LCCC - Learning and Adaptation for Sensorimotor Control,
October 24-26, Lund University, Sweden

Measuring Motion Complexity and Its Applications to Learning of Motion Skills

Hanyang University, Seoul, Korea

October 24, 2018

Il Hong Suh

한양대학교
HANYANG UNIVERSITY

LUNDS UNIVERSITET

Contents

1. What motion will be more complex?

2. What motion skill will be better learned first?

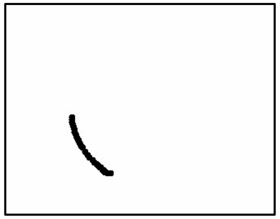
3. What and where to attend to learn from demonstrations?

1. What motion will be more complex?

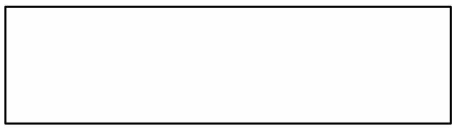
2. What motion skill will be better learned first?

3. What and where to attend to learn from demonstrations?

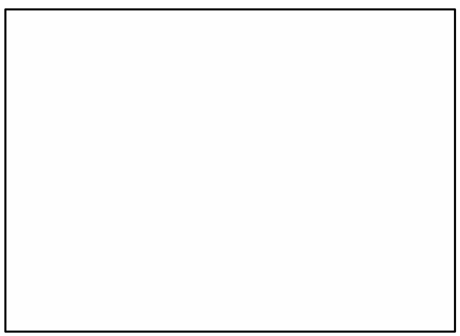
Do you think what motion is complex?



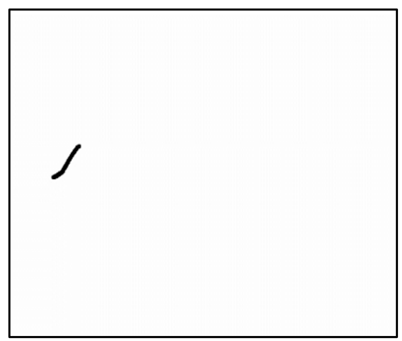
circle



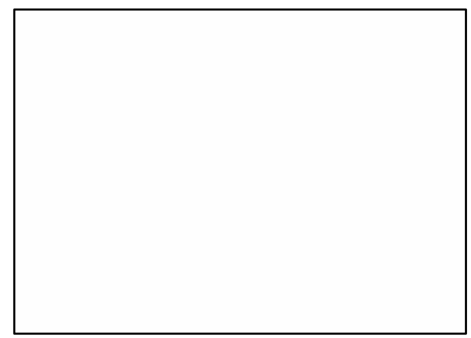
line



rectangle



alphabets



Random stroke

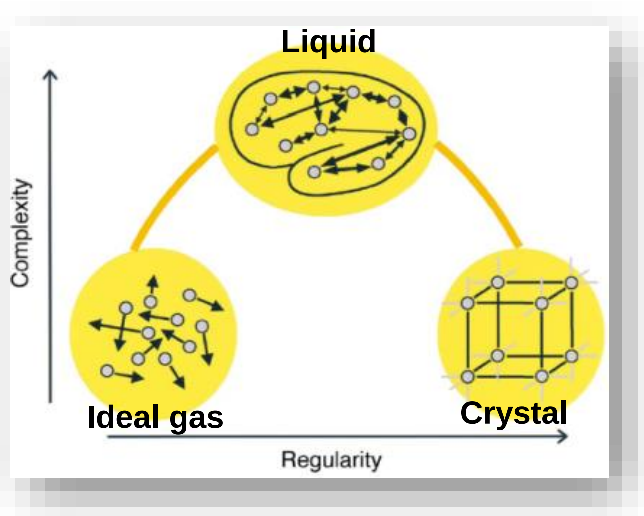
Low complexity

High complexity

Low complexity

order → **disorder**

Neural Complexity Measure (1/2)



Neural Complexity (G. Tononi, Science 1998)

$$C_N(\mathbf{X}) = \sum_{k=1}^{n}\left[\frac{k}{n} I(\mathbf{X}) - \left\langle I\left(X_j^k\right)\right\rangle\right]$$

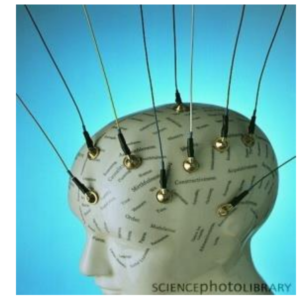
$I(X)$: integration (a measure of dependency among elements in X)

$\boldsymbol{X}$: a whole system

X_j^k : a subsystem including k elements

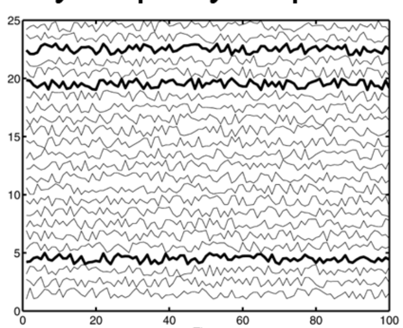
$<X_j>$: ensemble average over j

$$I(\mathbf{X}) = \sum_{i=1}^{n} H(x_i) - H(\mathbf{X})$$



Random

25-Dimension Brain signals: Nearly completely independent

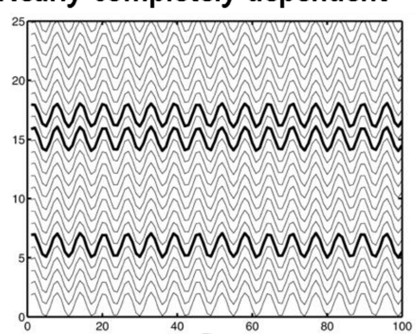


No correlation among 25 signals

$$\underset{\mathbf{L}}{\frac{k}{25} I(\mathbf{X})} - \underset{\mathbf{L}}{\left\langle I\left(X_j^k\right)\right\rangle} \approx 0$$

Regular

25-Dimension Brain signals: Nearly completely dependent

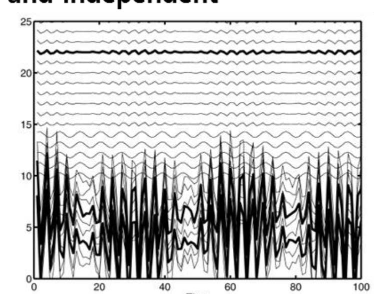


High correlation among 25 signals

$$\underset{\mathbf{H}}{\frac{k}{25} I(\mathbf{X})} - \underset{\mathbf{H}}{\left\langle I\left(X_j^k\right)\right\rangle} \approx 0$$

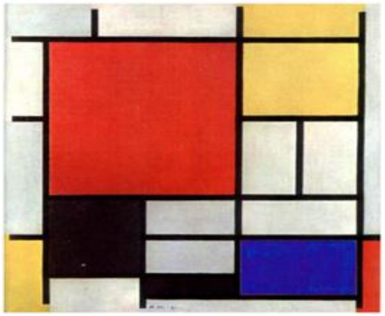
Random + Regular

25-Dimension Brain signals: A balance between dependent and independent



The whole system $\boldsymbol{X}$ is moderately correlated,
Some subsystems are **less** correlated
(i.e. a *subsystem with signal 5,14,23)*
Some subsystems are **highly** correlated,
(i.e. a *subsystem with signal 20,21,22)*

Neural Complexity Measure (2/2)



Mondrian

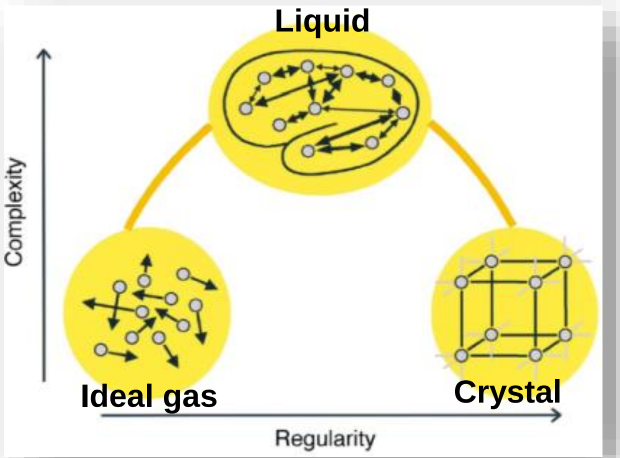
Low Randomness
Simple !

Pollock

High Randomness
Simple !

Bosch

High Randomness
Complex !



