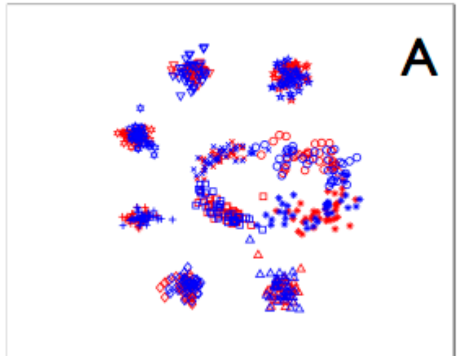
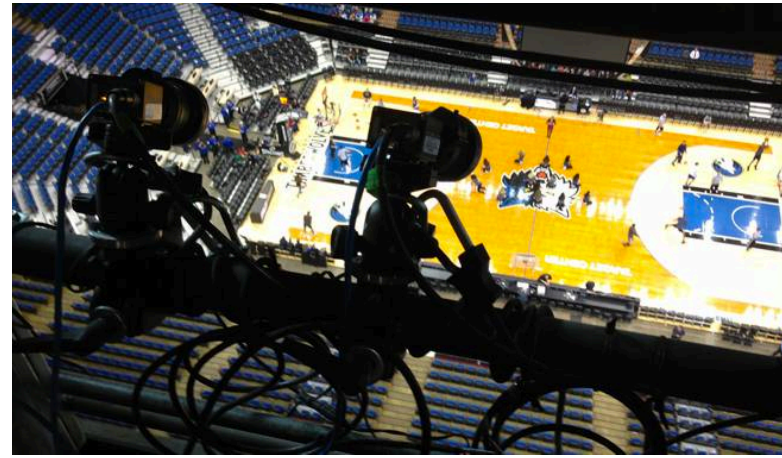
Caltech

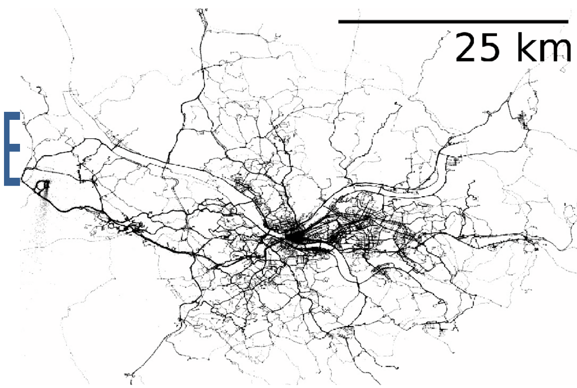
New Frontiers in Imitation Learning

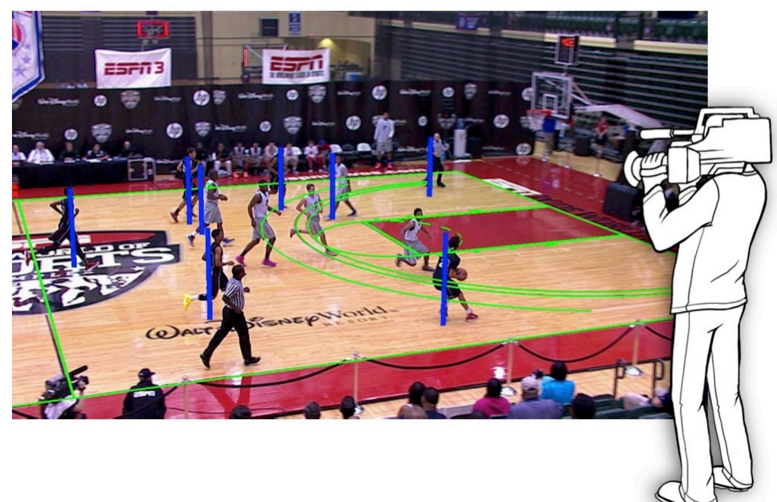
Yisong Yue


A




LUNGE


25 km



How to Use Behavioral Tracking Data?

- Analyze & Understand Behavior
  - "Interpretable" Machine Learning
  - Causal Relationships

...and a little bit of this

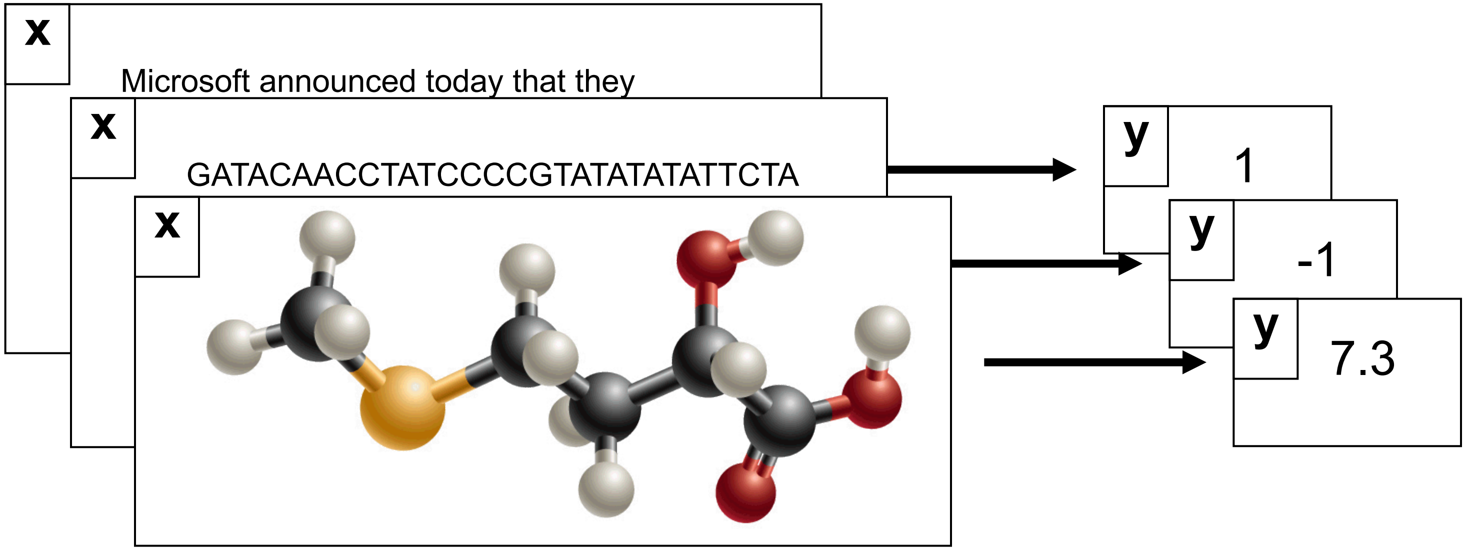
- **Predictive Modeling**
  - **Predict next action (or sequence of actions)**
  - **Multiple predictions**
    - **Multi-agent systems**
    - **Multiple modalities**

Warm Up: Supervised Learning

- Find function from input space X to output space Y

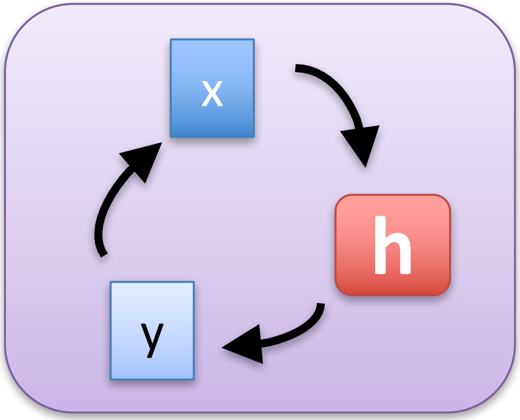
$$h : X \longrightarrow Y$$

such that the prediction error is low.



Imitation Learning

- Input:
  – Sequence of contexts/states:



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

Imitation Learning

Cost

Policy

$$\operatorname*{argmin}_{h} E_{\underbrace{s\sim P_h}_{\text{Estimate Empirically}}}\left[c_s(h(s))\right]$$

States

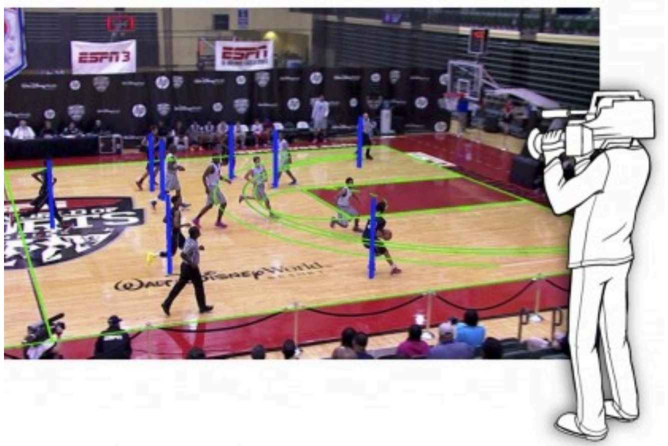
- Violates IID assumption
  - Policy induces state distribution
- Al
  -
  -

- Many approaches
  - SEARN
  - DAgger / DaD / AggreVaTe
  - Apprenticeship

Previous Work:
- Minimal assumptions
- Inefficient in complex & structured settings

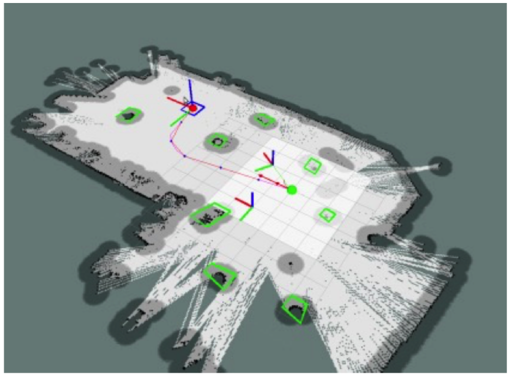
What to Imitate?

Human Demonstrations



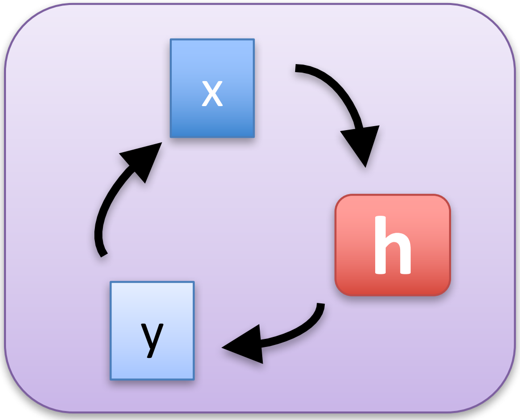
Animal Demonstrations

Computational Oracle



This Talk

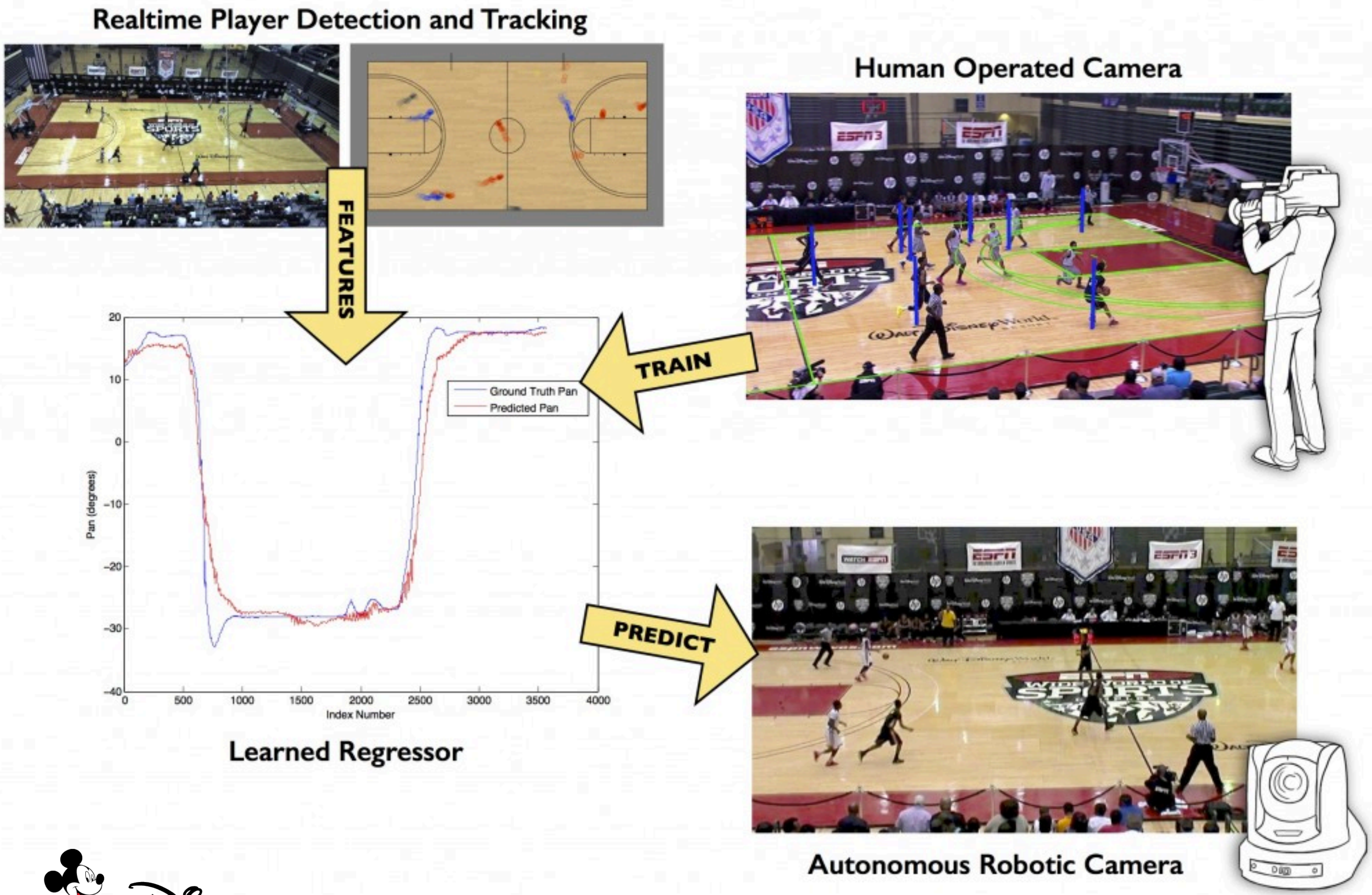
- Example Applications
  - Camera Control
  - Speech Animation
  - Hierarchical Behaviors
  - Multi-Agent Behaviors
  - Learning to Optimize



