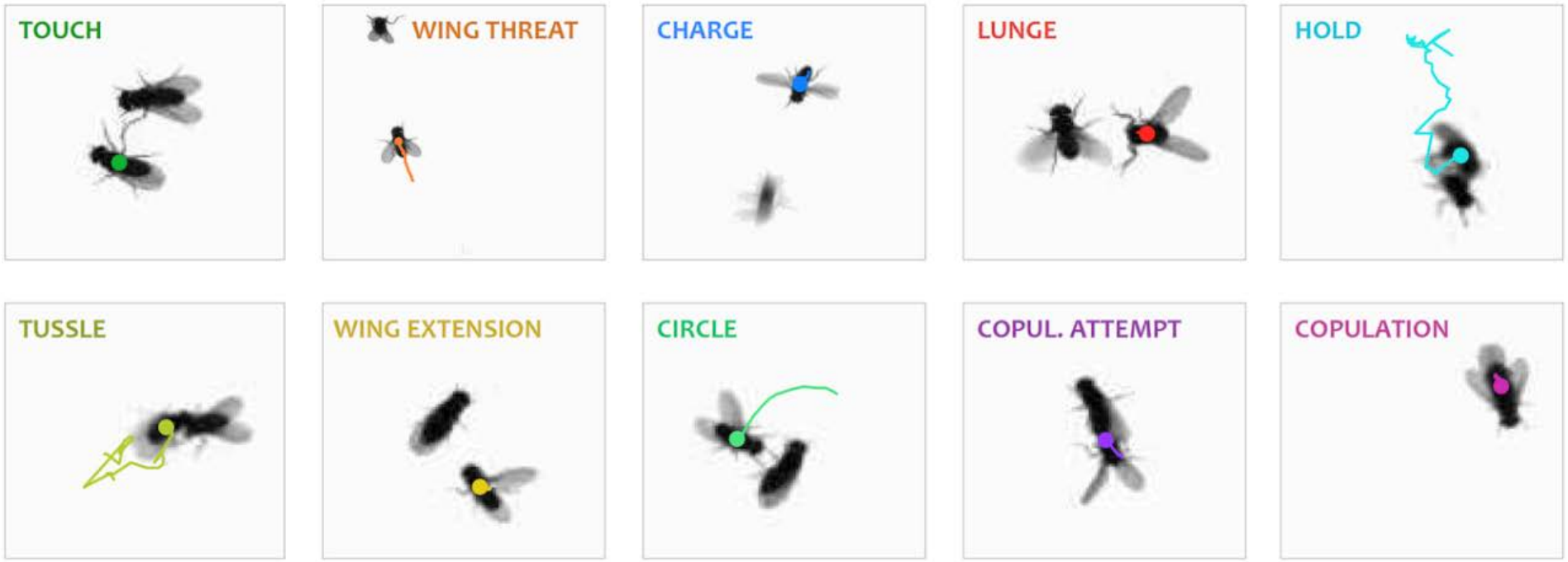




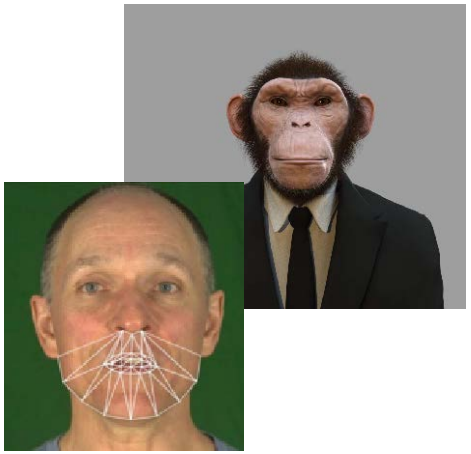
Eyrun
Eyolfsdottir

Drosophila Behavior

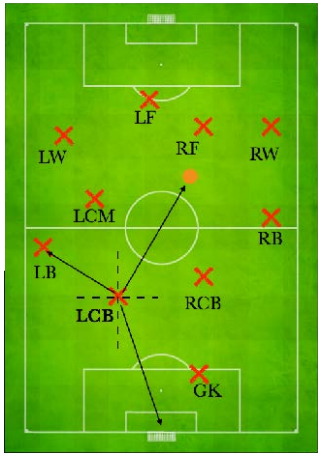
Activity Labels



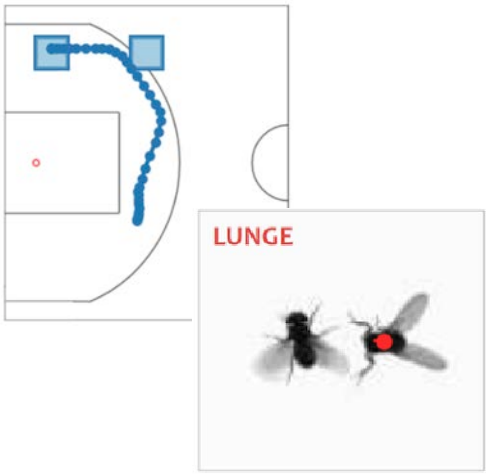
Learning recurrent representations for hierarchical behavior modeling
Eyrun Eyolfsdottir, Kristin Branson, Yisong Yue, Pietro Perona, ICLR 2017