[Objective]

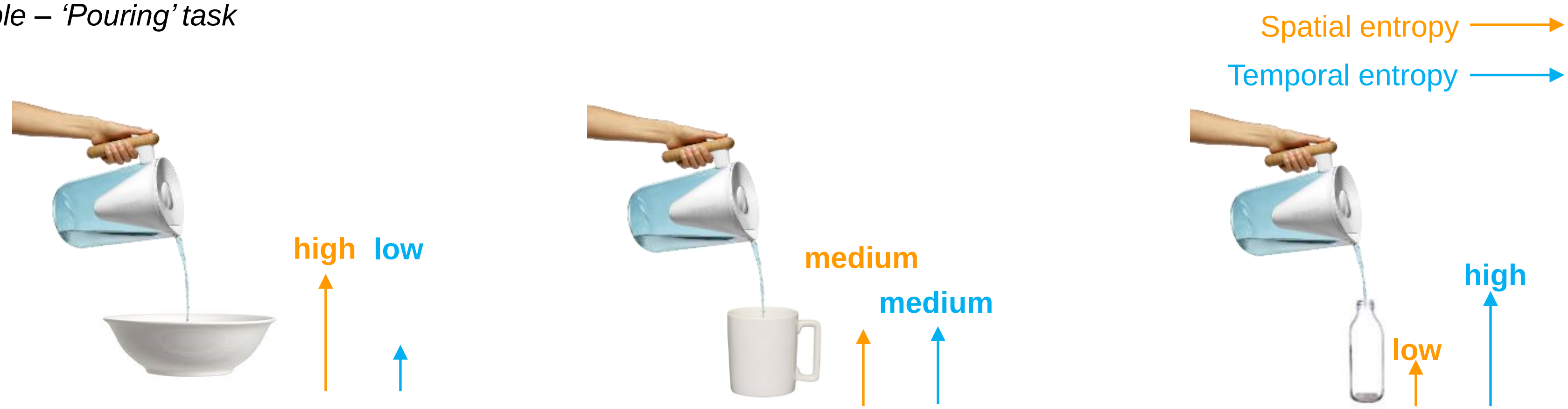
Calculating Motion Complexity

[Problem]

[Neural Complexity] →Intractable computation complexity
(ensemble average of all possible subsystems)
* in time-varying motion trajectories

Motion Complexity and Motion Significance

Example – 'Pouring' task



Quick pouring water into a bowl, which has **a large-size mouth**

Normally pouring water into a cup, which has **a medium-size mouth**

Slow pouring water into a bottle, which has **a small-size mouth**

- Definitions -

Motion significance indicates the relative significance of each motion frame to accomplish the goal of a task at every time index of human demonstrations. ***Motion complexity*** indicates how complex a whole set of human demonstrations is to learn.

- How to measure -

Motion significance is measured by considering both spatial entropy and temporal entropy of a motion frame, based on the analysis of Gaussian mixtures. ***Motion complexity*** is defined by measuring the averaged amount of motion significance involved in an entire set of human demonstrations.

ST-GMM based Motion Complexity/Motion Significance

Motion Significance

Temporal Entropy

$$H_i^{\tau} = \sum_{i=1}^{K} \omega_i \cdot \left(-\log(\omega_i) + \tfrac{1}{2}\log(2\pi e)|\Sigma_i^{\tau}|\right)$$

Spatial Entropy

$$H_i^{X} = \sum_{i=1}^{K} \omega_i \cdot \left(-\log(\omega_i) + \tfrac{1}{2}\log(2\pi e)^{D}|\Sigma_i^{X}|\right)$$

Motion Significance

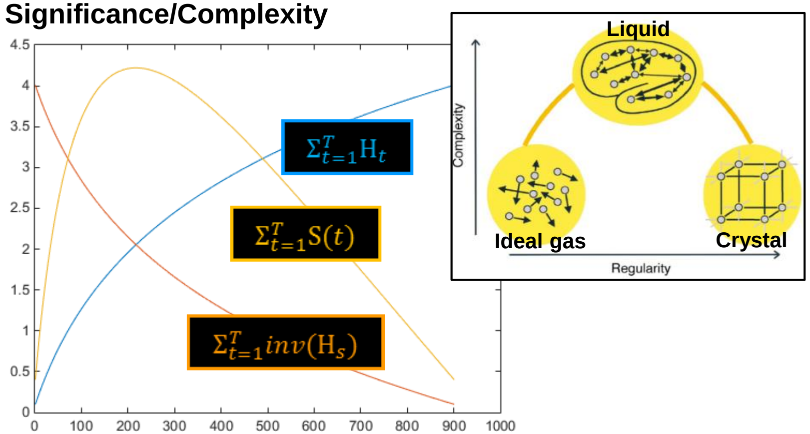
$$S(t) = \frac{zscore(H^{\tau}(t))}{zscore(H^{X}(t))}$$

$*\ H^{\tau}(t),\ H^{X}(t)$: *Interpolated temporal and spatial entropies of all GMMs*

Motion Complexity

$$\mathrm{C} = \frac{1}{T}\sum_{t=1}^{T} S(t)$$

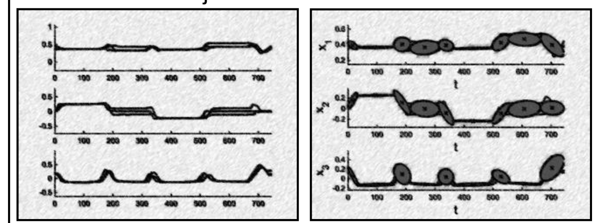
Significance/Complexity



Regularity

Three Motion Trajectories

Gaussian Mixture Model

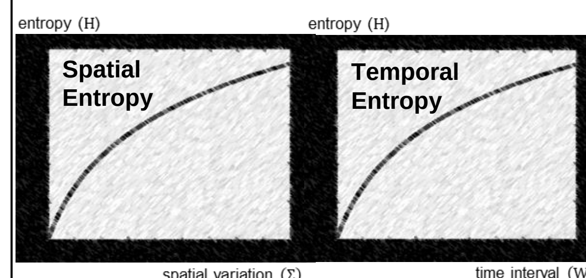


for temporal entropy

$$P(\Psi) = \sum_{i=1}^{K} \omega_i \cdot N(\Psi|\mu_i, \Sigma_i)$$

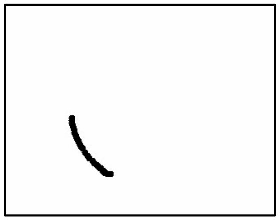
where $\mu_i = (\mu_i^{\tau} \quad \mu_i^{X}),\ \Sigma_i = \begin{pmatrix} \Sigma_i^{\tau} & \Sigma_i^{\tau X} \\ \Sigma_i^{X\tau} & \Sigma_i^{X} \end{pmatrix}$

for spatial entropy



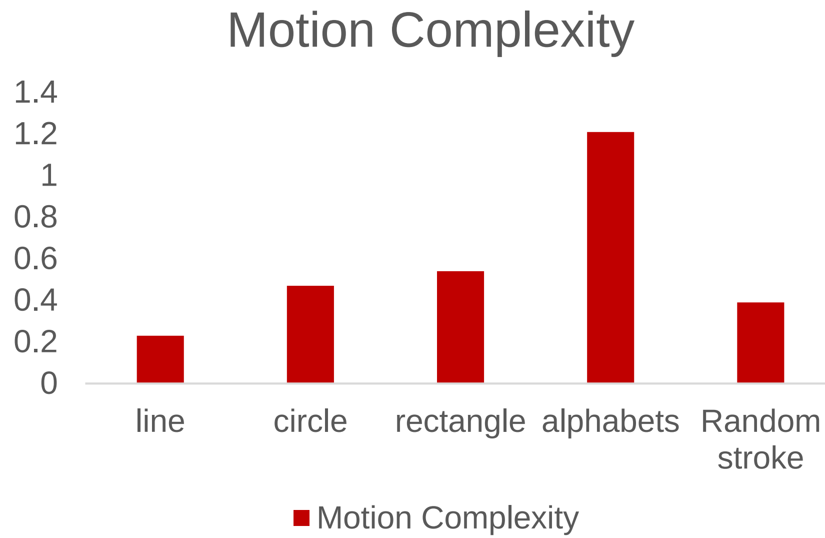
Reference Paper: Il Hong Suh, Sang Hyong Lee, Nam Jun Cho, Woo Young Kwon,
Measuring Motion Significance and Motion Complexity, Journal of Information Science ,Vol388-389, May 2017

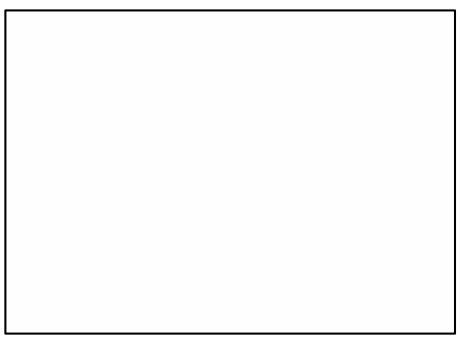
Do you think what motion is complex?