- Research Questions
  - Structure of Input & Output Spaces

Camera Control

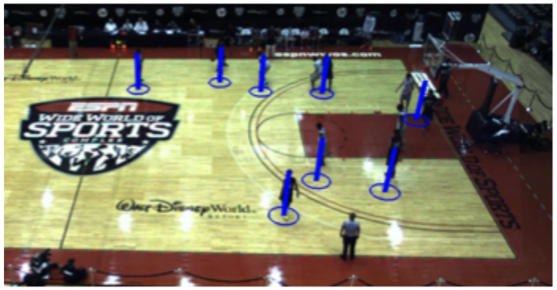
(Smooth Imitation Learning)



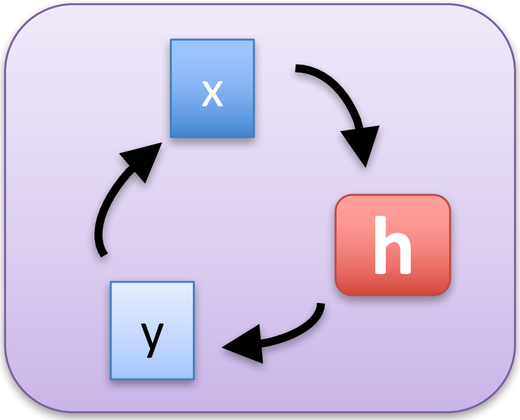
Disney Research

Problem Formulation

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



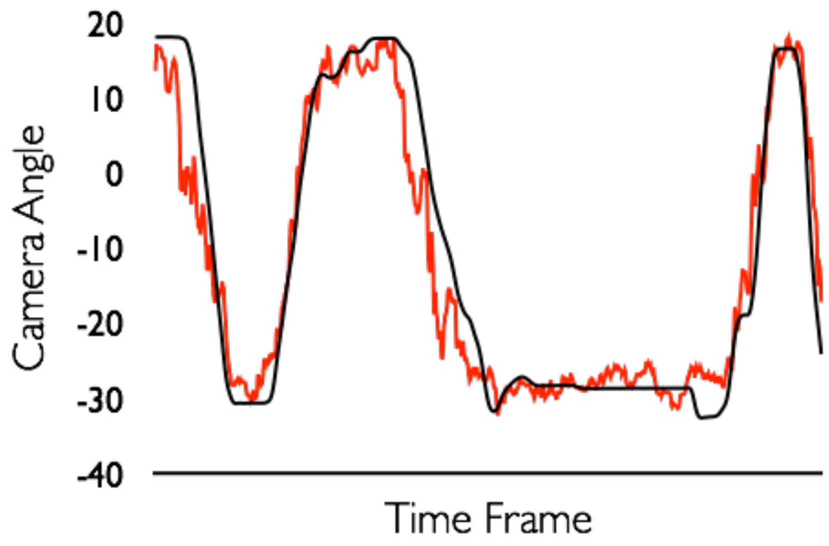
- State s = (x,a)
  - Recent detections and actions



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

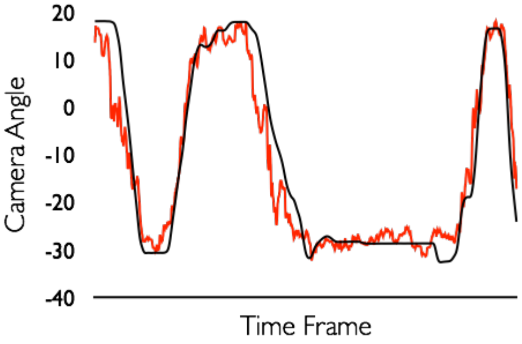
Naïve Approach

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

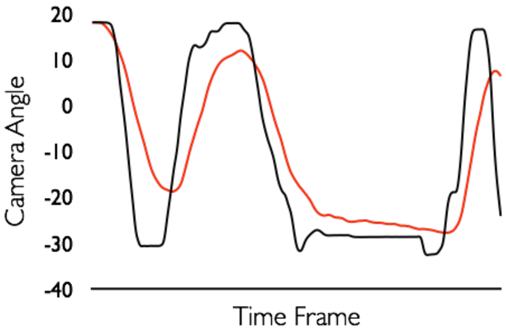


What is the Problem?

- 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

Conventional Models

- Black-box approaches
  - Assumption free, underspecified
  - Requires a lot of training data
  - E.g., random forest, deep neural network
- **Best of both worlds?**

New Policy Class

Hoang
Le

Jimmy
Chen

Policy

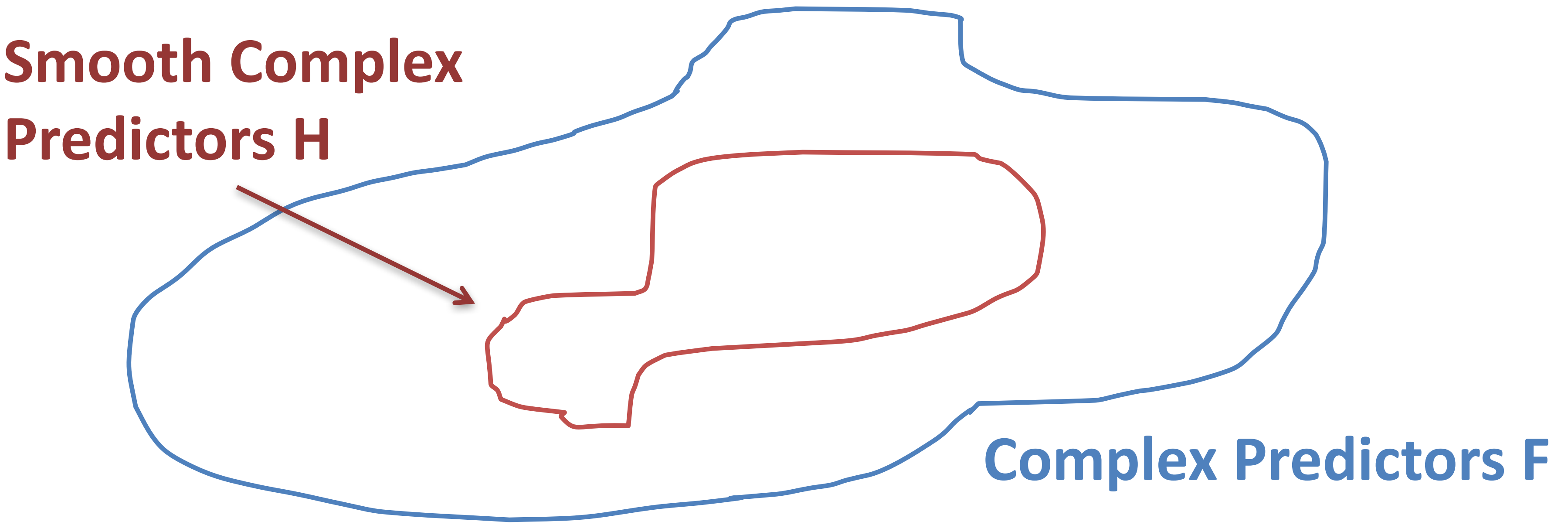
Black Box Predictor

Smooth Model

$$h(s \equiv (x,a)) = \underset{a'}{\operatorname{argmin}} \left(f(s) - a'\right)^2 + \lambda\left(g(a) - a'\right)^2$$

$$= \frac{f(s) + \lambda g(a)}{1+\lambda}$$

Visual Interpretation of Policy Class

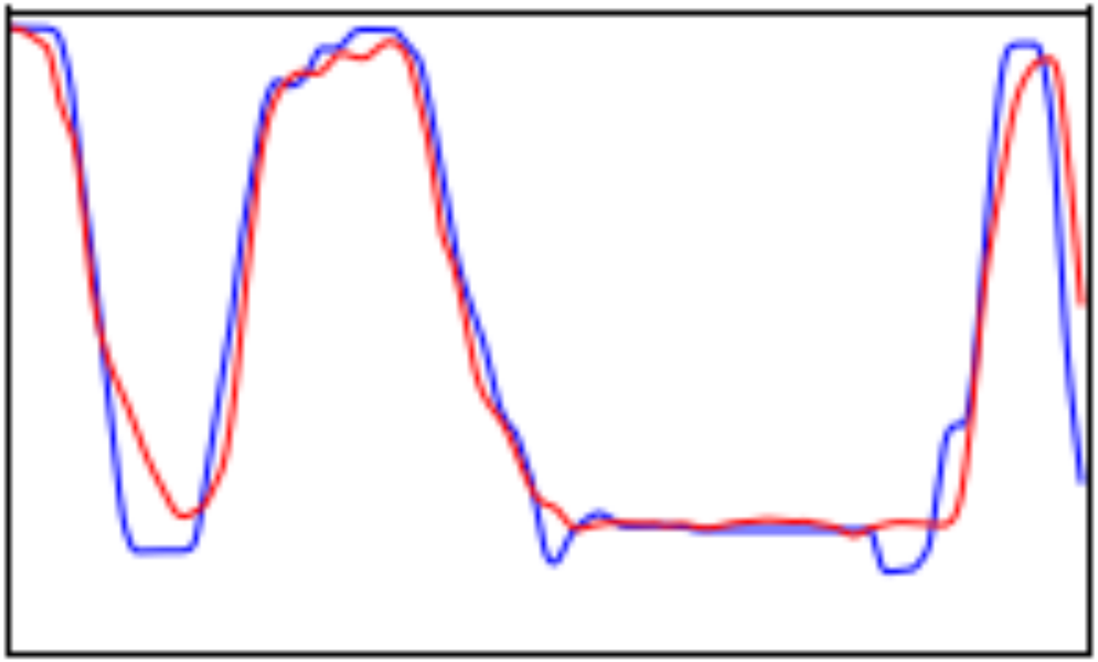


$$h(s \equiv (x,a)) = \underset{a'}{\operatorname{argmin}} \left(f(s) - a'\right)^2 + \lambda\left(g(a) - a'\right)^2$$

$$= \frac{f(s) + \lambda g(a)}{1+\lambda}$$

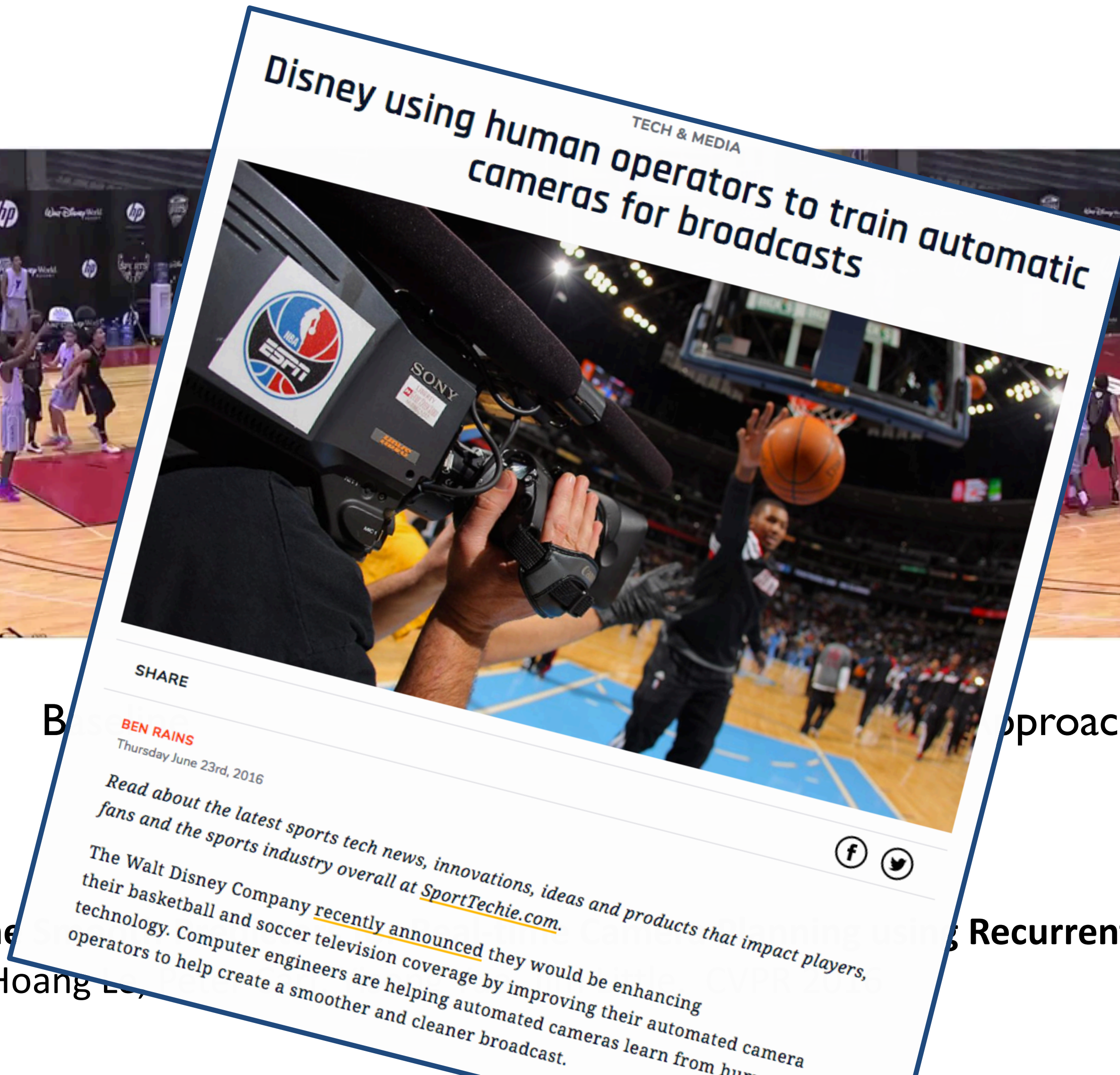
Smooth Imitation Learning for Online Sequence Prediction
Hoang Le, Andrew Kang, Yisong Yue, Peter Carr. ICML 2016

Our Result



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

Qualitative Comparison


TECH & MEDIA

Disney using human operators to train automatic cameras for broadcasts

SHARE

BEN RAINS
Thursday June 23rd, 2016

Read about the latest sports tech news, innovations, ideas and products that impact players, fans and the sports industry overall at SportTechie.com.