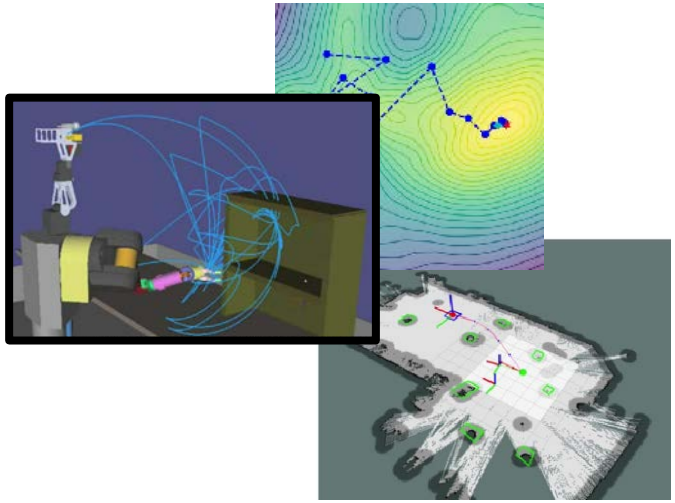
Speech Animation



Coordinated Learning



Hierarchical Behaviors (Generative)



Learning to Optimize

Smooth Imitation Learning

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: $(x{=}\text{MILP}, y{=}\text{solution/search}{-}\text{trace})$
- 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]

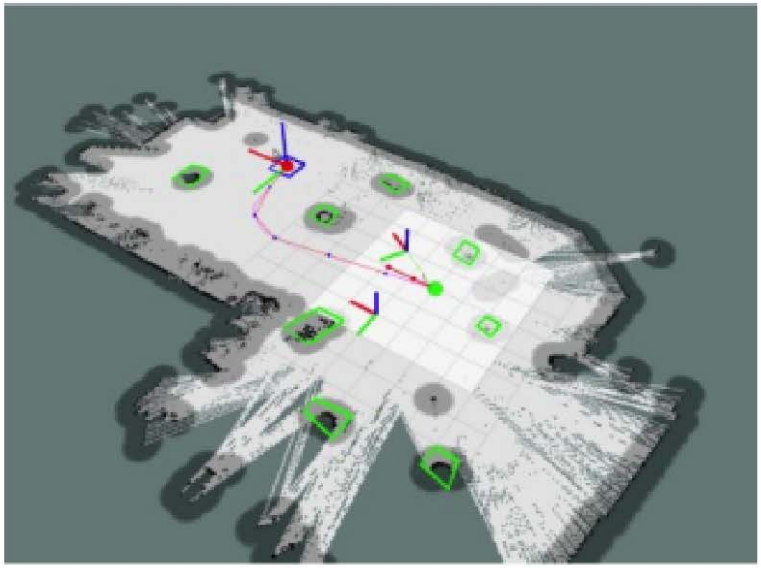

Joe Marino

Ravi Lanka

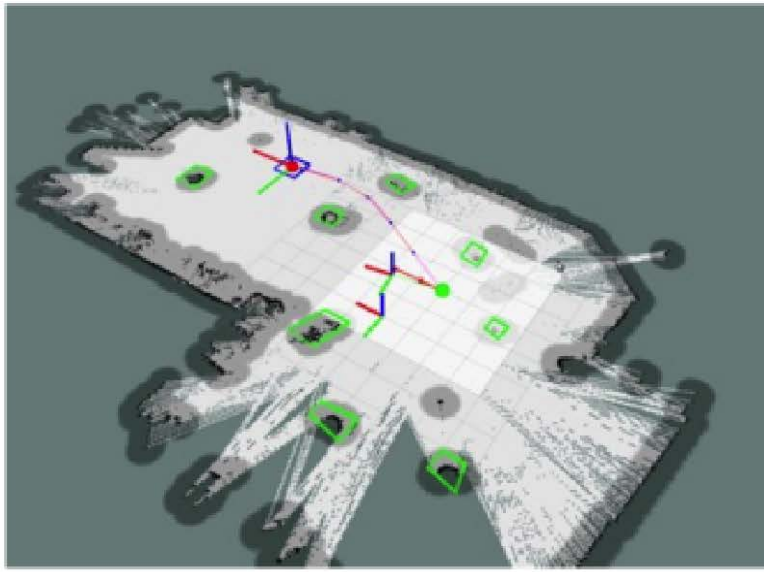
Ongoing Research

Risk-Aware Planning

Jialin Song



Low Risk

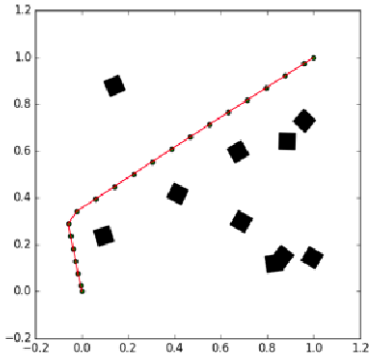


High Risk

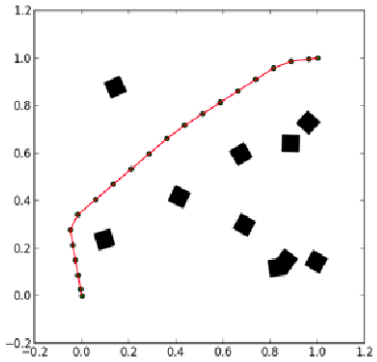
- Compiled as mixed integer program
- Challenging optimization problem

NASA JPL

Preliminary Results



Optimal Solution
(Gurobi solver)