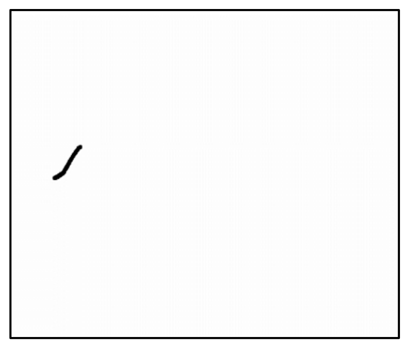
circle

line

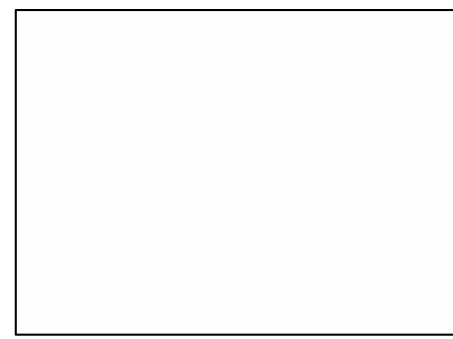




rectangle



alphabets



Random stroke

Low complexity

High complexity

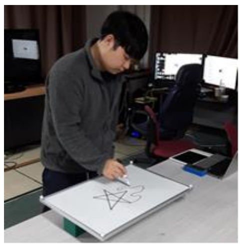
Low complexity

order

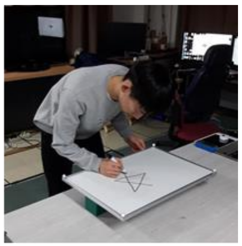
disorder

What motion will be more complex and significant?

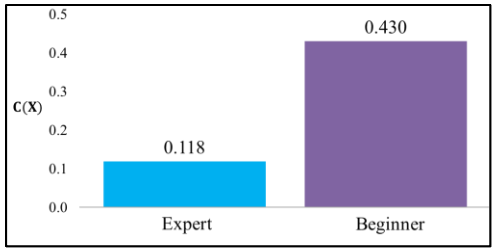
Motion Complexity

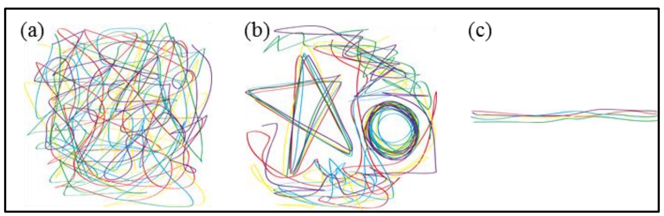


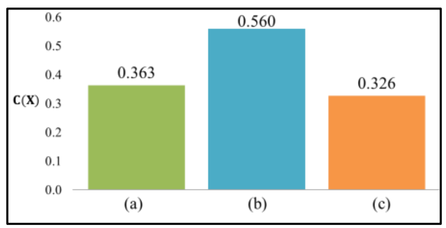
expert

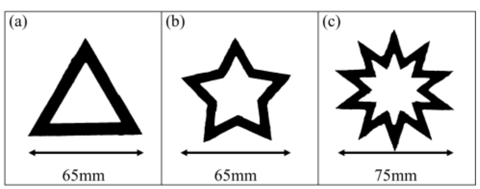


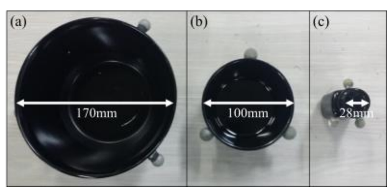
beginner

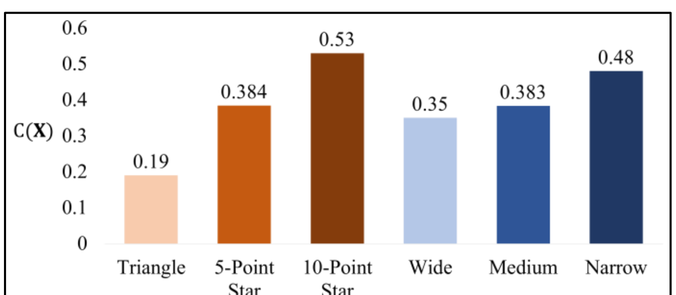




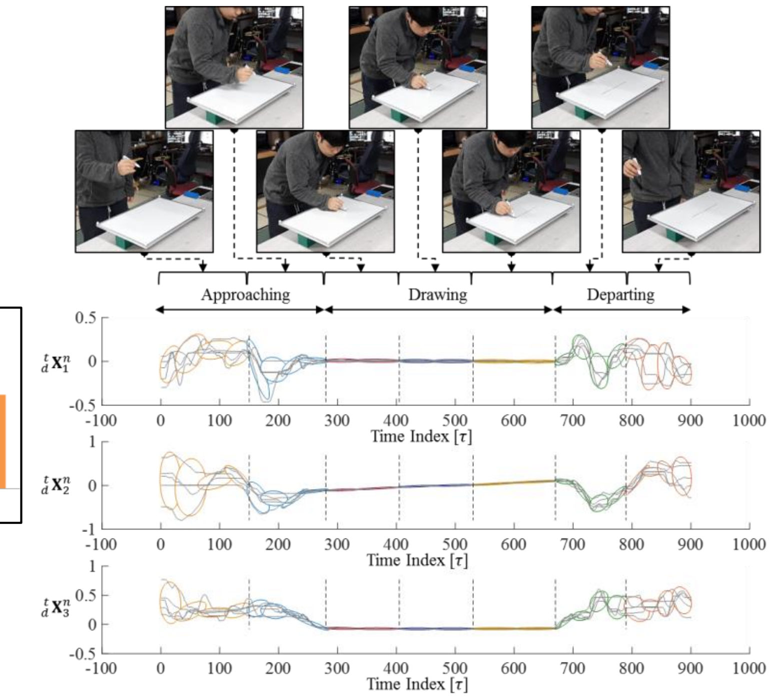


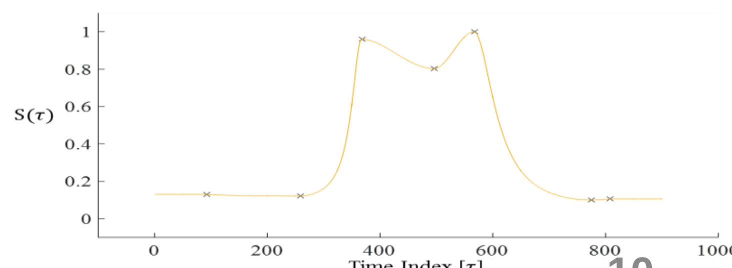






Motion Significance





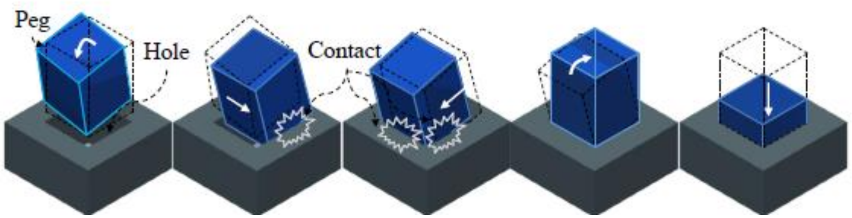
I. H. Suh et al., "Measuring motion significance and motion complexity," Information Sciences, 388, 84-98, 2017.

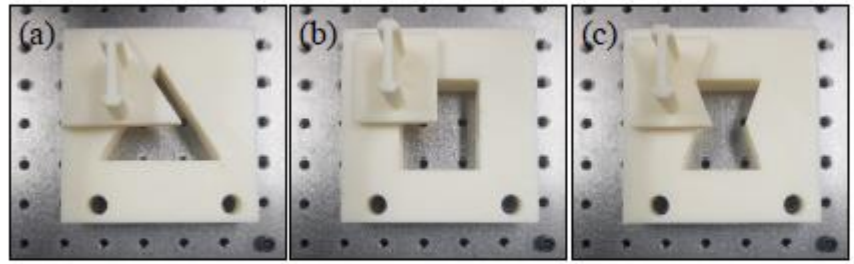
1. What motion will be more complex?

2. What motion skill will be better learned first?

3. What and where to attend to learn from demonstrations?

What motion skill will be better learned first in fitting task?