The Walt Disney Company recently announced they would be enhancing their basketball and soccer television coverage by improving their automated camera technology. Computer engineers are helping automated cameras learn from hu[...] operators to help create a smoother and cleaner broadcast.

B[...] [...]pproach

Learning Online [...] Recurrent Decision Trees
Jianhui Chen, Hoang [...]

Lessons Learned

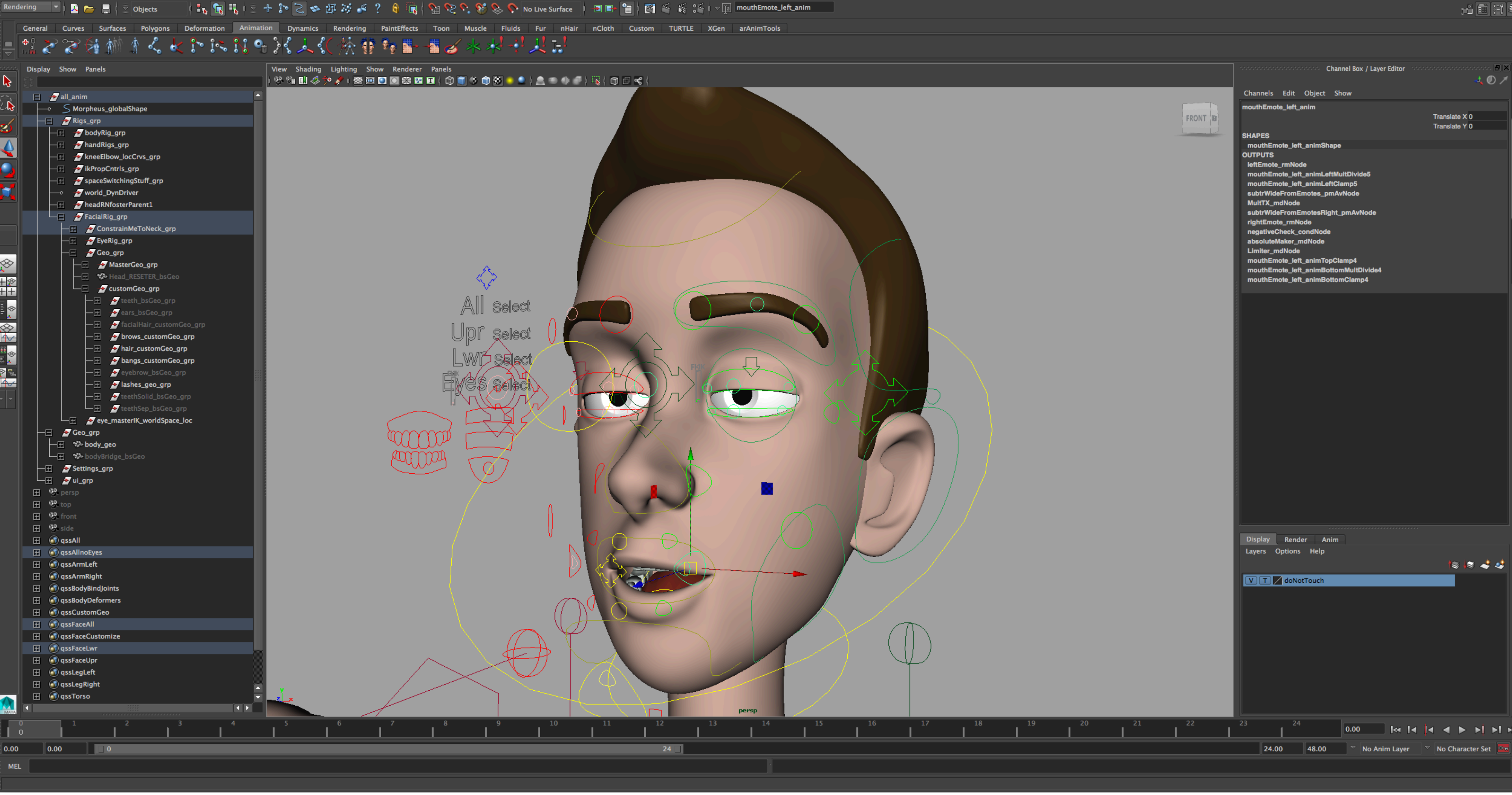
- **Intuition:** Let model do most of work
  – Black box (deep neural net) adds flexibility
  – "Regularization" improves learning
    - Exponentially faster convergence compared to SEARN

Exploit Lipschitz from smooth temporal dynamics

- Applicable to other approaches?
  – Synthesize program + black box
  – Optimal controller + black box

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

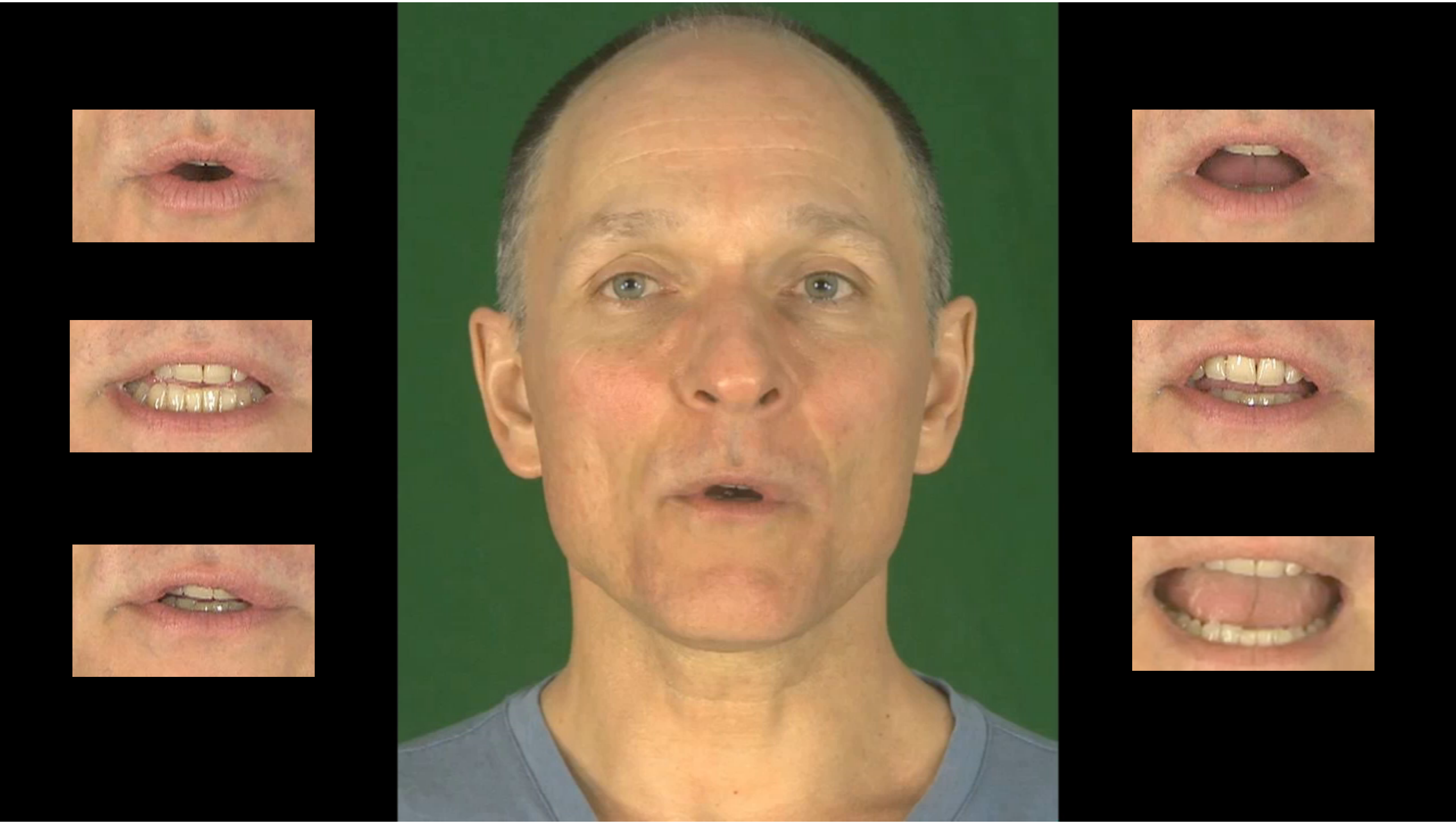
Speech Animation



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

We'll focus on mouth & speech.

Co-Articulation is Hard



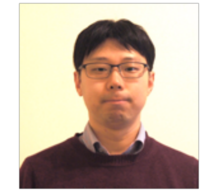
/k/

Automatically Animate to Input Audio?

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

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 $h : X \rightarrow Y$

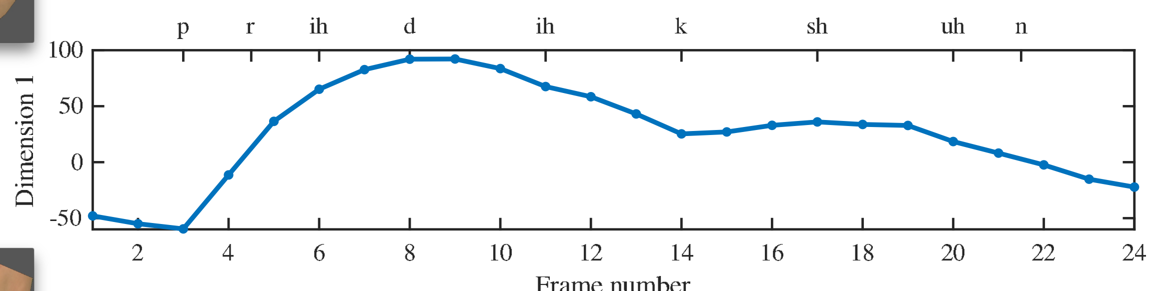
X

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

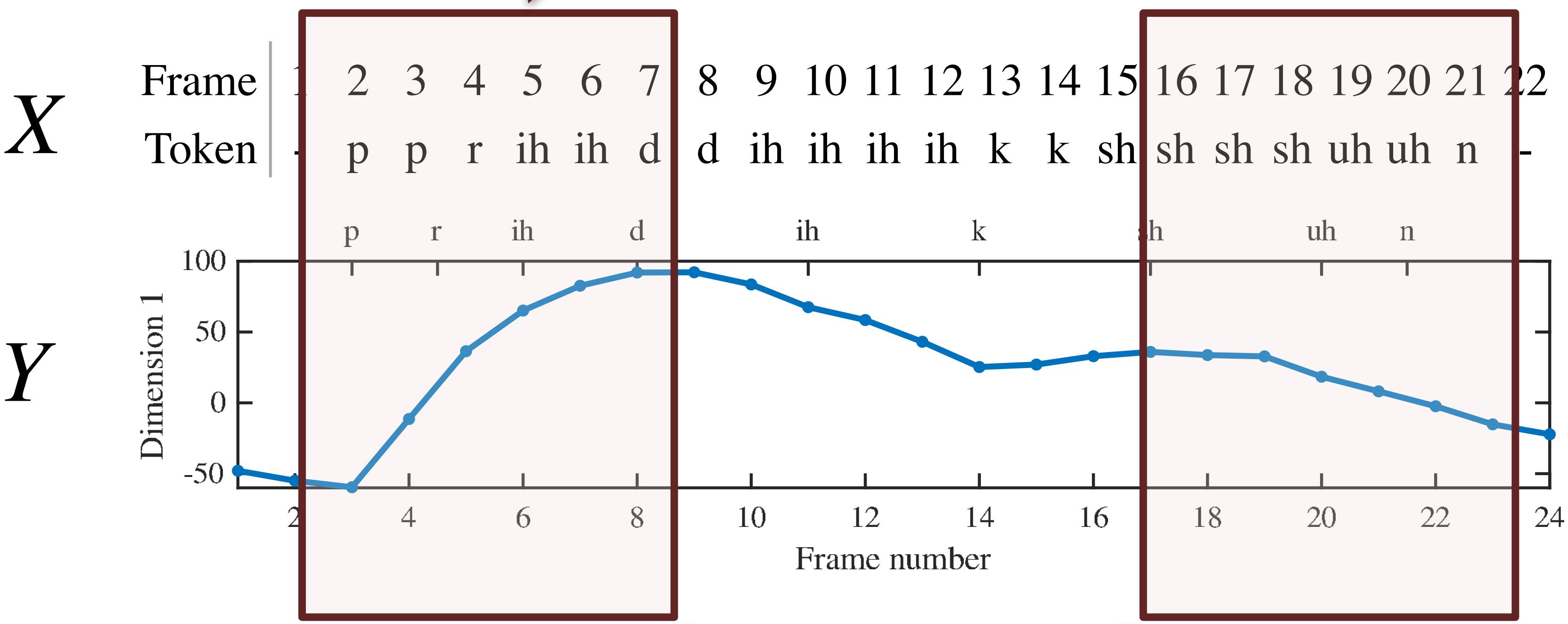
Y



Sequence of face configurations

Temporal curvature can vary smoothly or sharply

(Depends on context – this is the co-articulation problem)