Our Approach

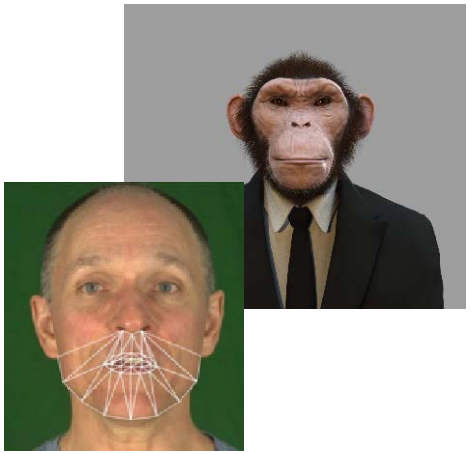
	Ours	Gurobi Solver
Train	1049	15241
Test	1127	25249

Avg Nodes Explored

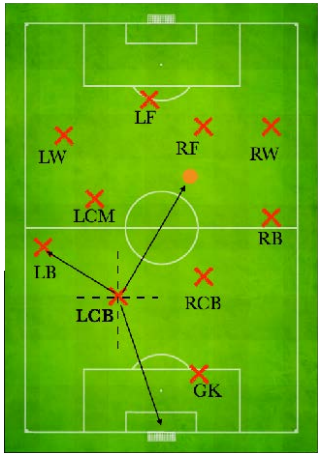
	Ours	Gurobi Solver
Train	0.732	0.305
Test	0.577	0.309

Avg Objective value

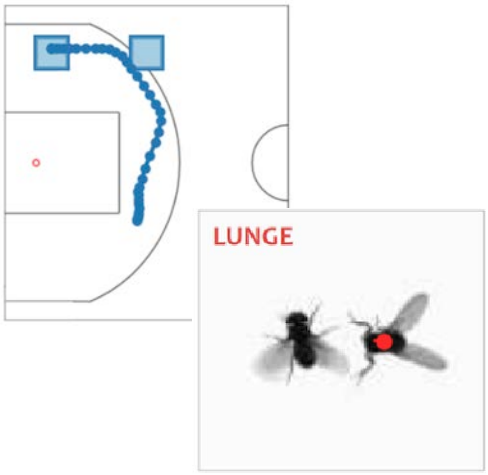
Learning to Search via Retrospective Imitation
R. Lanka, J. Song, A. Zhao, Y. Yue, M. Ono. (under review)



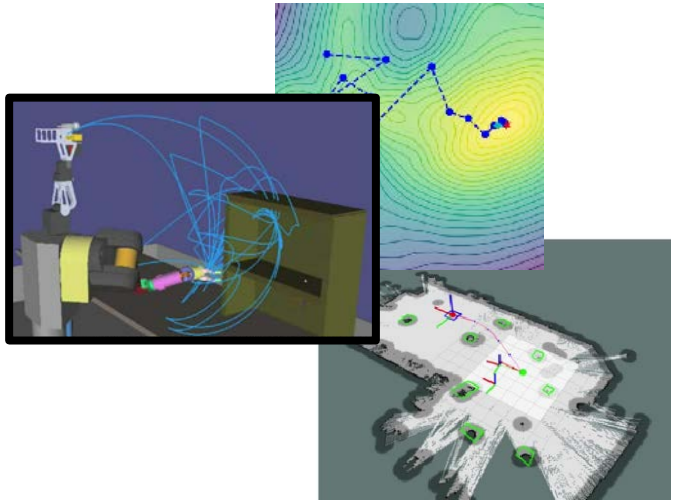
Speech Animation



Coordinated Learning



Hierarchical Behaviors (Generative)