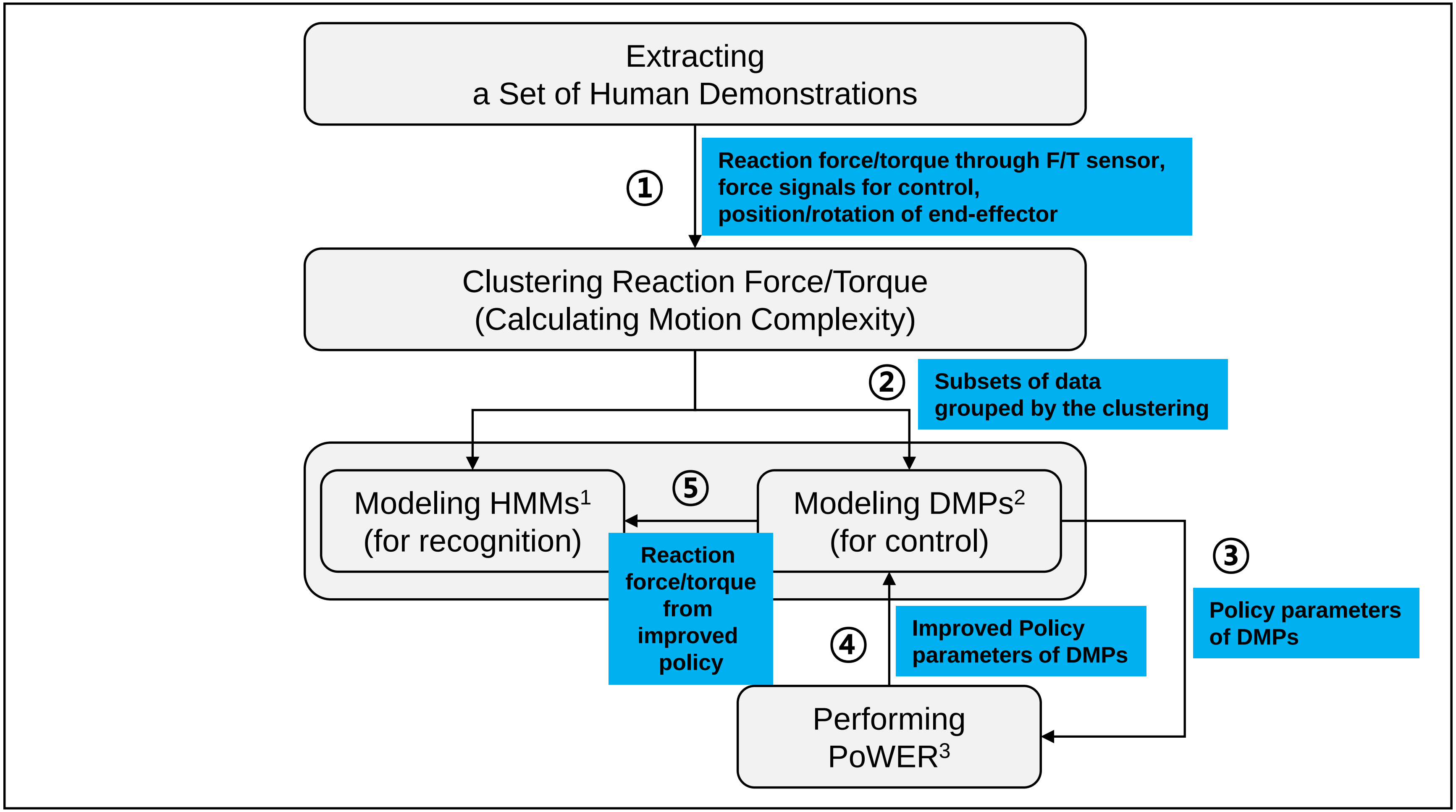
triangle rectangle irregular concave hexagon

Objective: When human demonstrates how to fit a shape, the robot has to learn fitting other two shapes by using pre-demonstrated motion as well as RL.

Q1) What fitting motion skill is more complex among triangle-, rectangle-, and hexagon-shaped fitting??

Q2) For effective learning and effective learning transfer,
Complex one needs to be learned first? Or simpler one needs to be learned first?

Overview of Learning Process



[1]HMM(Hidden Markov Model): to model reaction force/torque according to the directions of inserting pegs

[2]DMP(Dynamic Movement Primitive): to model control signals

[3]PoWER(Policy Learning by Weighting Exploration with the Returns): to improve policy parameters through RL

DMP and PoWER for RL

Representation of Motor Skills

Dynamic Movement Primitives

$$\dot{v} = K(x_g - x) + Dv + (x_g - x_0)\zeta,$$

$$\tau\dot{x} = v,$$

$$\zeta(s) = \frac{\omega_i \psi_i(s) s}{\sum_{i=1}^{L} \psi_i(s)},$$

Reward Function for RL

$$r(t) = \exp\left(\begin{array}{r} -\alpha(|\bar{r}_x^f - r_x^f(t)| + |\bar{r}_y^f - r_y^f(t)|) + |\bar{r}_z^f - r_z^f(t)|) \\ -\beta(|\bar{r}_x^m - r_x^m(t)| + |\bar{r}_y^m - r_y^m(t)| + |\bar{r}_z^m - r_z^m(t)|) \\ -\gamma(|\bar{P}_z - P_z(t)|) \end{array} \right),$$

Representation of Motor Skills

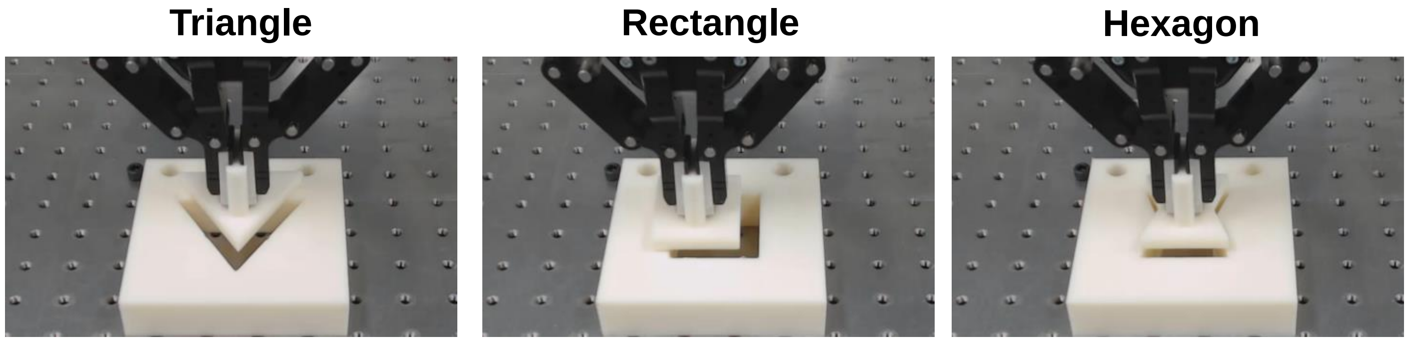
Extension of Policy Learning by Weighting Exploration with the Returns (PoWER) to Optimize and Transfer Motor Skills

Input: initial policy parameters Ω_0
(Here, $a = \Omega^T \Psi(x)$ and T is the original length of initial policy)

Repeat
Sample: Using an initial target x_g in Equation (5),
Generate rollout (x) using action (i.e. DMP)
$a = (\Omega + \varepsilon_t)^T \Psi(x, t)$ with exploration $[\varepsilon_t]_{ij} \sim N(0, \sigma_{ij}^2)$ as stochastic policy and collect all $(t, x_t, a_t, x_{t+1}, \varepsilon_t, r_{t+1})$ for $t = \{1, 2, \ldots, \tilde{T} + 1\}$,
where $\tilde{T} = \mathrm{argmax}_t\, r(t)$ and $x_g = \begin{cases} a(\tilde{T}), & \text{if } r_t > r_{\max} \\ a(T^*), & \text{if } r_t < r_{\max} \end{cases}$.
Here, the value $r_{\max}$ indicates the highest reward in all rollouts, and the target $a(T^*)$ is the action in the rollout of the value $r_{\max}$.
Estimate: Use unbiased estimate of the value function
$\hat{Q}^\pi(x, a, t) = \sum_{\tilde{t}}^{\tilde{T}} r(x_{\tilde{t}}, a_{\tilde{t}}, x_{\tilde{t}+1}, \tilde{t})$.
Reweight: rollouts, discard low-reward rollouts.
Update policy using
$\Omega_{k+1} = \Omega_k + \langle \sum_{t=1}^{\tilde{T}} \varepsilon_t Q^\pi(x, a, t) \rangle / \langle \sum_{t=1}^{\tilde{T}} Q^\pi(x, a, t) \rangle$