Minimal long-range dependencies

(prediction = construction = election…)

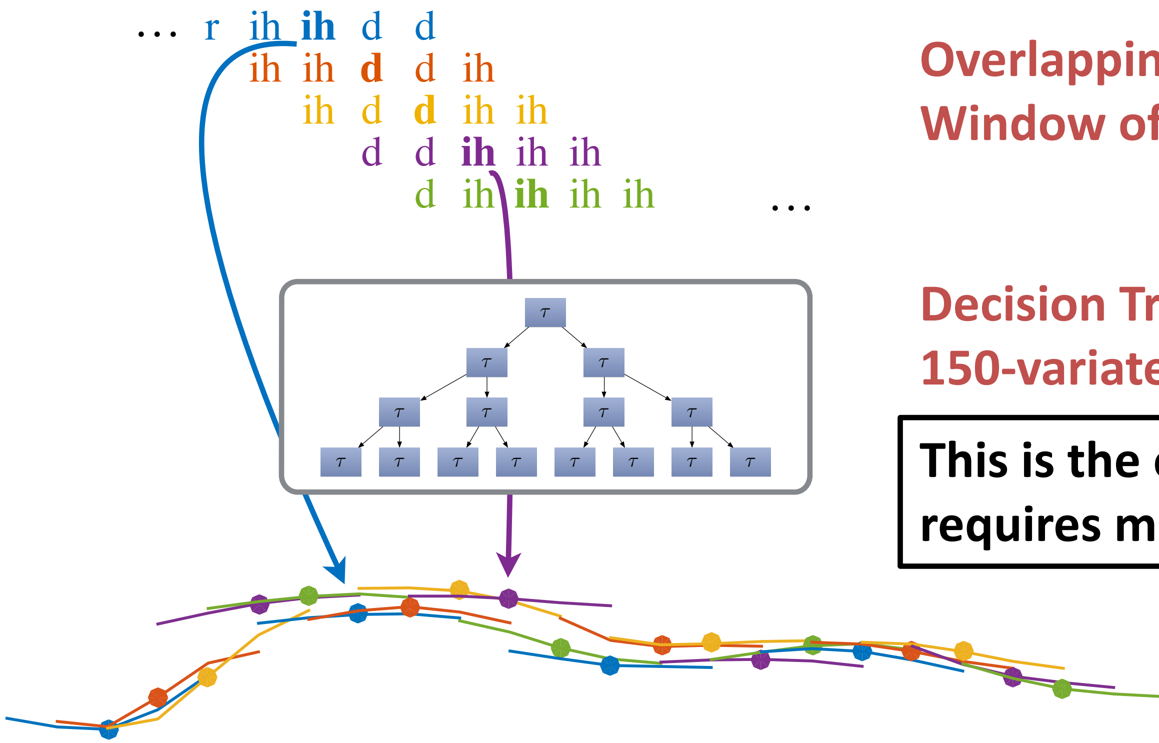
Input speech: “ P R E D I C T I O N ”

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

$\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots$

$h(\hat{\mathbf{x}})$

$\hat{\mathbf{y}}_1, \hat{\mathbf{y}}_2, \ldots$



Overlapping Sliding Window of Inputs

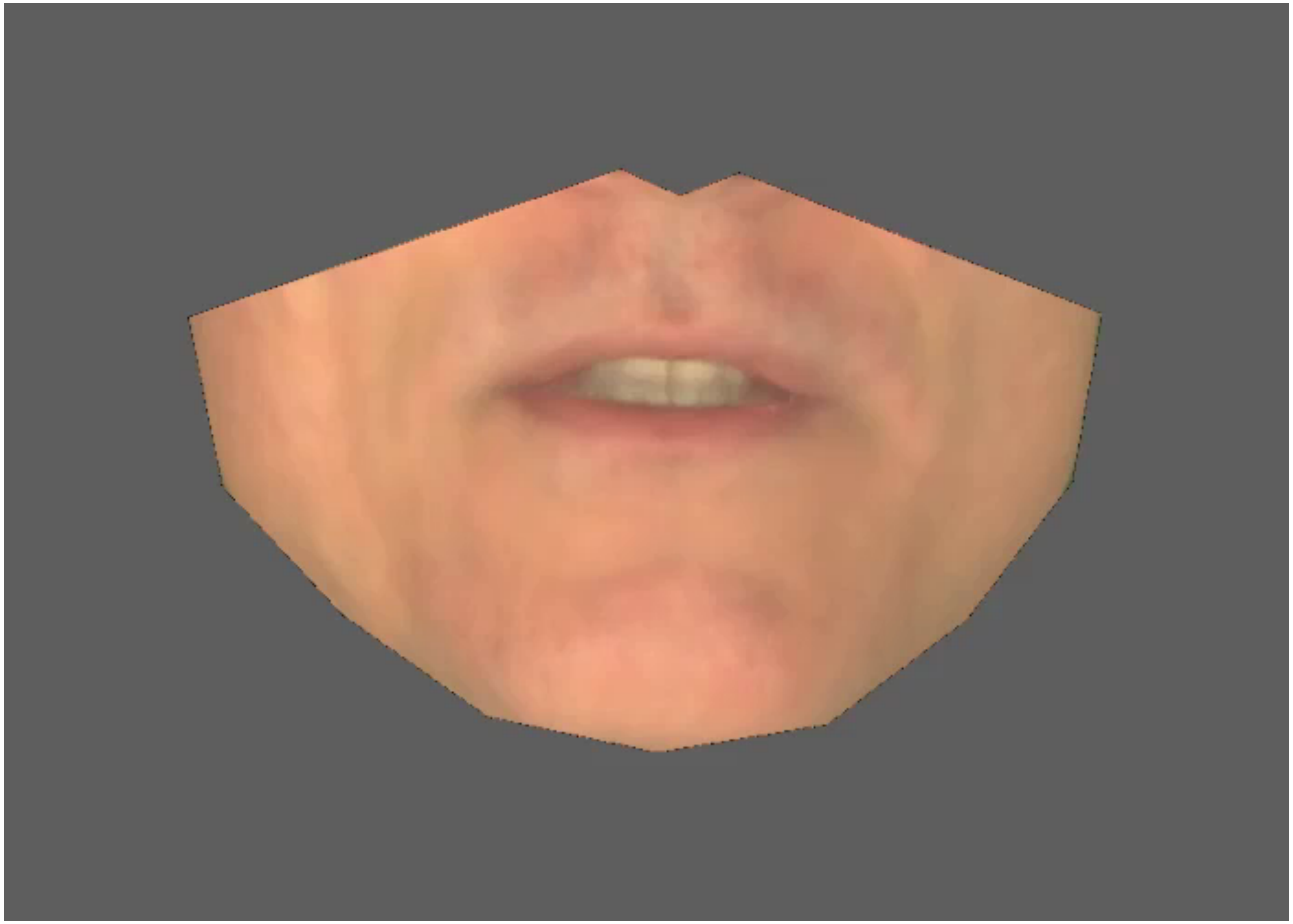
Decision Tree Model
150-variate regression

This is the only thing that requires machine learning!

Aggregate Outputs

Very fast!

Prediction for Very Different Language



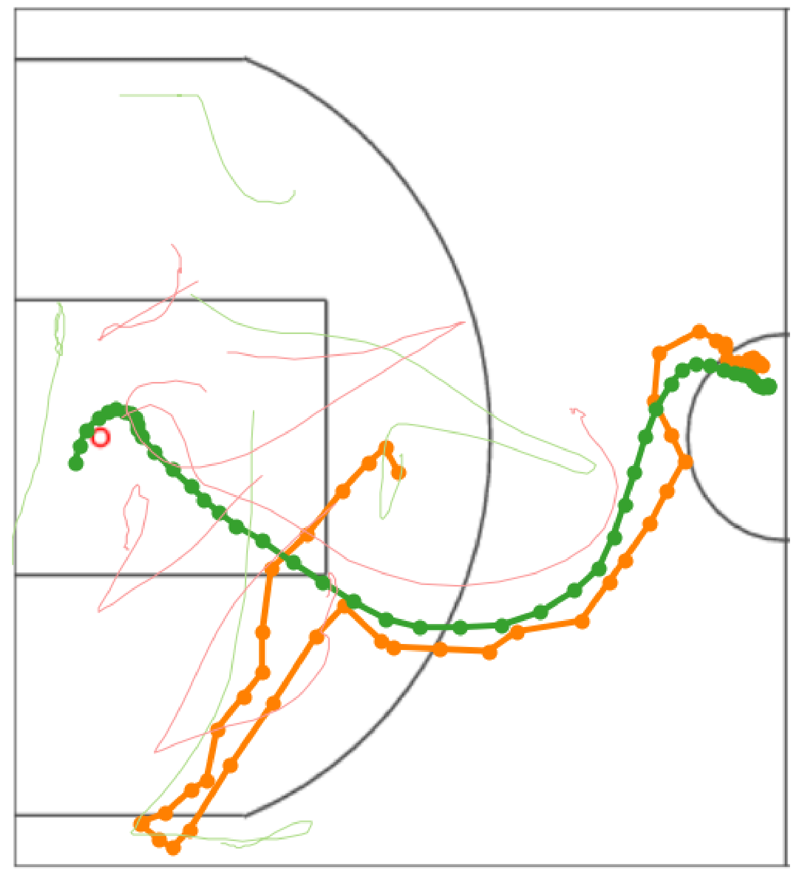
Behind the Scenes of Pandora - The World of Avatar https://youtu.be/URSOqWtLix4

Hierarchical Behaviors

(New Model Classes)

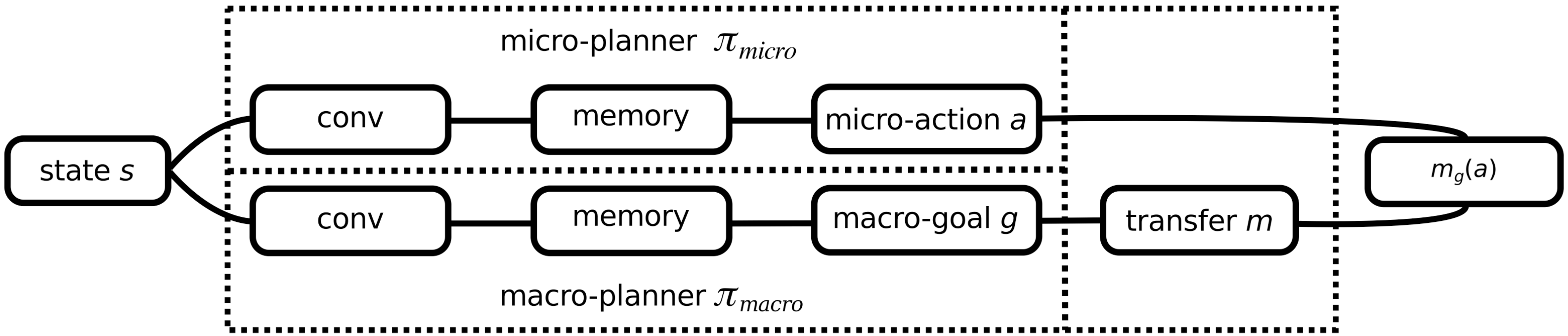
Strategy vs Tactics

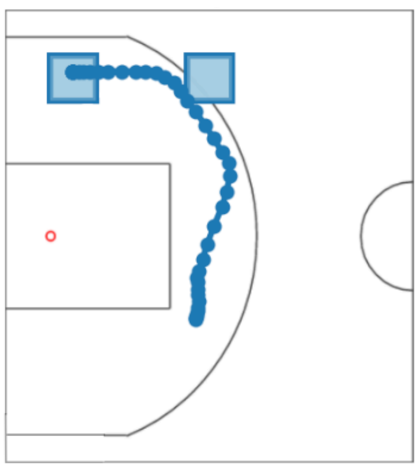
- Long-term Goal:
  - Curl around basket
- Tactics
  - Drive left w/ ball
  - Pass ball
  - Cut towards basket