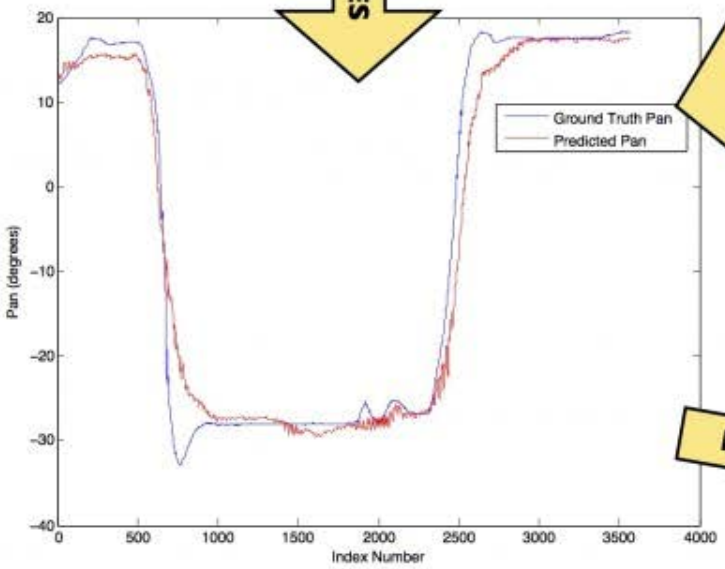


Learning to Optimize

Smooth Imitation Learning

Realtime Player Detection and Tracking

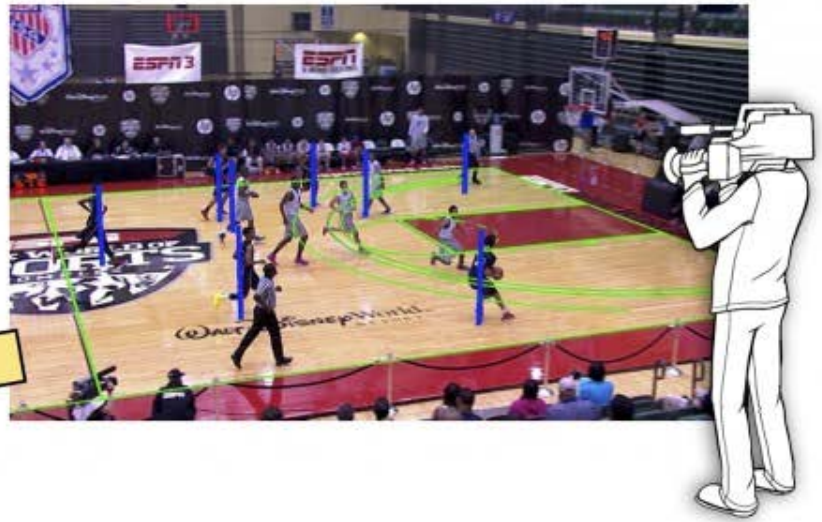
FEATURES



Learned Regressor

TRAIN

Human Operated Camera



PREDICT

Autonomous Robotic Camera

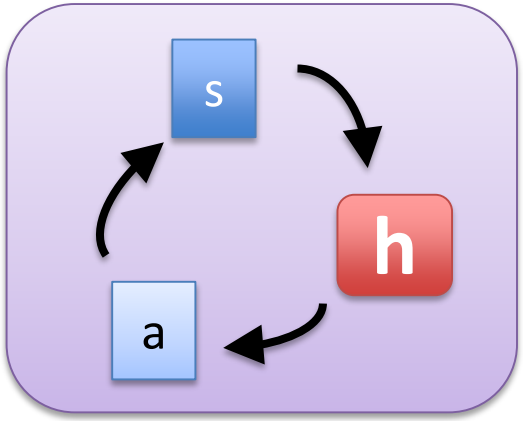
Disney Research

Problem Formulation

- Input: stream of x_t
  - E.g., noisy player detections

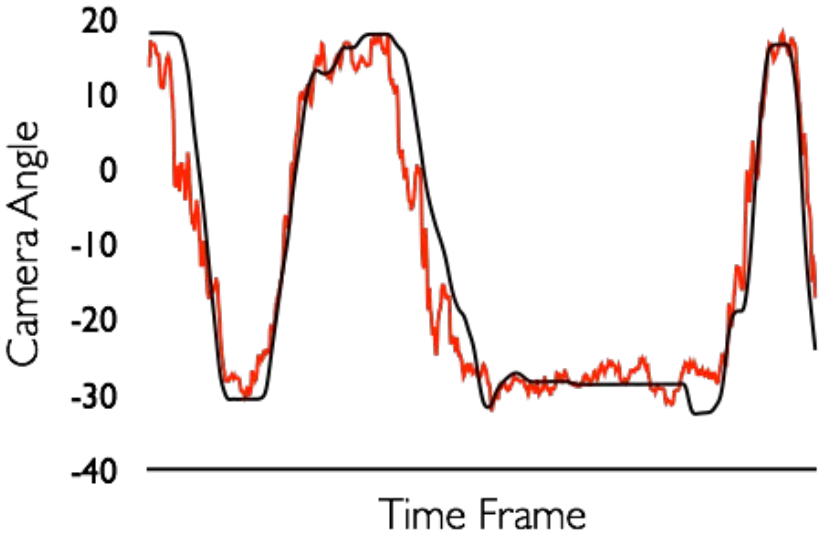
- State $s_t = (x_{t:t-K}, a_{t-1:t-K})$
  - Recent detections and actions

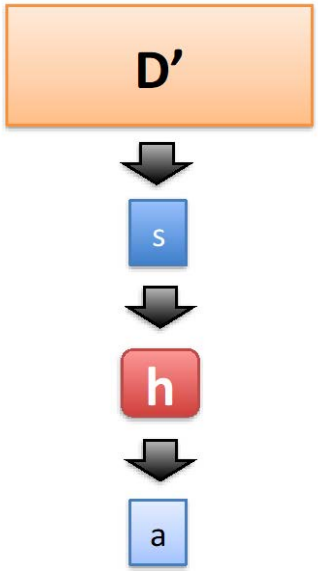
- Goal: learn $h(s_t) \rightarrow a_t$
  - Imitate expert



Naïve Approach

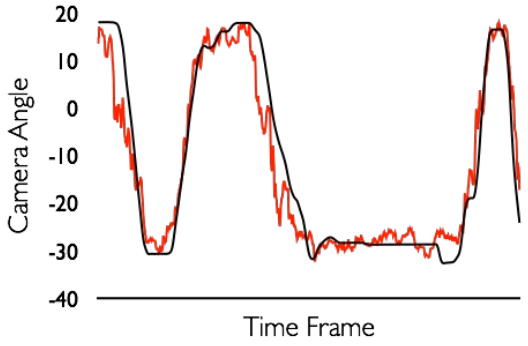
- Supervised learning of demonstration data
  - Train predictor per frame
  - Predict per frame



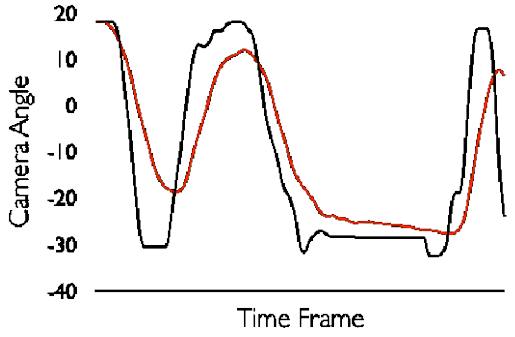


What is the Problem?