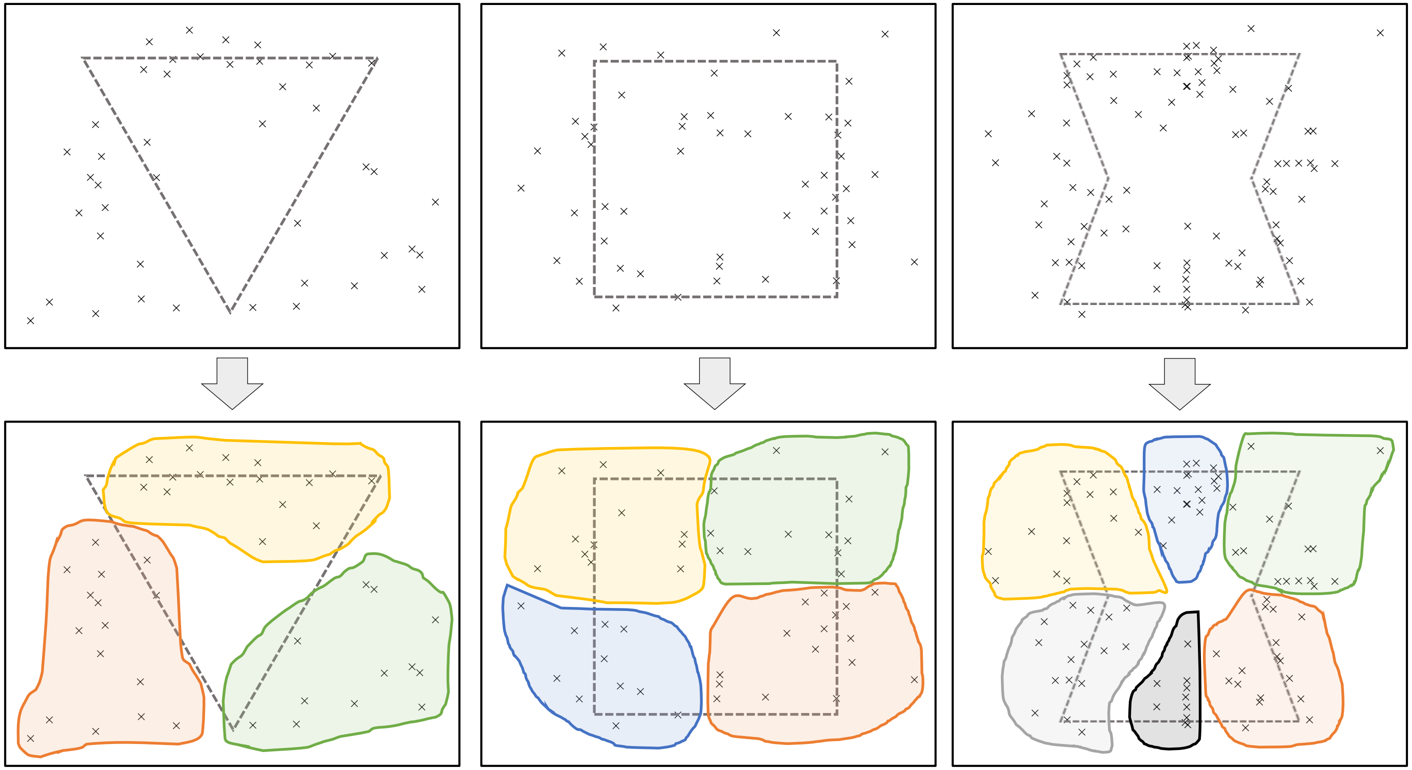
Until convergence $\Omega_{k+1} \approx \Omega_k$

Clustering Reaction F/T Signals in Fitting Task



With only
Reaction F/T signals

x: Initial Robot End-Effector Position



Motion Complexity in Fitting Tasks

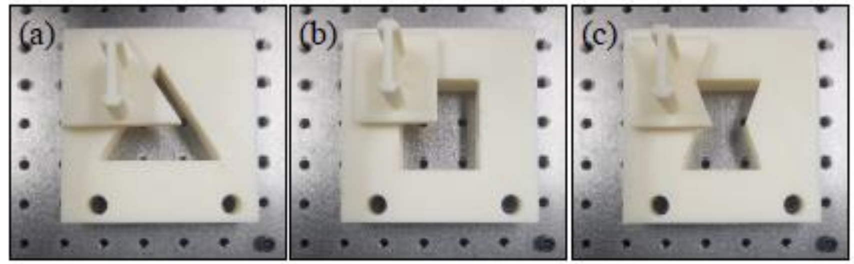
Reaction Force/Torque

Clustering

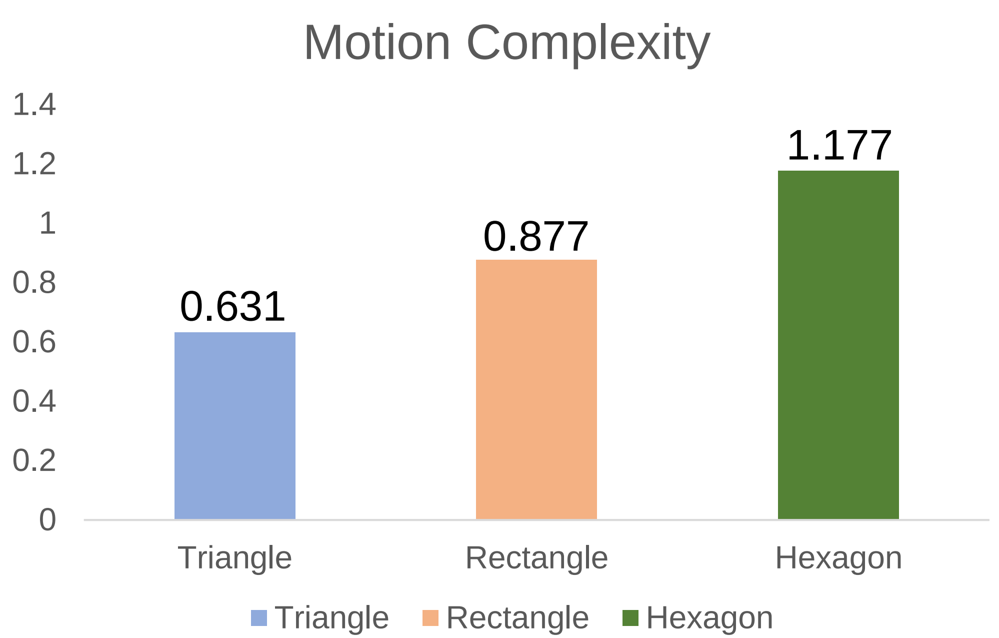
Calculating temporal and spatial entropies in every cluster

Calculating motion complexity in every cluster

Calculating motion complexity of a task by summing all motion complexities

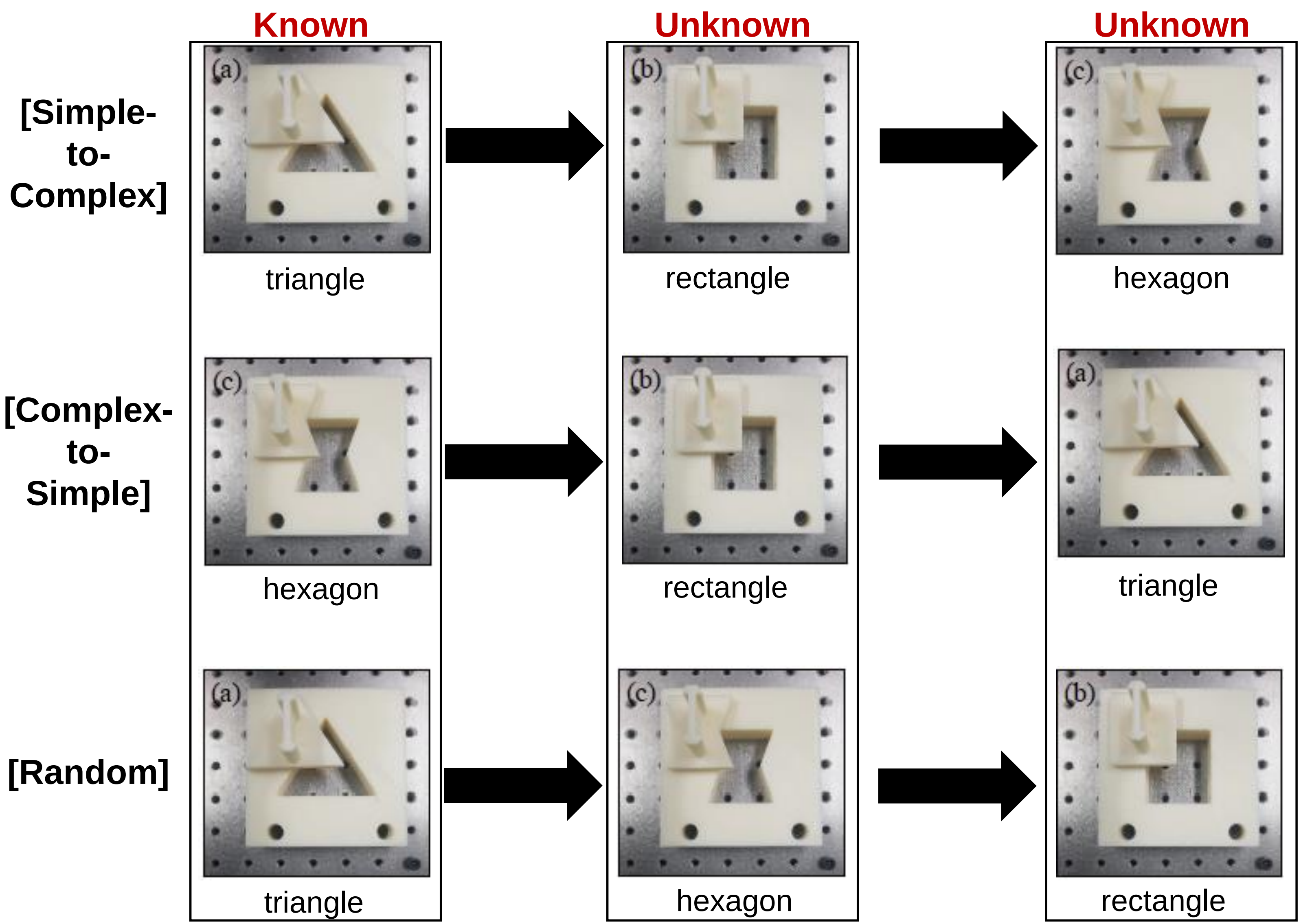


triangle | rectangle | irregular concave hexagon



*** Motion complexity calculated using reaction force/torque signals**

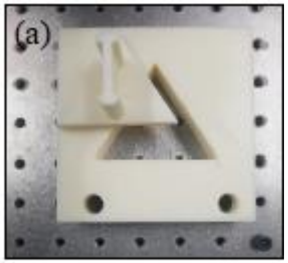
Three Sequences of Task Transfer through RL (1/6)