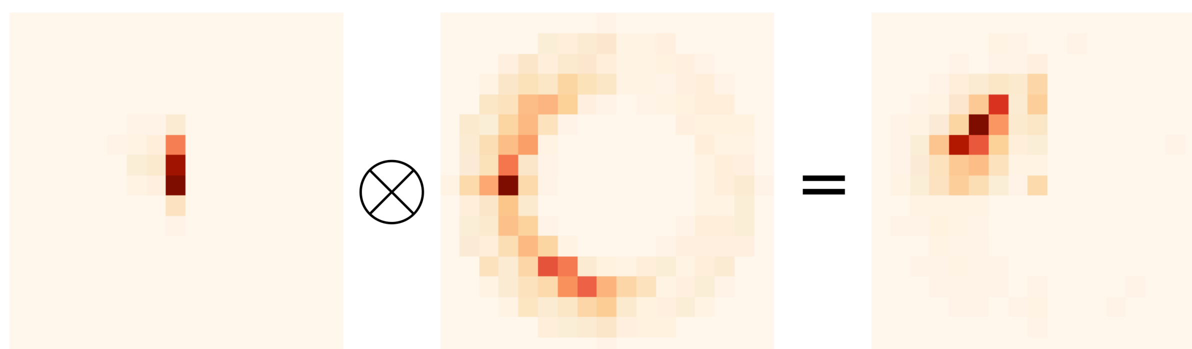


Stephan Zheng

Macro Goals & Micro Actions

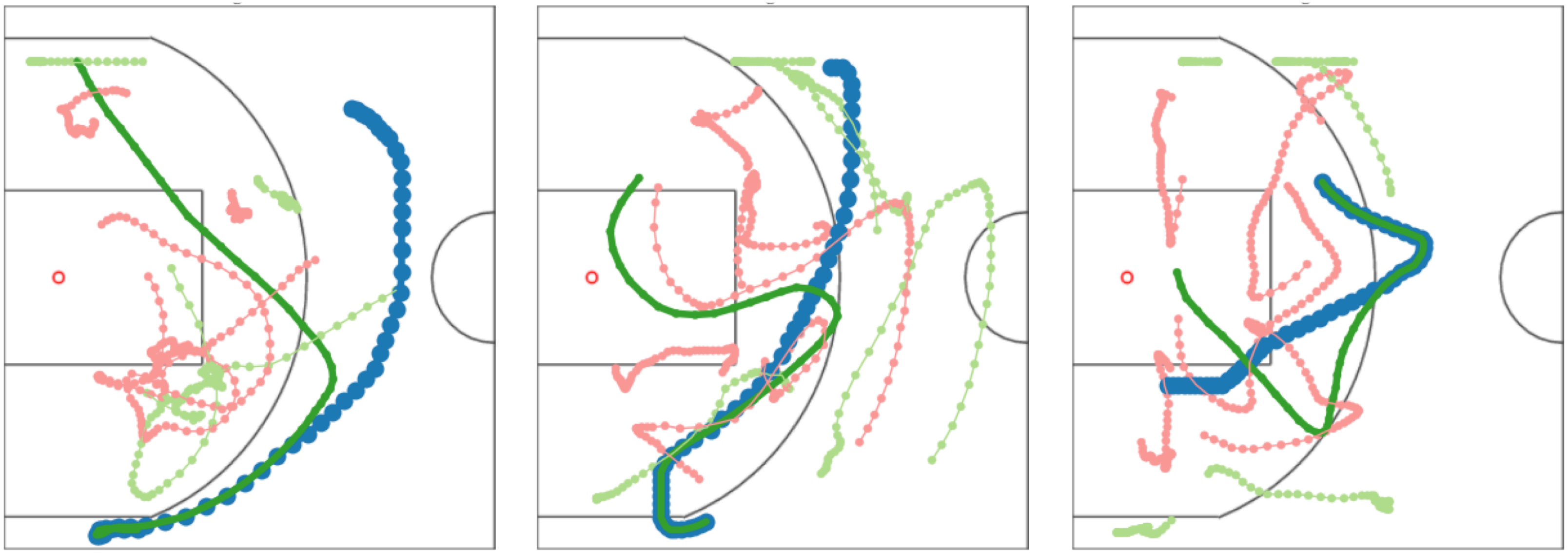


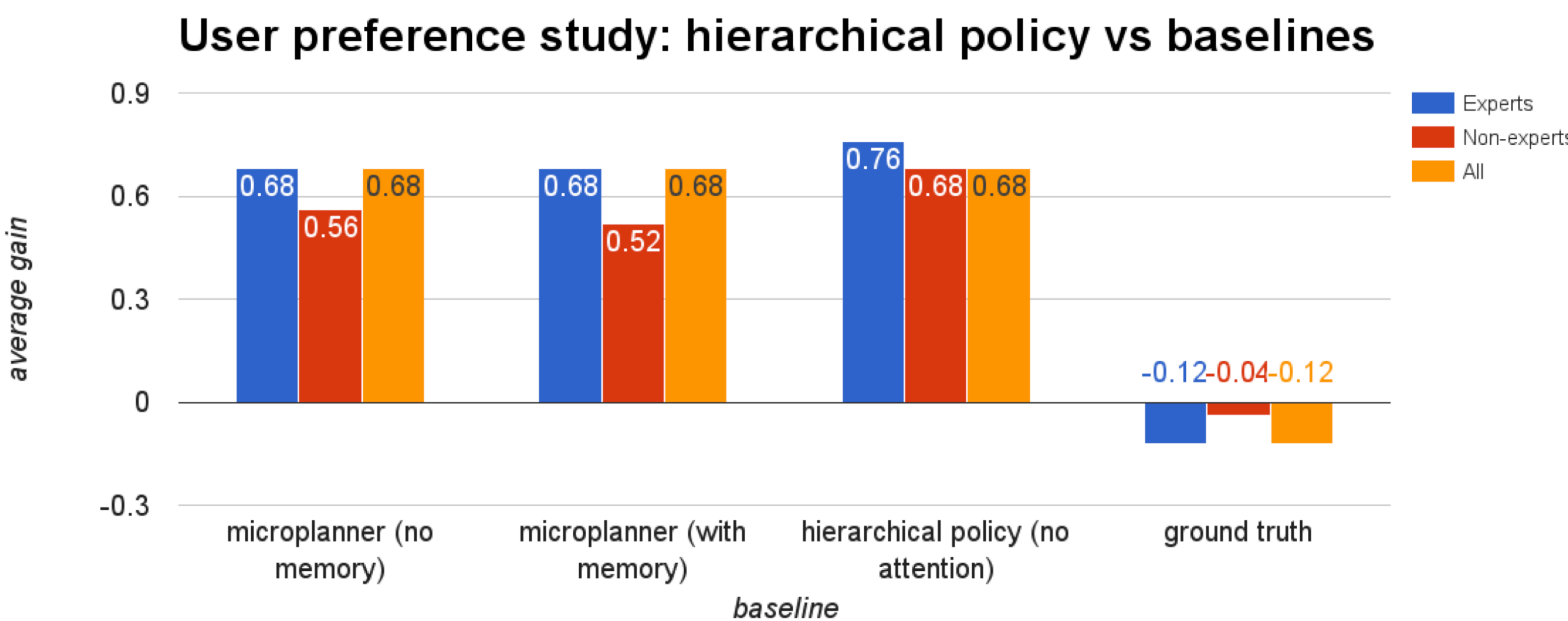




Micro-action Attention Final prediction

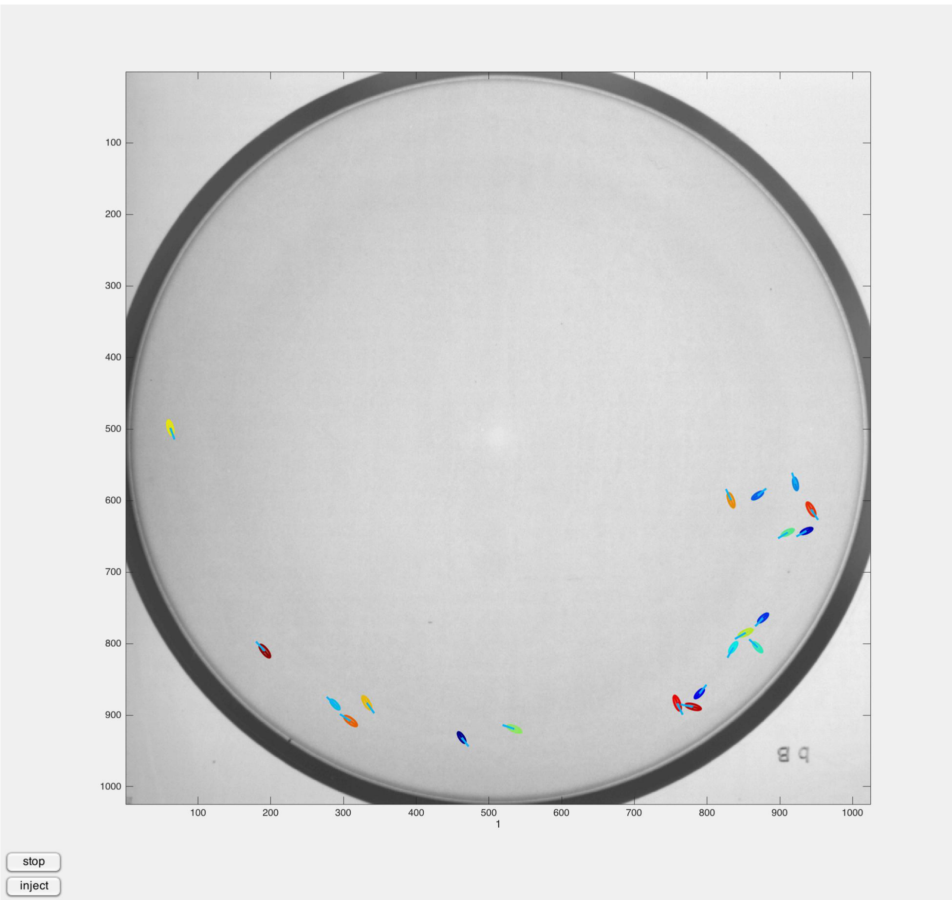
Generating Long-term Trajectories using Deep Hierarchical Networks
Stephan Zheng, Yisong Yue, Patrick Lucey. NIPS 2016

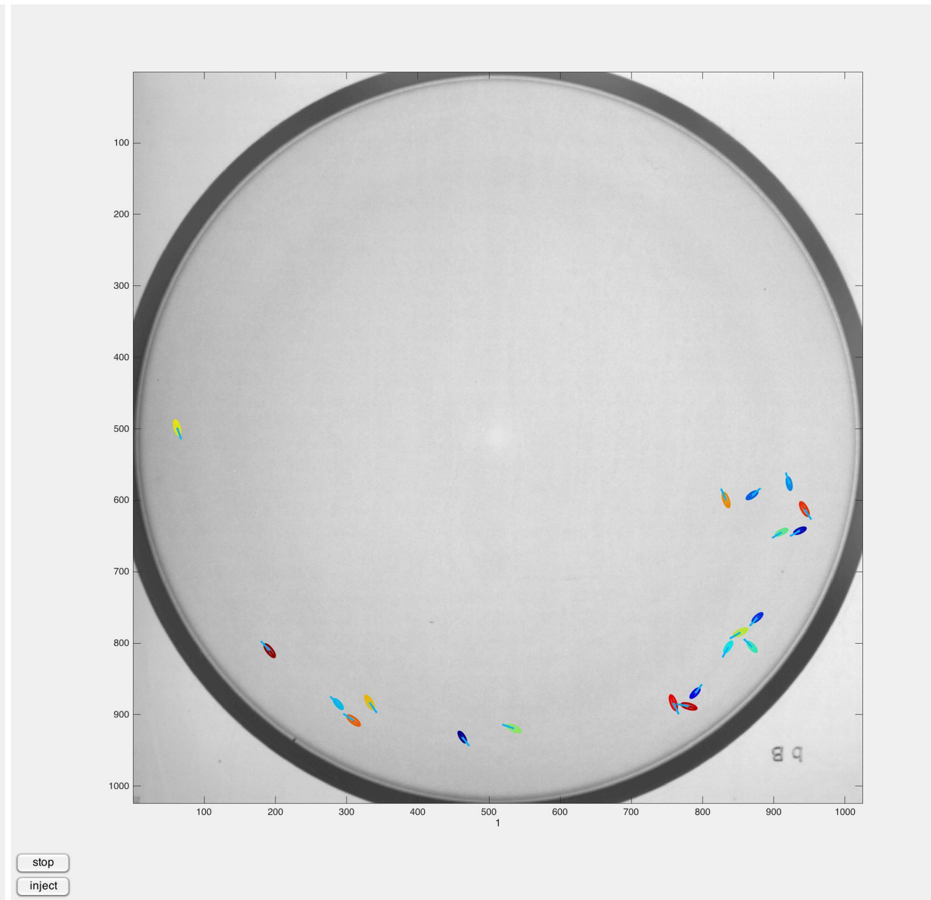



User preference study: hierarchical policy vs baselines
0.9
0.6
0.3
0
-0.3
average gain
0.68
0.56
0.68
0.68
0.52
0.68
0.76
0.68
0.68
-0.12
-0.04
-0.12
microplanner (no memory)
microplanner (with memory)
hierarchical policy (no attention)
ground truth
baseline
Experts
Non-experts
All