- Basically takes “infinite” training data to train smooth model.
  - Via input/output examples



- In practice, people do post-hoc smoothing



Cannot Rely 100% on Learning!

- People have models of smoothness!
  - Kalman Filters
  - Linear Autoregressors
  - Etc...
- Pure ML approach throws them away!
  - ”black box”

Hybrid Model-Based + Black-Box

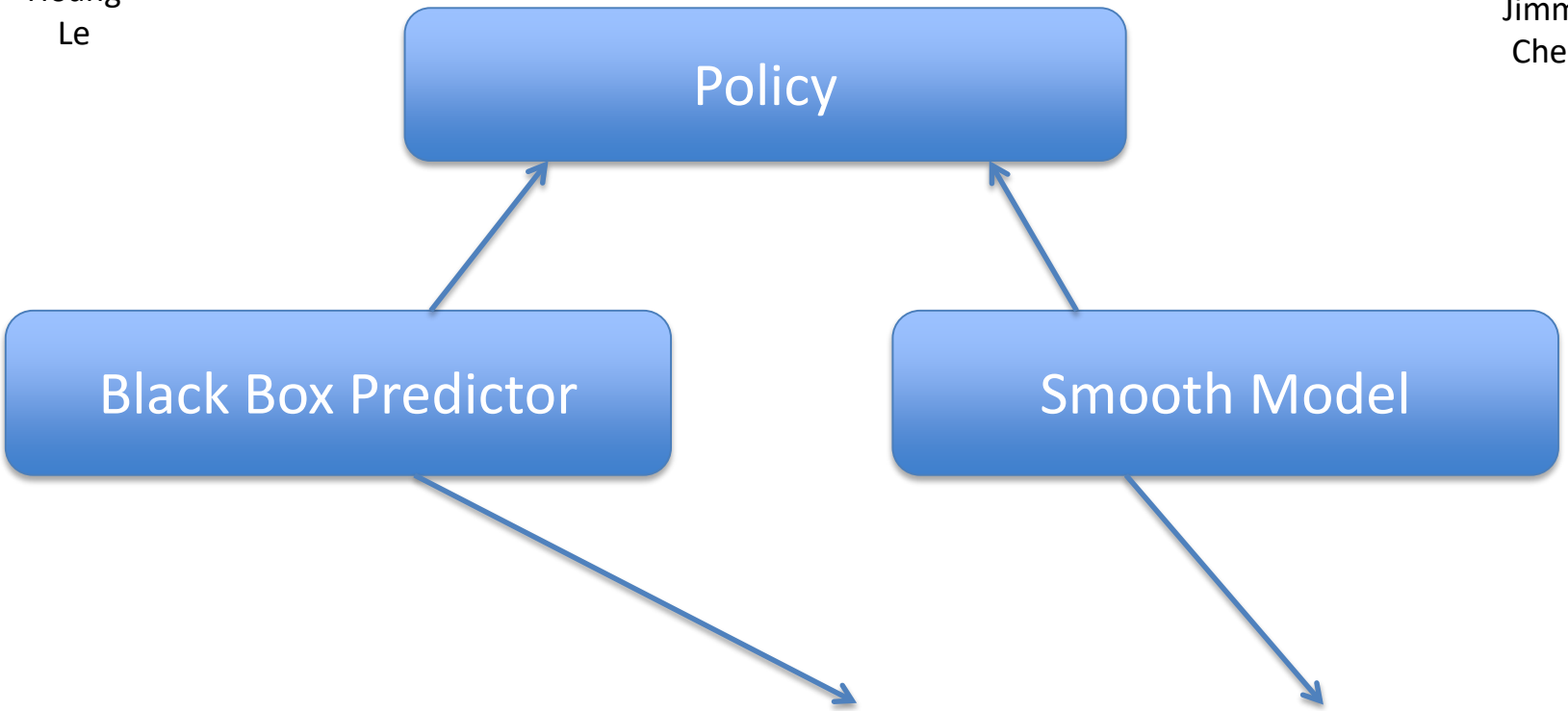
- Model-based approaches
  - Strong assumptions, well specified
  - Lacks flexibility
  - E.g., Kalman Filter, Linear Autoregressor
- Black-box approaches
  - Assumption free, underspecified
  - Requires a lot of training data
  - E.g., random forest, deep neural network
- **Best of both worlds?**