Three Sequences of Task Transfer through RL (2/6)

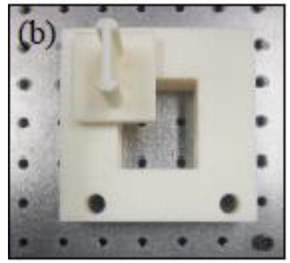
[Simple-to-Complex]

Known



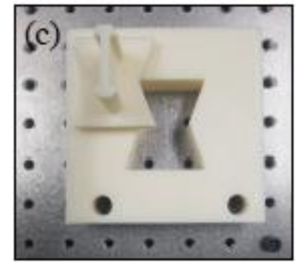
triangle

Unknown



rectangle

Unknown



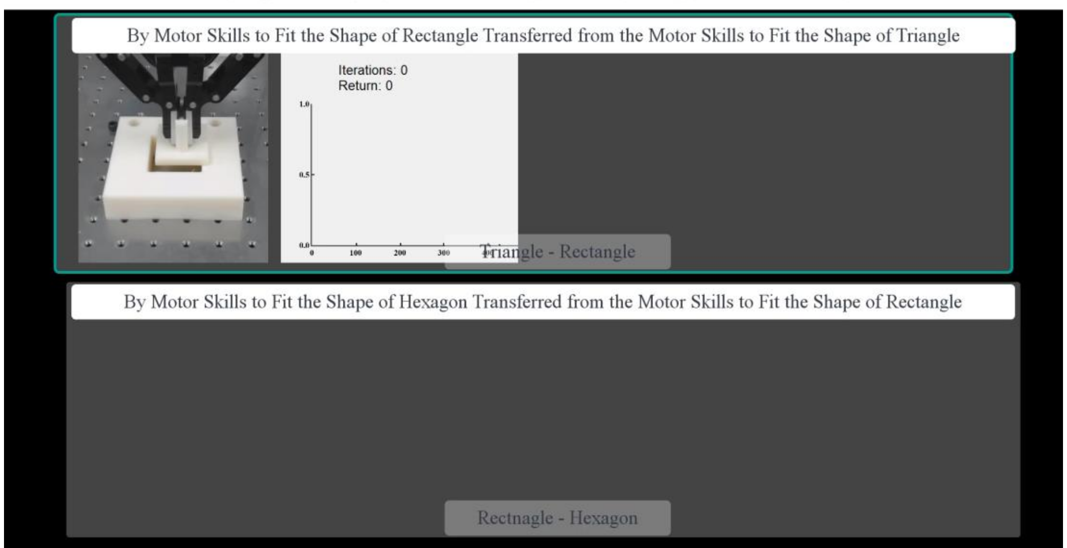
hexagon

Total 368

# of iterations 190 (A)

# of iterations 178 (B)

Thr [Simple-to-Complex] Order (i.e., Triangle - Rectangle - Irregular Concave Hexagon)



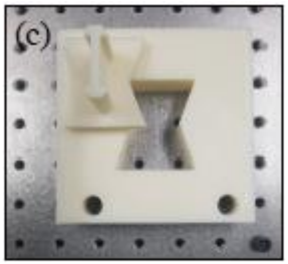
(A)

(B)

Three Sequences of Task Transfer through RL (3/6)

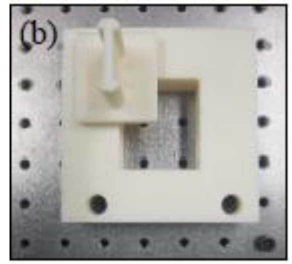
[Complex-to-Simple]

Known



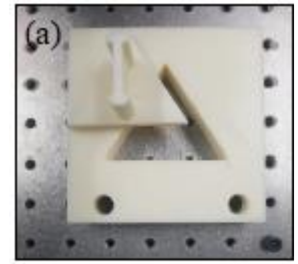
hexagon

Unknown



rectangle

Unknown



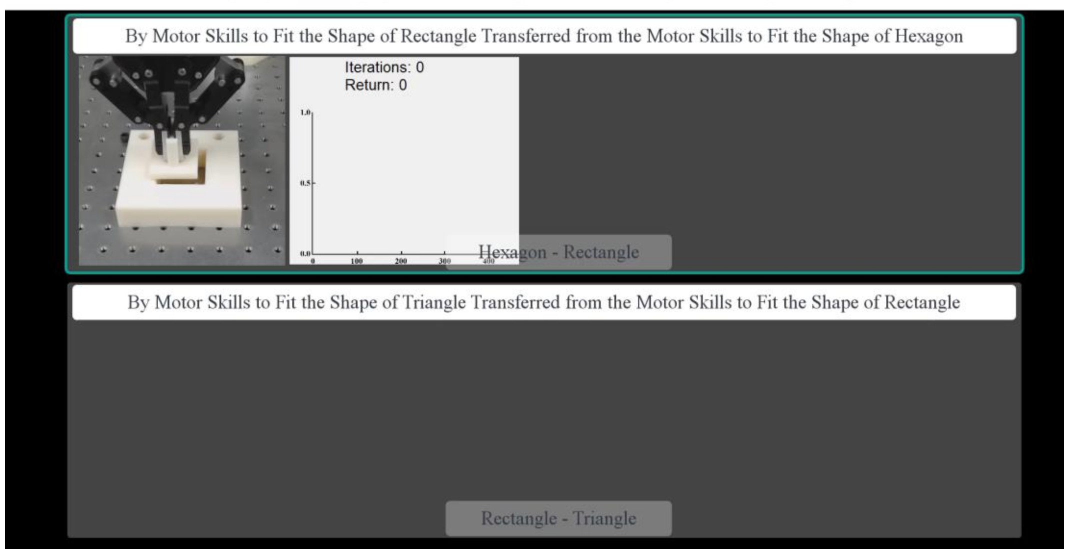
triangle

Total 237

# of iterations 136 (A)

# of iterations 101 (B)

Thr [Complex-to-Simple] Order (i.e., Irregular Concave Hexagon - Rectangle - Triangle)



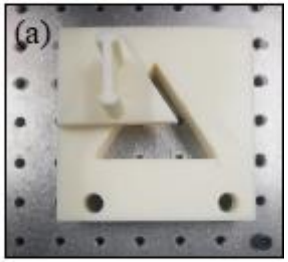
(A)

(B)

Three Sequences of Task Transfer through RL (4/6)

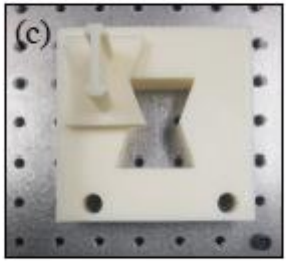
[Random]