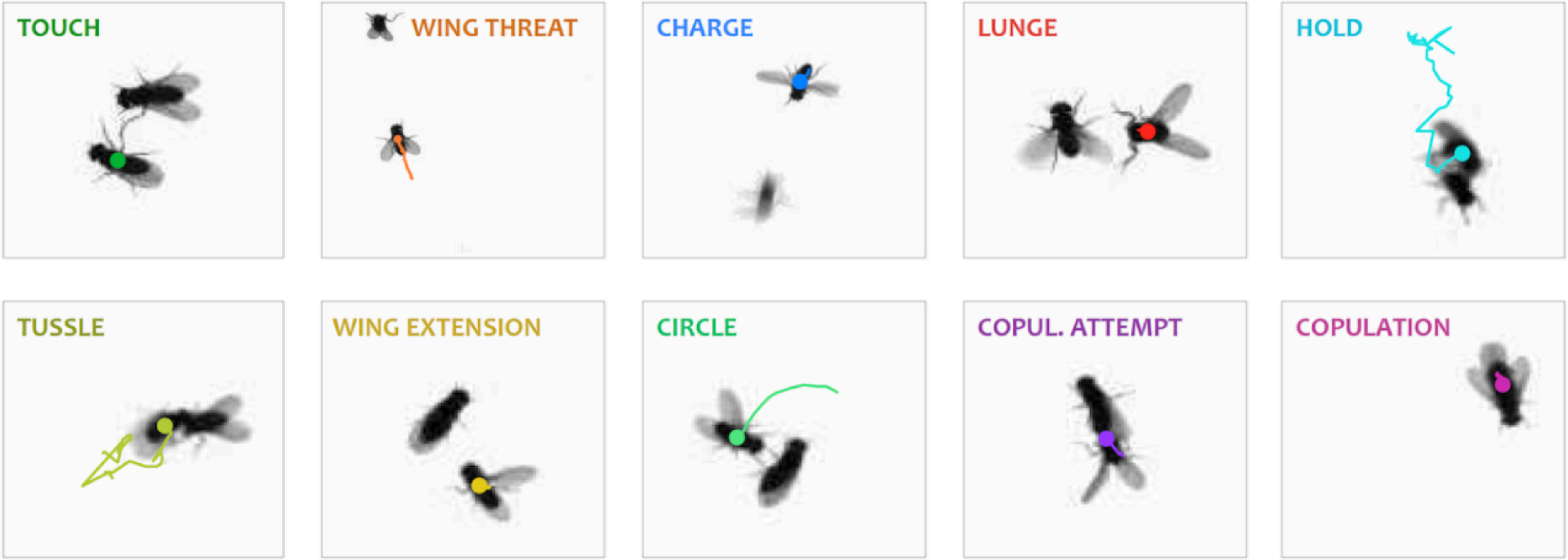
Eyrun
Eyolfsdottir

Drosophila Behavior



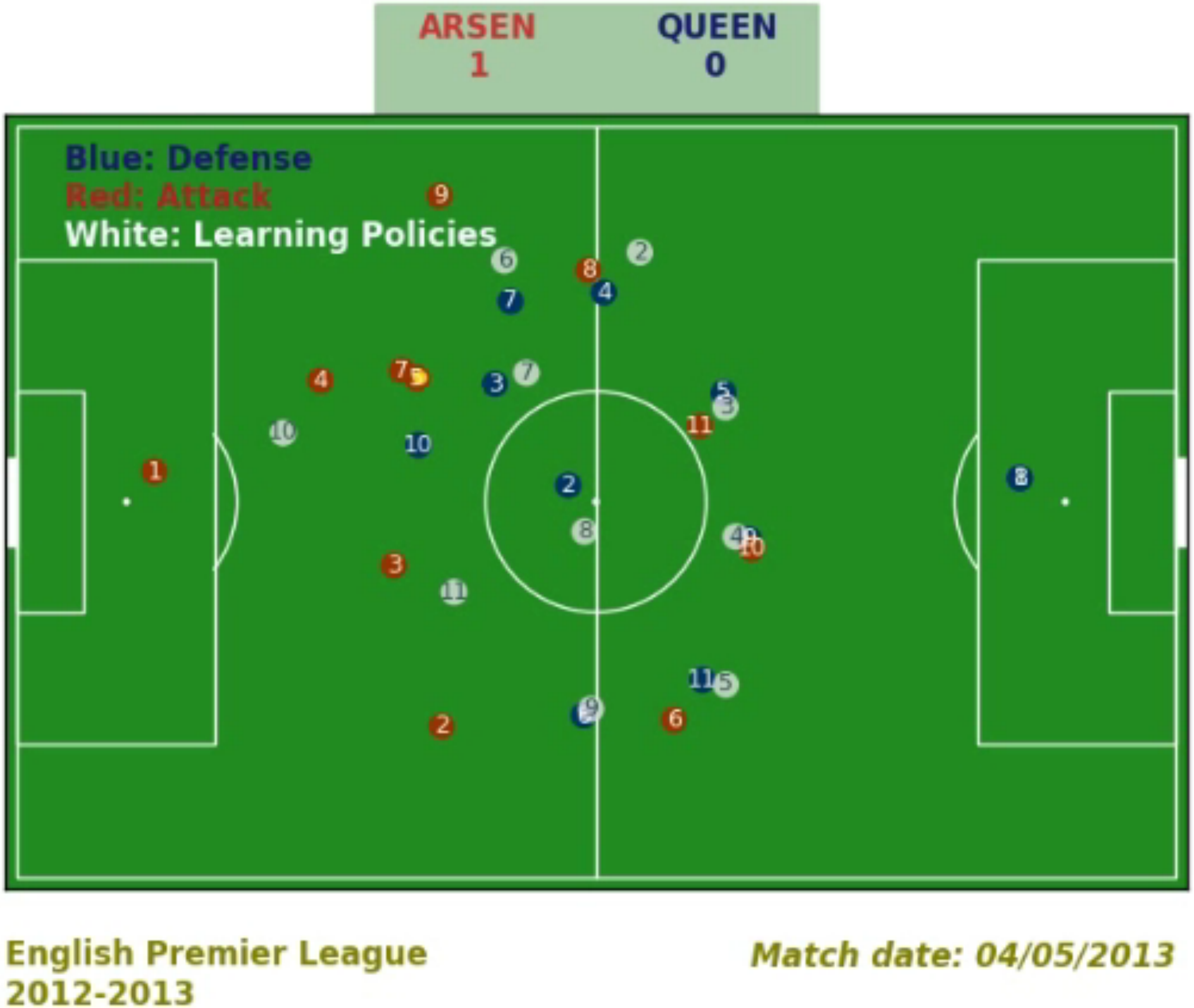


Activity Labels



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

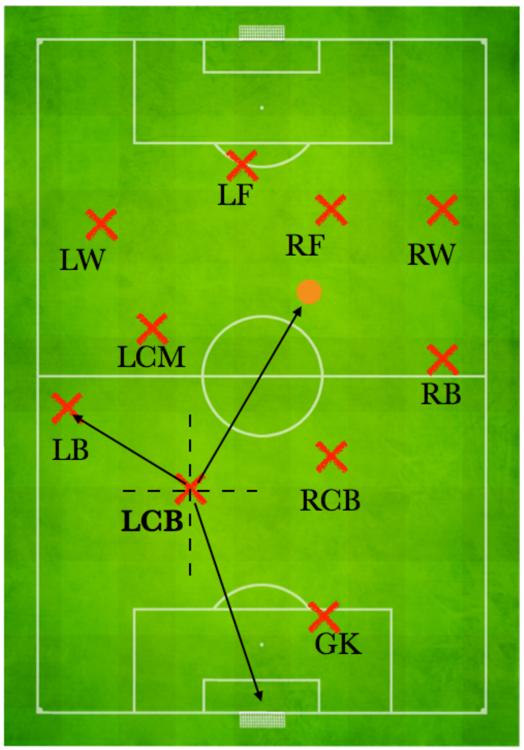
Multi-Agent Systems
(Coordinated Imitation Learning)



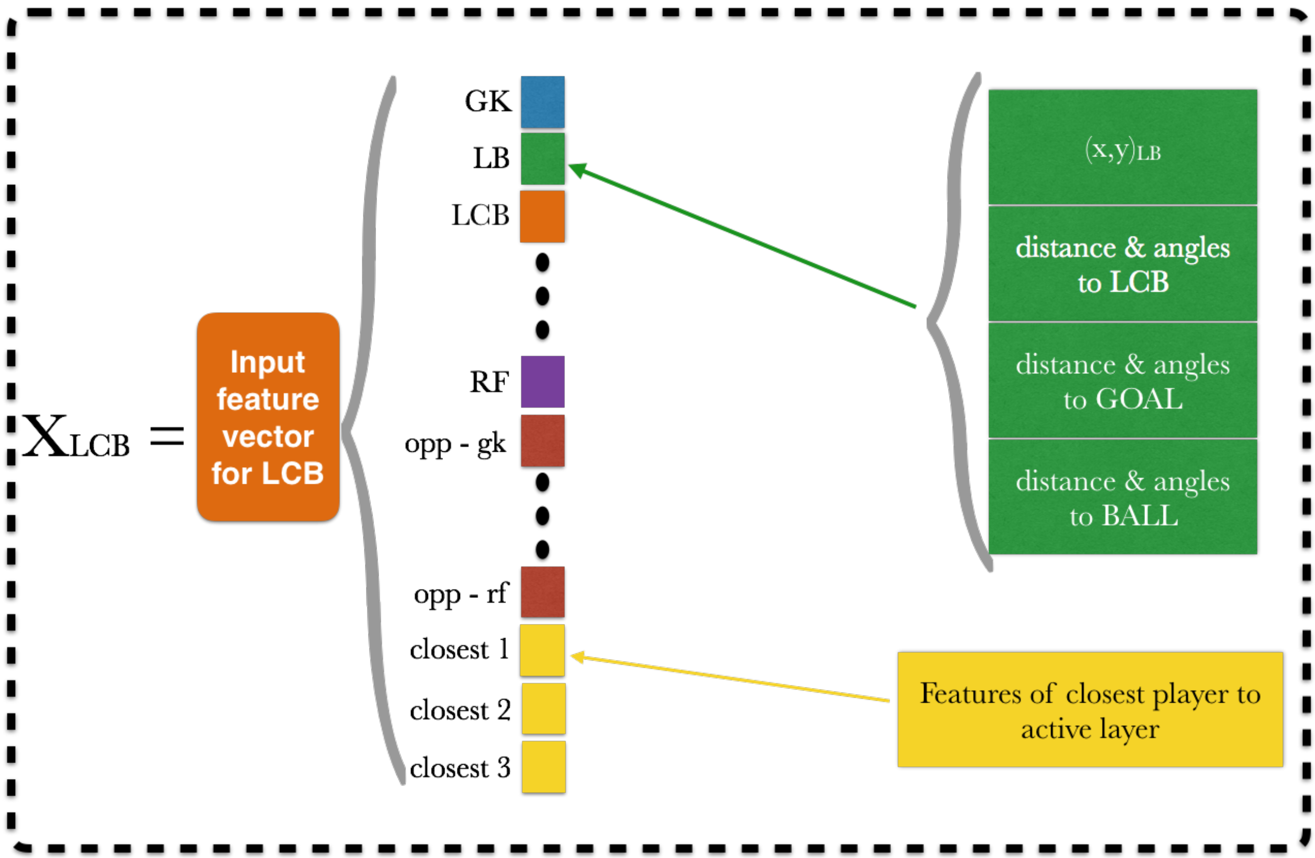
Data-Driven Ghosting using Deep Imitation Learning

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

State Representation



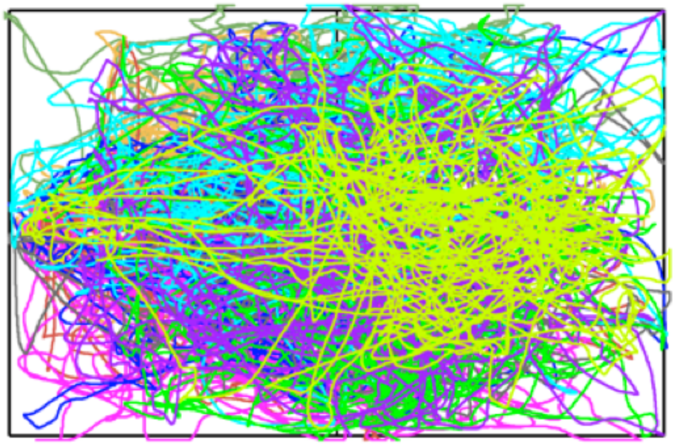
Geometric features computed



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.