Conventional Models

Hoang Le

New Policy Class

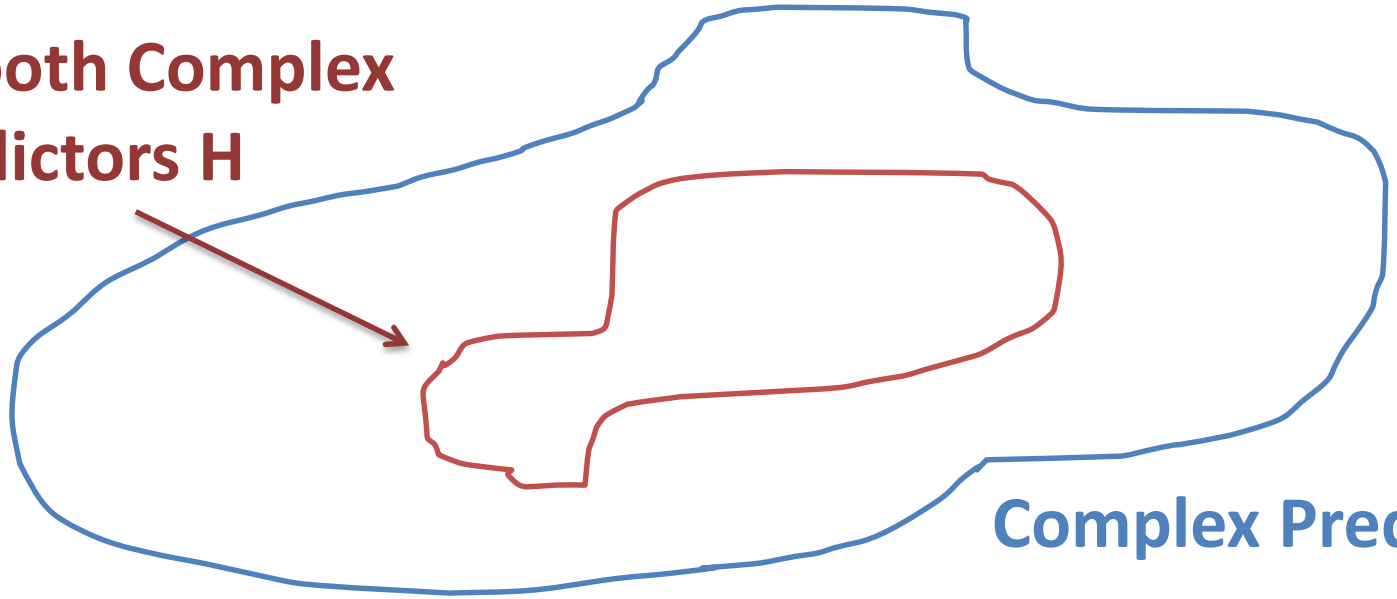
Jimmy Chen



$$h\left(s_t \equiv (x_{t:t-K}, a_{t-1:t-K})\right) = argmin_{a'}(f(s_t) - a')^2 + \lambda(g(a_{t-1:t-K}) - a')^2$$

$$= \frac{f(s_t) + \lambda g(a_{t-1:t-K})}{1+\lambda}$$

Functional Regularization

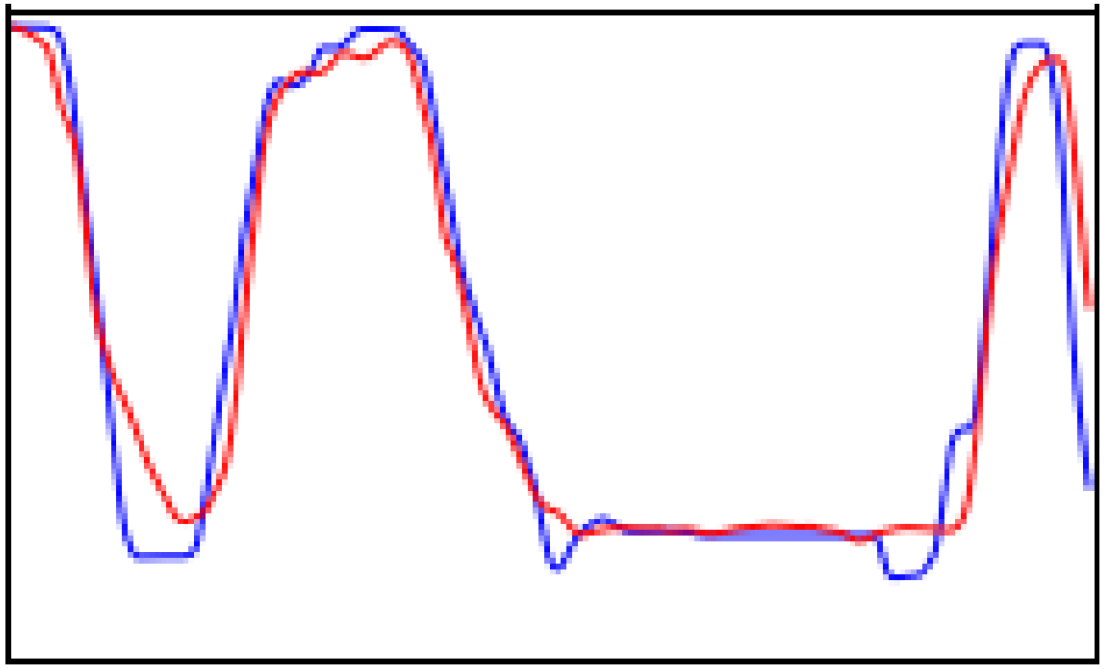


$$h\left(s_t \equiv (x_{t:t-K}, a_{t-1:t-K})\right) = argmin_{a'}(f(s_t) - a')^2 + \lambda(g(a_{t-1:t-K}) - a')^2$$

$$= \frac{f(s_t) + \lambda g(a_{t-1:t-K})}{1+\lambda}$$

Smooth Imitation Learning for Online Sequence Prediction
Hoang Le, Andrew Kang, Yisong Yue, Peter Carr. ICML 2016