Known



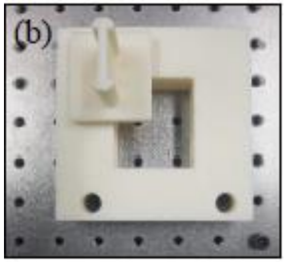
triangle

Unknown



hexagon

Unknown



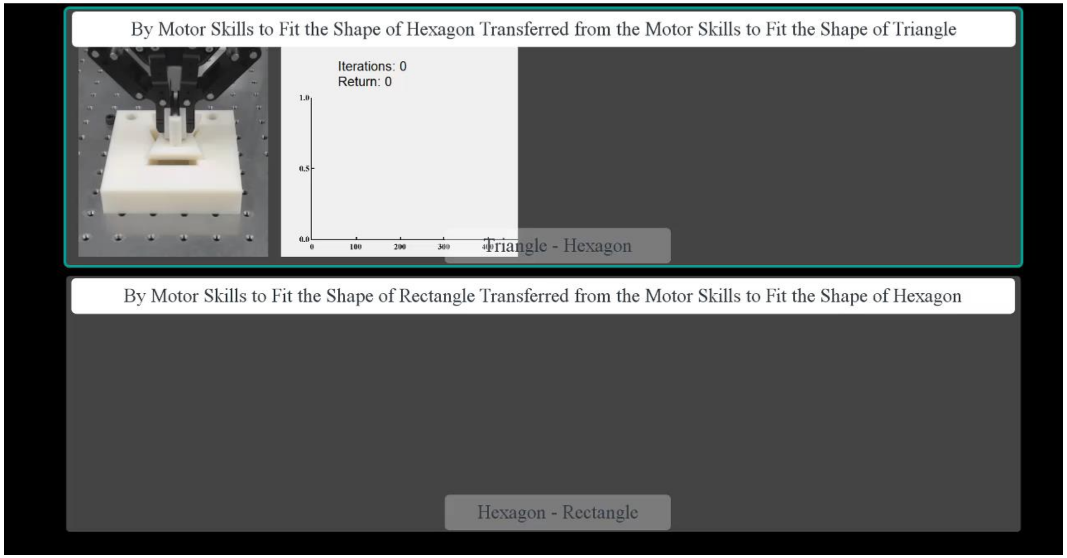
rectangle

Total 539

# of iterations 431 (A)

# of iterations 108 (B)

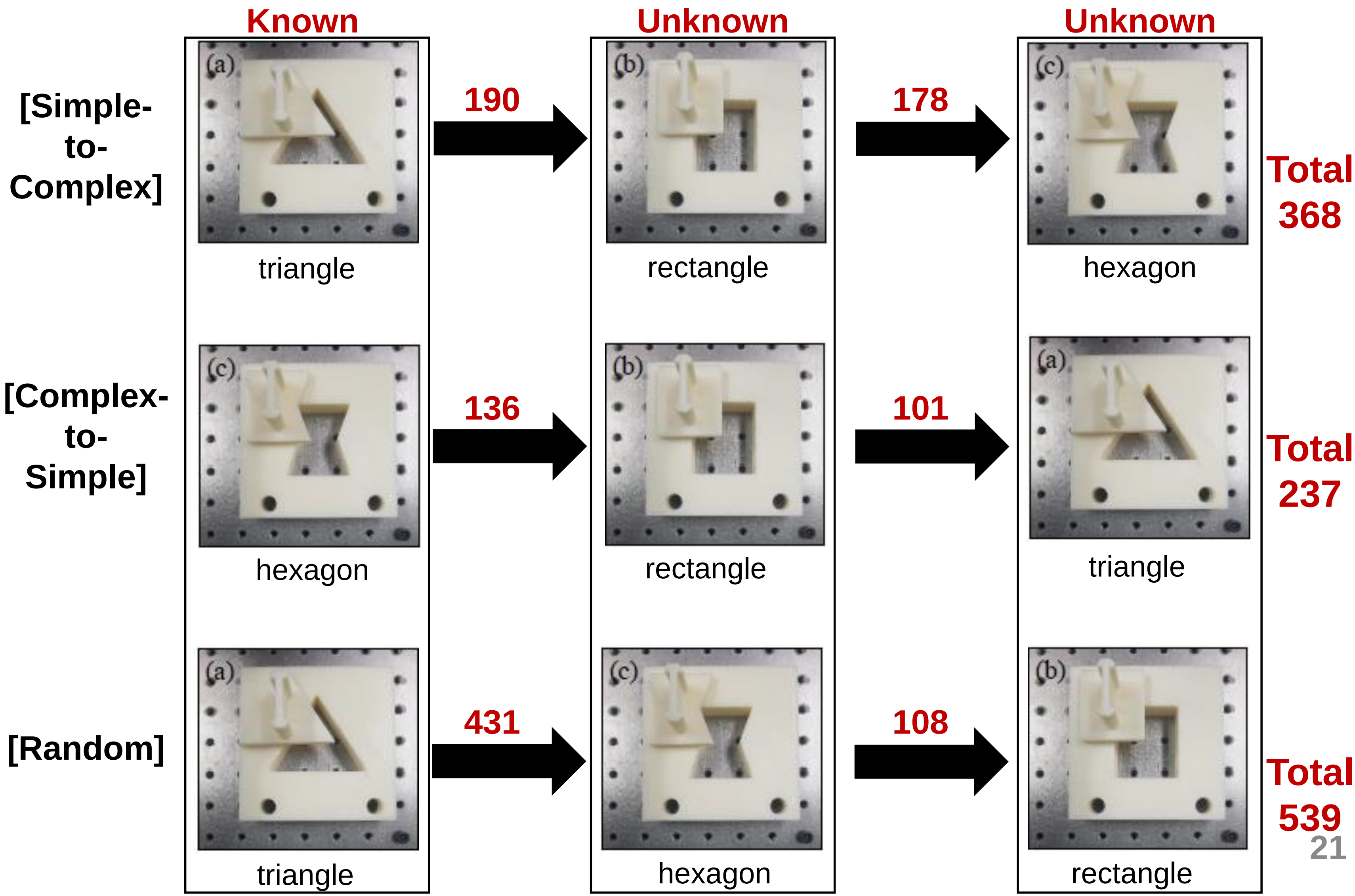
Thr [Random] Order (i.e., Triangle - Irregular Concave Hexagon - Rectangle)



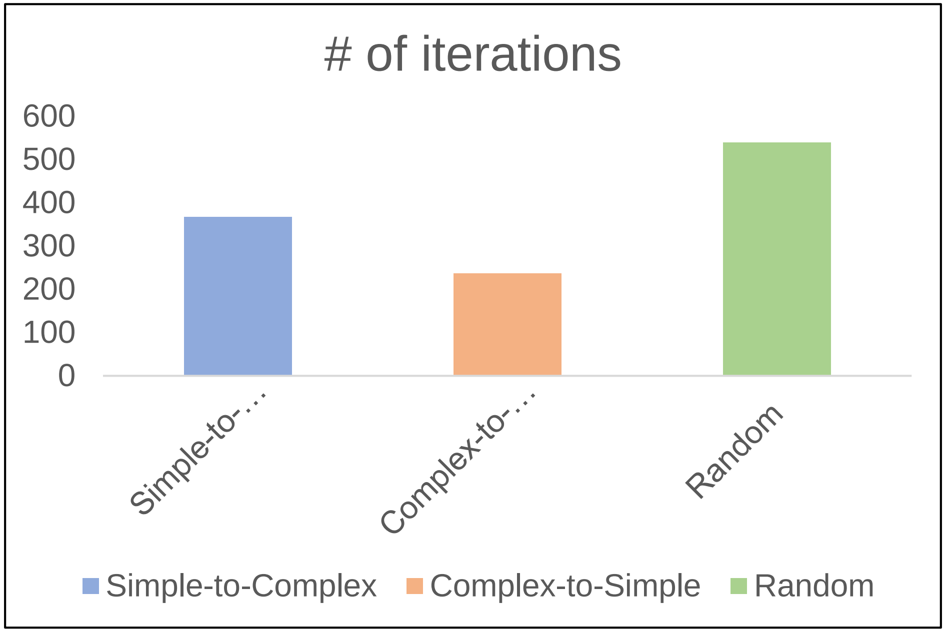
(A)

(B)

Three Sequences of Task Transfer through RL (5/6)

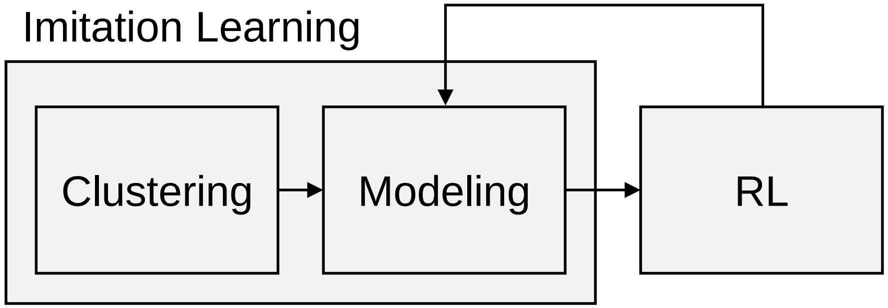


Three Sequences of Task Transfer through RL (6/6)



- **When human can provide demonstrations:**
  Transfer task skills through the sequence of [Complex-to-Simple].
- **When human cannot provide demonstrations:**
  Transfer task skills through the sequence of [Simple-to-Complex].