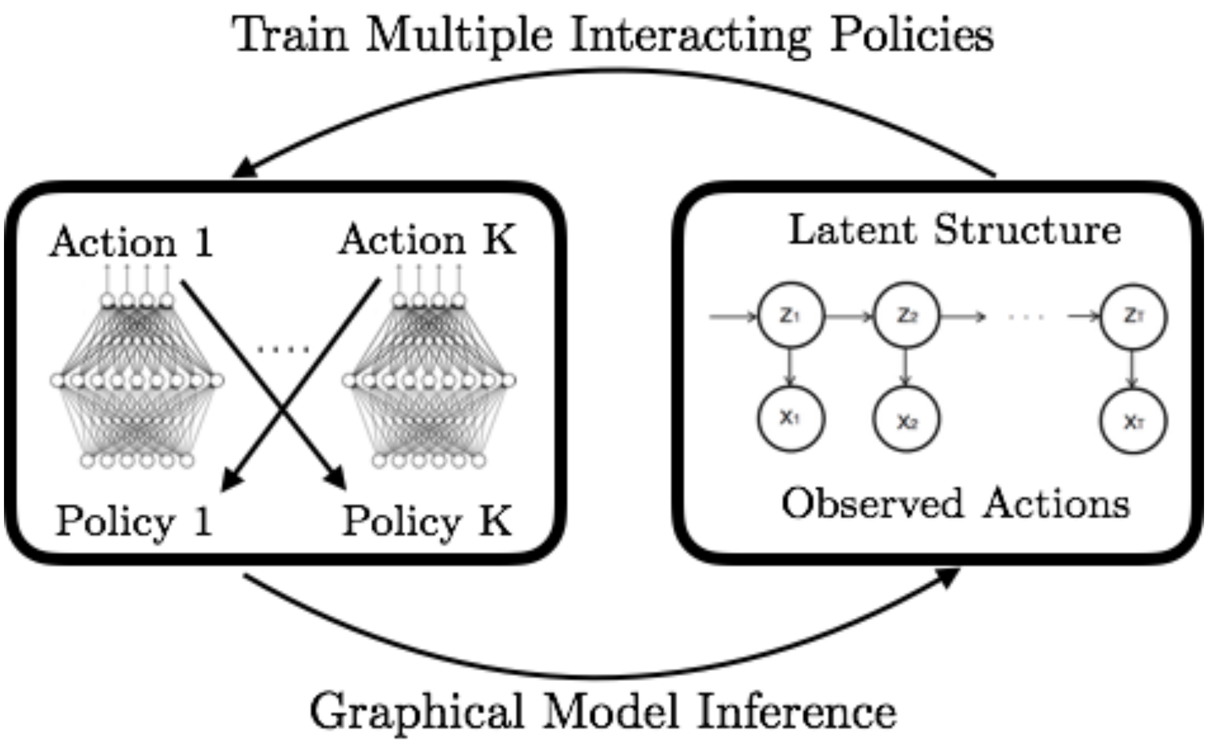
- Need to solve a permutation problem
  - **What happens if we ignore this?**

English Premier League
2012-2013

Match date: 04/05/2013

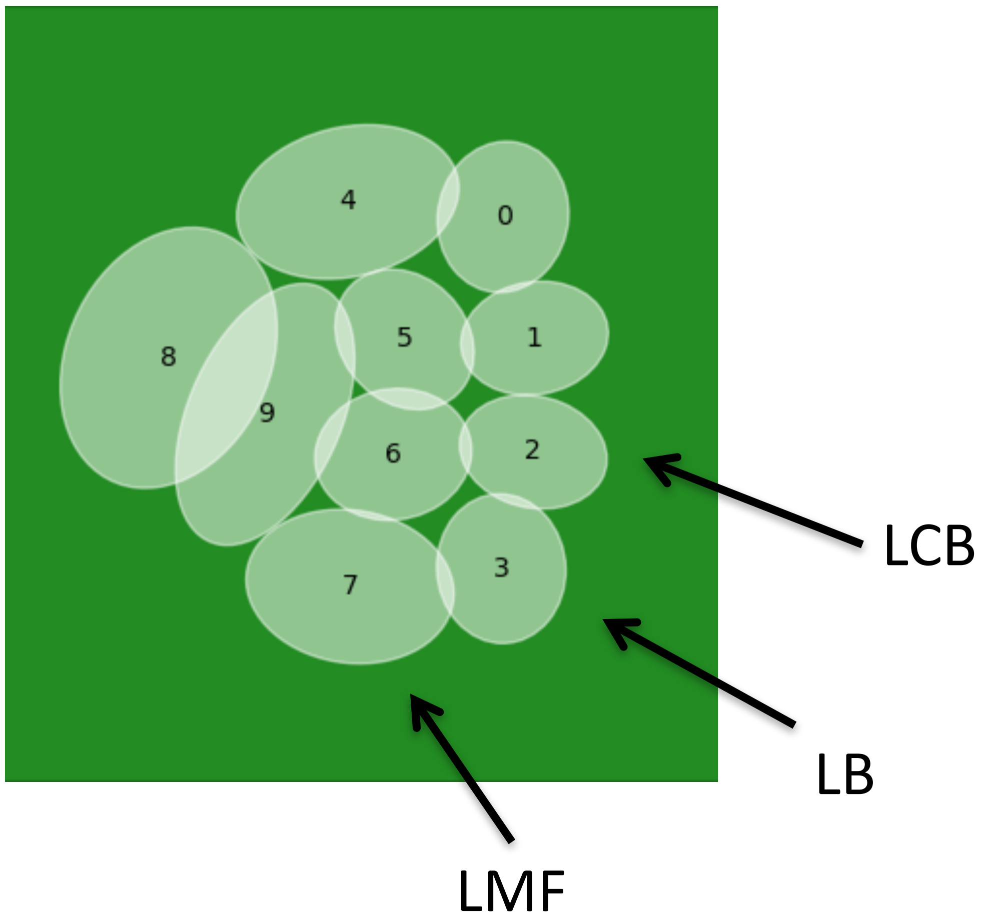
Hoang
Le

Coordination Model



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

Learned Roles



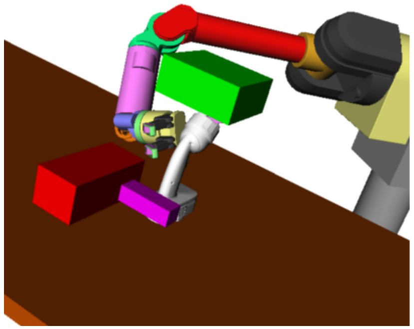
Learning to Optimize

(Combinatorial) Optimization

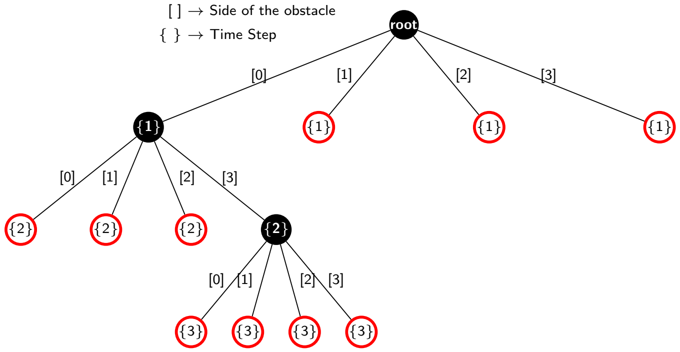
- Find good feasible solutions
  - Within combinatorial search space
- Examples:
  - Mixed Integer Programming
  - Submodular Optimization
  - Boolean Satisfiability
  - Etc…
- Typically solved using local search heuristics

Challenges in Optimization

- Expensive oracles
  - Computational intensive
  - Not always available



- Weak search heuristics
  - Long search time
  - Low solution quality

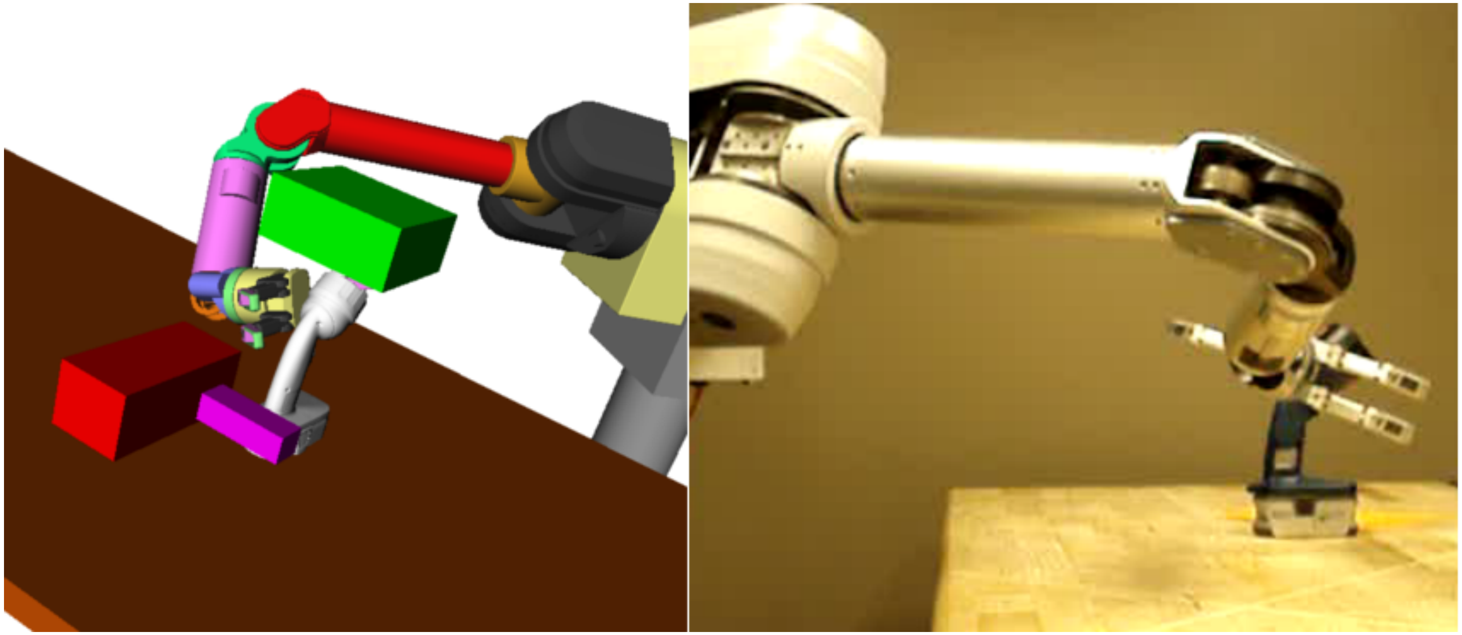


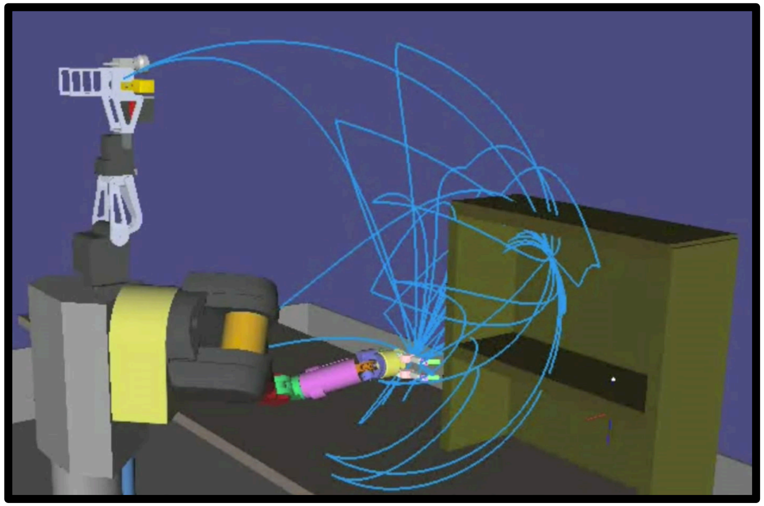
Sequential Decision Making

- Many solvers are sequential:
  - Greedy
  - Search heuristics
- Can view as solver as “agent”
  - State = intermediate solution
  - Find a state with high reward (solution)

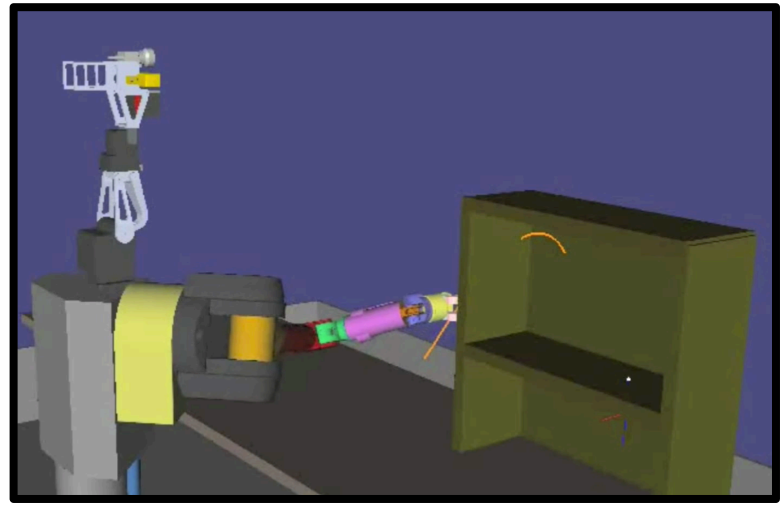
Grasp Trajectory Prediction

- Quickly identify successful trajectory
- Requires high-fidelity simulator (slow)

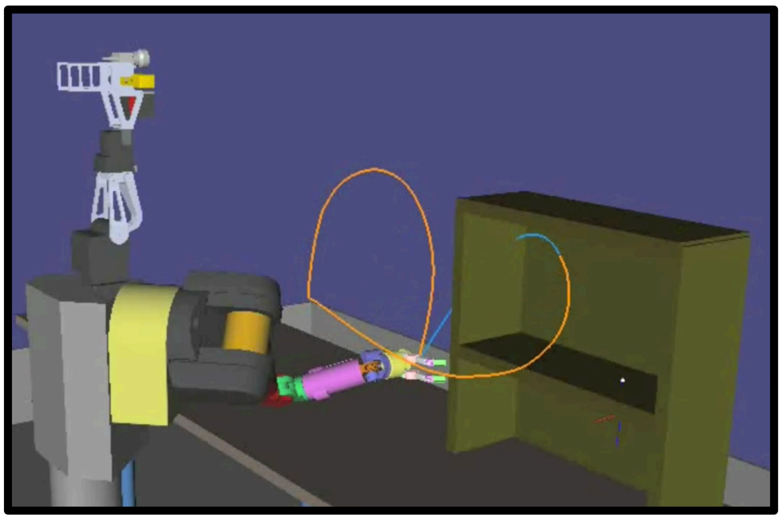




High-fidelity Simulator



First trajectory fails



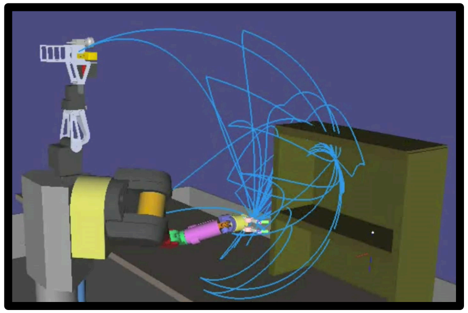
Second trajectory diverse

Requires Statistical Model of Diversity!