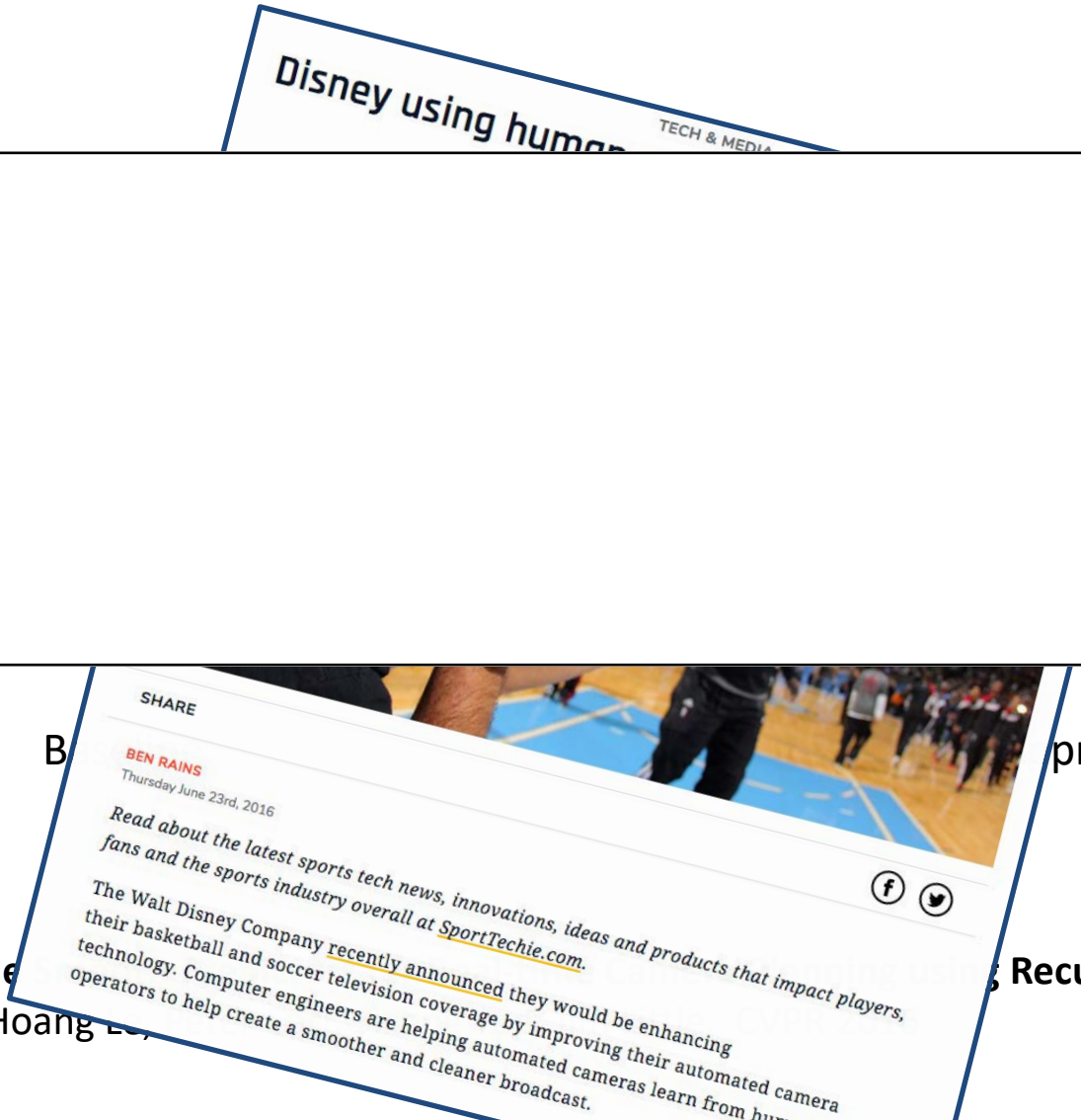
Our Result



$$h\left(s_t \equiv (x_{t:t-K}, a_{t-1:t-K})\right) = \frac{f(s_t) + \lambda g(a_{t-1:t-K})}{1+\lambda}$$

Smooth Imitation Learning for Online Sequence Prediction
Hoang Le, Andrew Kang, Yisong Yue, Peter Carr. ICML 2016

Qualitative Comparison



B proach

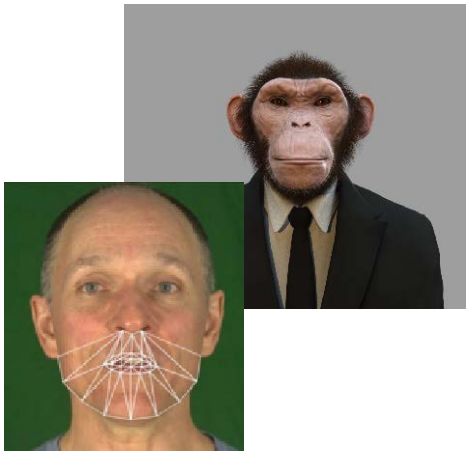
Learning Online Recurrent Decision Trees
Jianhui Chen, Hoang

Lessons Learned

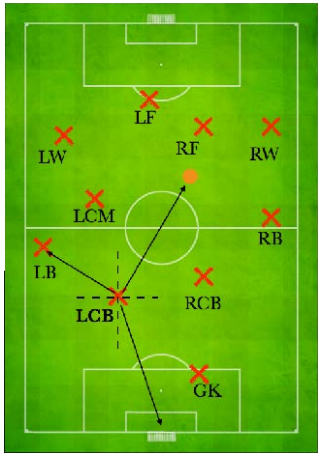
- **Intuition:** Let model do most of work
  - Black box (deep neural net) adds flexibility
  - "Regularization" improves learning
    - Exponentially faster convergence compared to SEARN