Learning Policies for Contextual Submodular Prediction
S. Ross, R. Zhou, Y. Yue, D. Dey, J.A. Bagnell. ICML 2013

Contextual Submodular Optimization

- Contextual submodular optimization
  - Tests are redundant
  - Depends on context

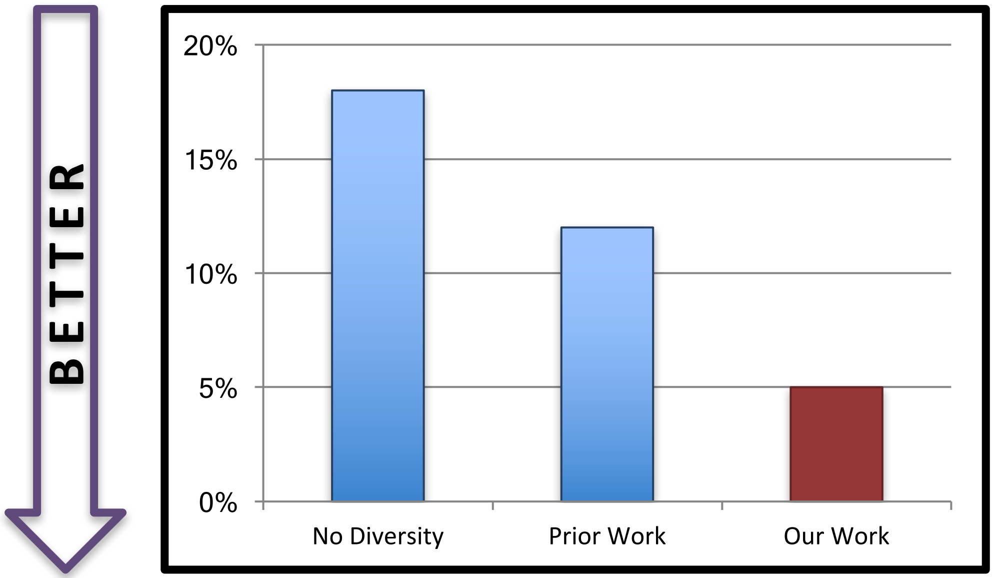


- Typically solved using greedy
  - If you know the submodular function

- **Goal:** imitate greedy algorithm
  - Decisions based on features
  - Minimize dependency on oracle (at test time)

Stephane
Ross

Robotic Trajectory Prediction



Failure Rate (4 evaluations)

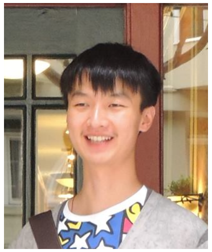
Learning Policies for Contextual Submodular Prediction
S. Ross, R. Zhou, Y. Yue, D. Dey, J.A. Bagnell. ICML 2013

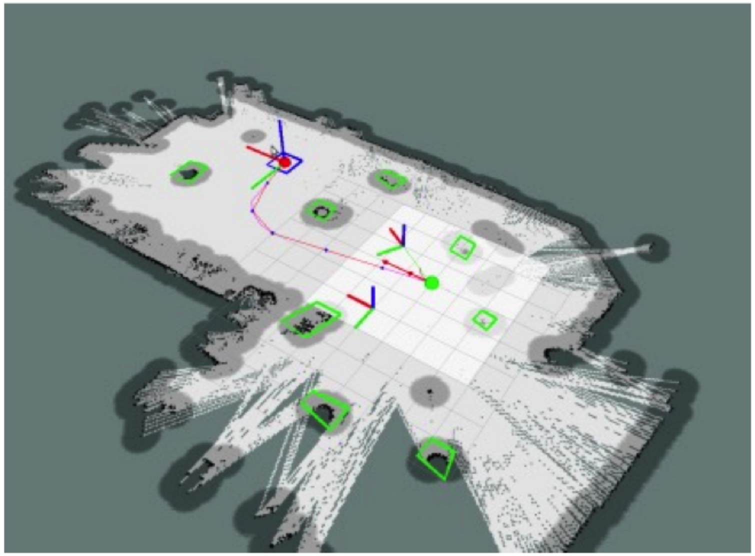
Ravi
Lanka

Ongoing Research

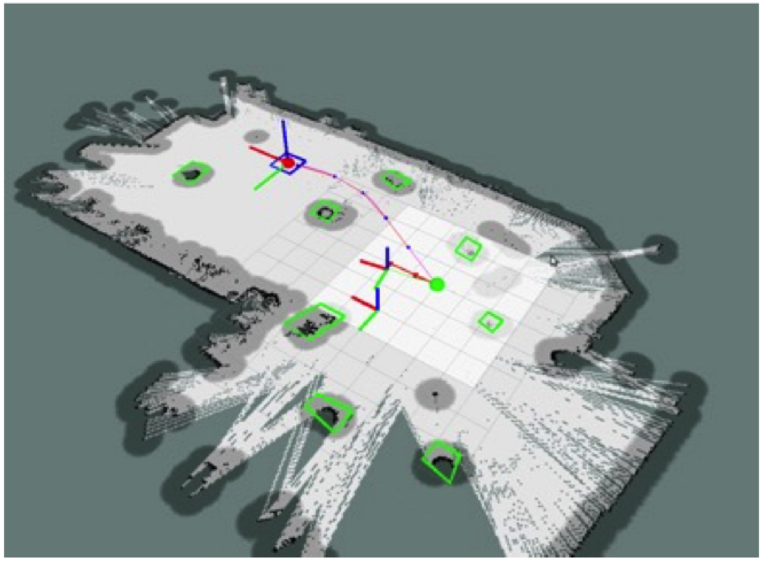
Risk-Aware Planning

Jialin
Song



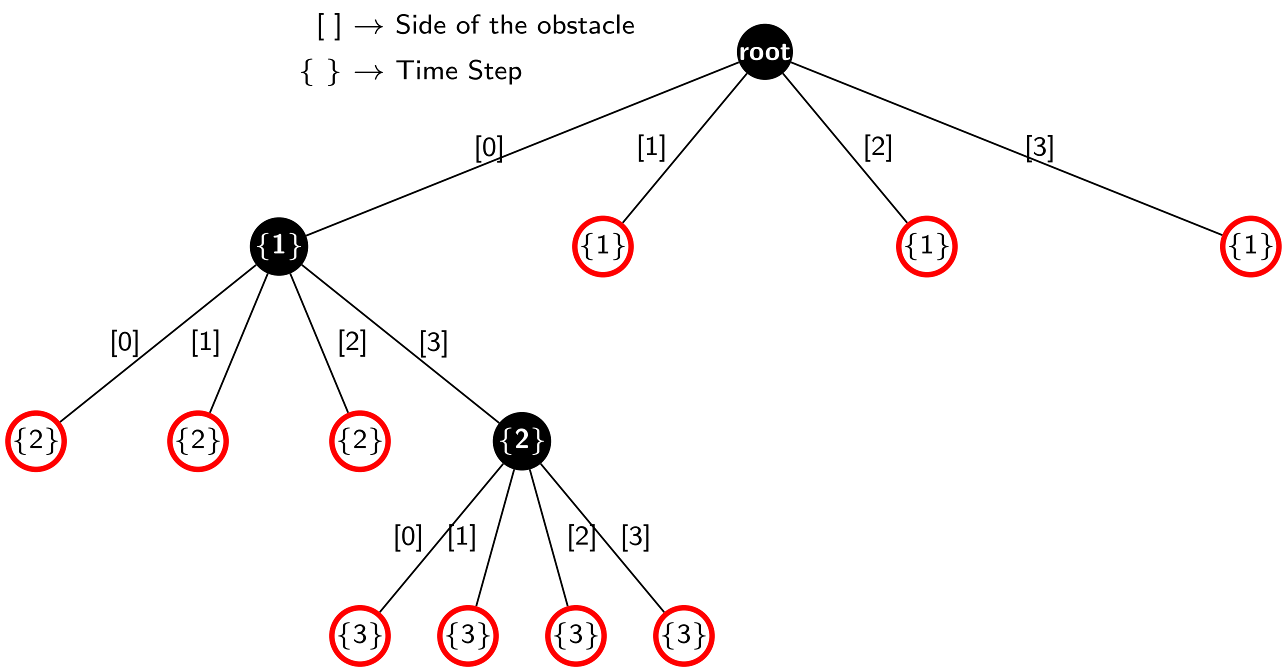


Low Risk



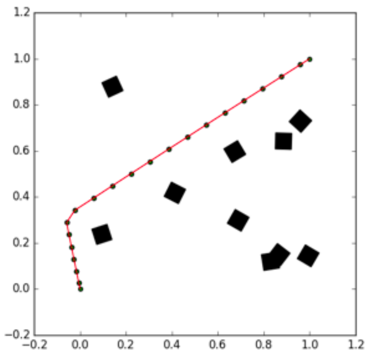
High Risk

- Compiled as mixed integer program
- Challenging optimization problem

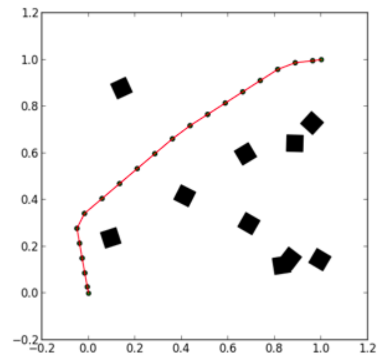


- Exponential search space
  - Local search heuristics (e.g., branch and bound)
- **Goal:** Learn statistical model of search space
  - Find feasible solutions much faster

Preliminary Results



Optimal Solution (Gurobi solver)



Our Approach

	Ours	Gurobi Solver
Train	1049	15241
Test	1127	25249

Avg Nodes Explored

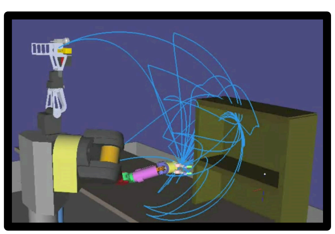
	Ours	Gurobi Solver
Train	0.732	0.305
Test	0.577	0.309

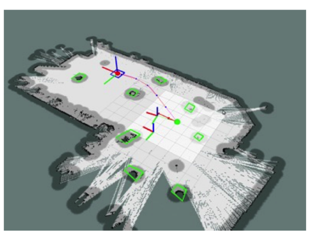
Avg Objective value

Learning to Search via Self-Imitation with Application to Risk-Aware Planning
R. Lanka, J. Song, A. Zhao, Y. Yue, M. Ono (under review)

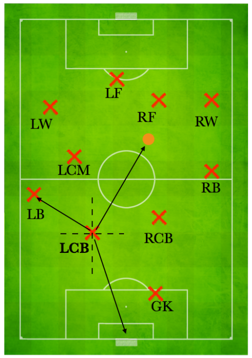
New Frontiers in Imitation learning

- **Incorporating Structure**
  - Smoothness of output space
  - Latent structure of input space
- **New Algorithmic Frameworks**
  - Black Box + Model-Based Planning
  - Black Box + Latent Graphical Models
- **Cool Applications!**







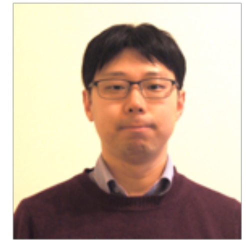


Eyrun
Eyolfsdottir

Jimmy
Chen

Stephan
Zheng

Hoang
Le



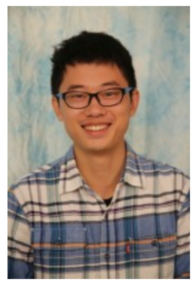
Taehwan
Kim

Stephane
Ross

Jialin
Song

Andrew
Kang

Debadeepta
Dey



Robin
Zhou

Albert
Zhao

Sarah
Taylor

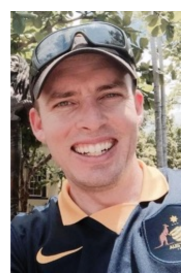
Ravi
Lanka

Kristin
Branson

Iain
Matthews

Jim
Little

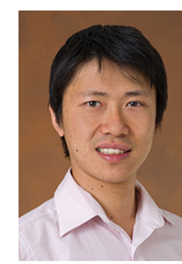
Pietro
Perona



Patrick
Lucey

Drew
Bagnell

Peter
Carr



Masahiro
Ono

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

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