Exploit Lipschitz from smooth temporal dynamics

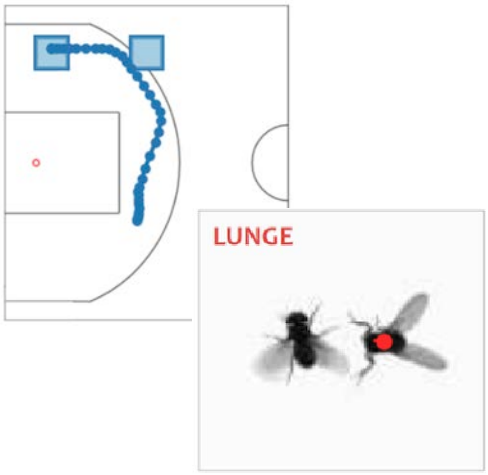
- Applicable to other approaches?
  - Deep learning + robust control?
    - w/ Aaron Ames @Caltech



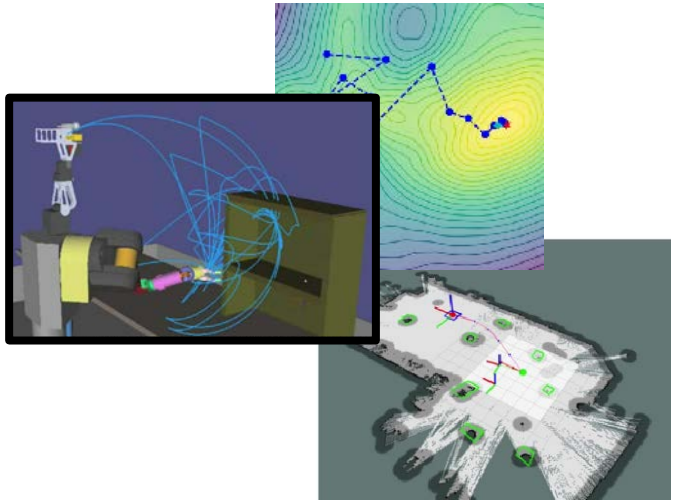
Speech Animation



Coordinated Learning



Hierarchical Behaviors (Generative)

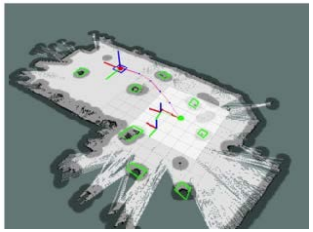


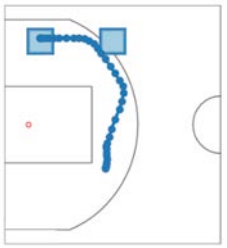
Learning to Optimize

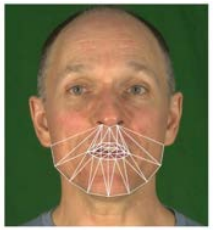
Smooth Imitation Learning

New Frontiers in Imitation learning

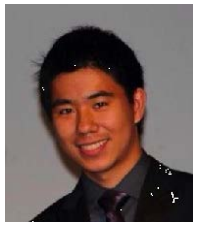
- **Incorporating Structure**
  - Smoothness of output space
  - Latent structure of input space
  - New feedback oracles
- **New Algorithmic Frameworks**
  - Black Box + Dynamics Models
  - Black Box + Graphical Models
  - Retrospective Imitation Learning
- **Cool Applications!**







Eyrun Eyolfsdottir



Eric Zhan

Stephan Zheng

Hoang Le

Taehwan Kim

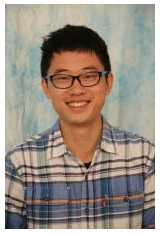
Stephane Ross

Jialin Song

Joe Marino

Andrew Kang

Debadeepta Dey



Robin Zhou

Albert Zhao

Jimmy Chen

Sarah Taylor

Ravi Lanka

Kristin Branson

Iain Matthews

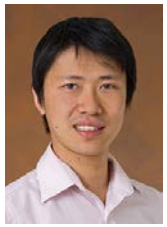
Jim Little

Pietro Perona

Patrick Lucey

Drew Bagnell

Peter Carr



Masahiro Ono

Stephan Mandt

References

http://www.yisongyue.com

Smooth Imitation Learning for Online Sequence Prediction

Hoang Le, Andrew Kang, Yisong Yue, Peter Carr. ICML 2016

Learning Smooth Online Predictors for Real-Time Camera Planning using Recurrent Decision Trees

Jianhui Chen, Hoang Le, Peter Carr, Yisong Yue, Jim Little. CVPR 2016

A Decision Tree Framework for Spatiotemporal Sequence Prediction

Taehwan Kim, Yisong Yue, Sarah Taylor, Iain Matthews. KDD 2015

A Deep Learning Approach for Generalized Speech Animation

Sarah Taylor, Taehwan Kim, Yisong Yue, Moshe Mahler, James Krahe, Anastasio Rodriguez, Jessica Hodgins, Iain Matthews. SIGGRAPH 2017

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

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

Learning Policies for Contextual Submodular Prediction

Stephane Ross, Jiaji Zhou, Yisong Yue, Debadeepta Dey, J. Andrew Bagnell. ICML 2013