RL Considering Task Execution Time in Fitting Task

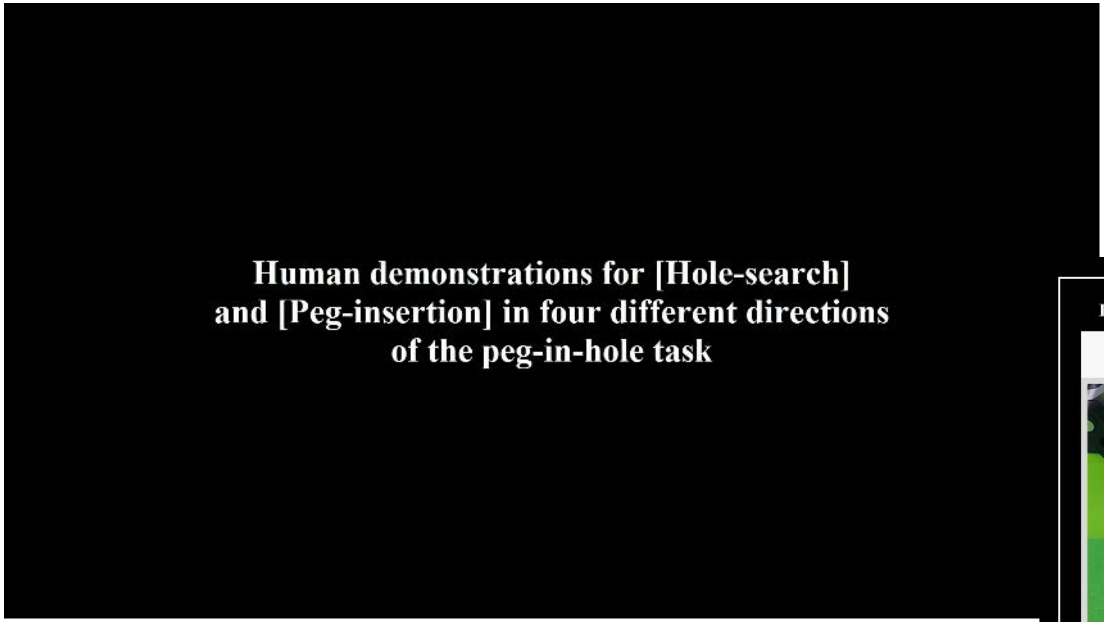


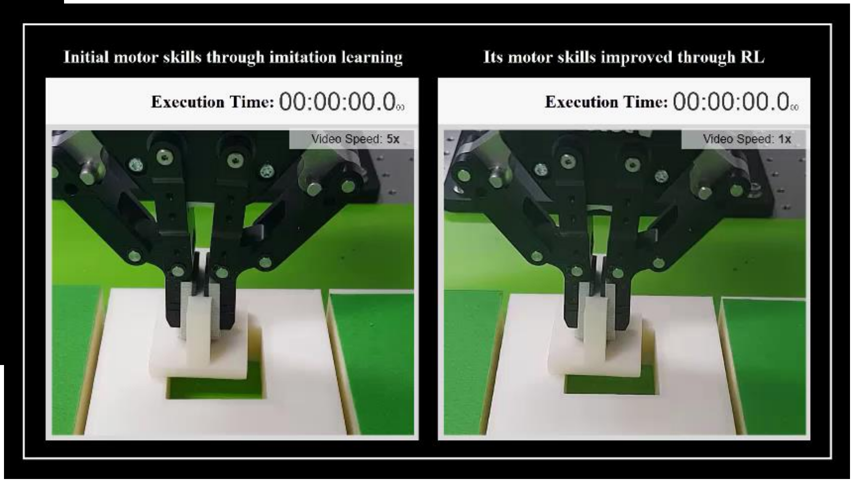
Policy Learning by Weighting Exploration with the Returns

Input: initial policy parameters Ω_0
(Here, $a = \Omega^T \Psi(x)$ and T is the original length of initial policy)

Repeat
Sample: Using an initial target x_g in Equation (1),
Generate rollout (x) using action (i.e. DMP)
$a = (\Omega + \varepsilon_t)^T \Psi(x, t)$ with exploration $[\varepsilon_t]_{ij} \sim N(0, \sigma_{ij}^2)$ as stochastic policy and collect all $(t, x_t, a_t, x_{t+1}, \varepsilon_t, r_{t+1})$
for $t = \{1,2, \ldots, \tilde{T} + 1\}$,
where $\tilde{T} = t_A$ or $\tilde{T} = t_F$ and $x_g = a(T)$; [Improvement] or
$\tilde{T} = \text{argmax}_t\ r(t)$ and $x_g = \begin{cases} a(\tilde{T}), & \text{if } r_t > r_{\max} \\ a(T^*), & \text{if } r_t < r_{\max} \end{cases}$;
[Generalization]. Here, the value $r_{\max}$ indicates the highest reward in all rollouts,
and the target $a(T^*)$ is the action in the rollout of the value $r_{\max}$.
Estimate: Use unbiased estimate of the value function
$$\hat{Q}^\pi(x, a, t) = \sum_{\tilde{t}}^{\tilde{T}} r(x_{\tilde{t}}, a_{\tilde{t}}, x_{\tilde{t}+1}, \tilde{t}).$$
Reweight: rollouts, discard low-reward rollouts.
Update policy using
$$\Omega_{k+1} = \Omega_k + \langle \textstyle\sum_{t=1}^{\tilde{T}} \varepsilon_t Q^\pi(x, a, t) \rangle / \langle \sum_{t=1}^{\tilde{T}} Q^\pi(x, a, t) \rangle$$

Until convergence $\Omega_{k+1} \approx \Omega_k$





1. What motion will be more complex?

2. What motion skill will be better learned first?

3. What and where to attend to learn from demonstrations?

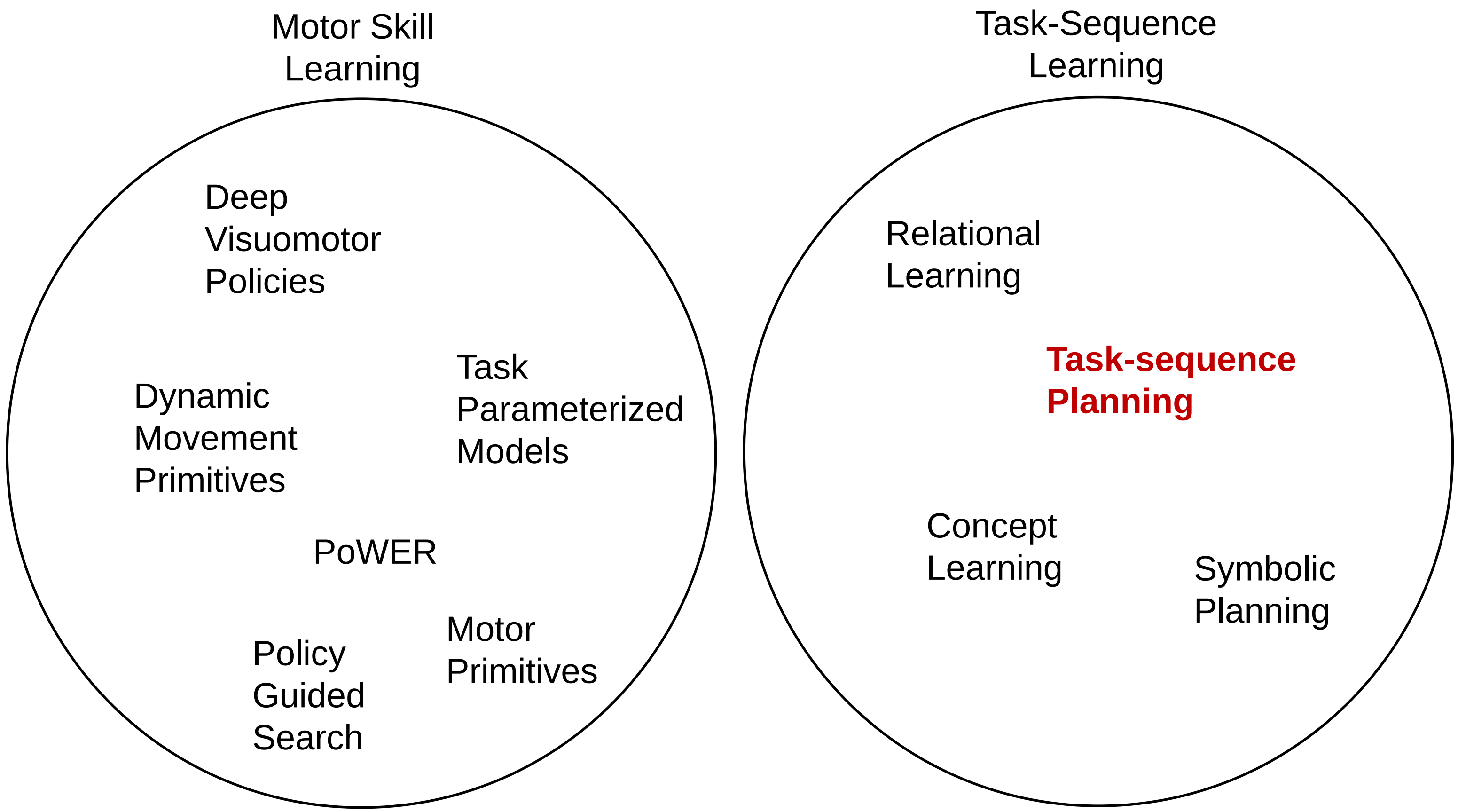
Where to Attend? What to Attend?

[00:00:45]

This ape should be able to find and learn attentive
and significant intentions(joint relations) in the human demonstration.

How to find this? and By what measure?

Two Paradigms of Existing PbD Approaches



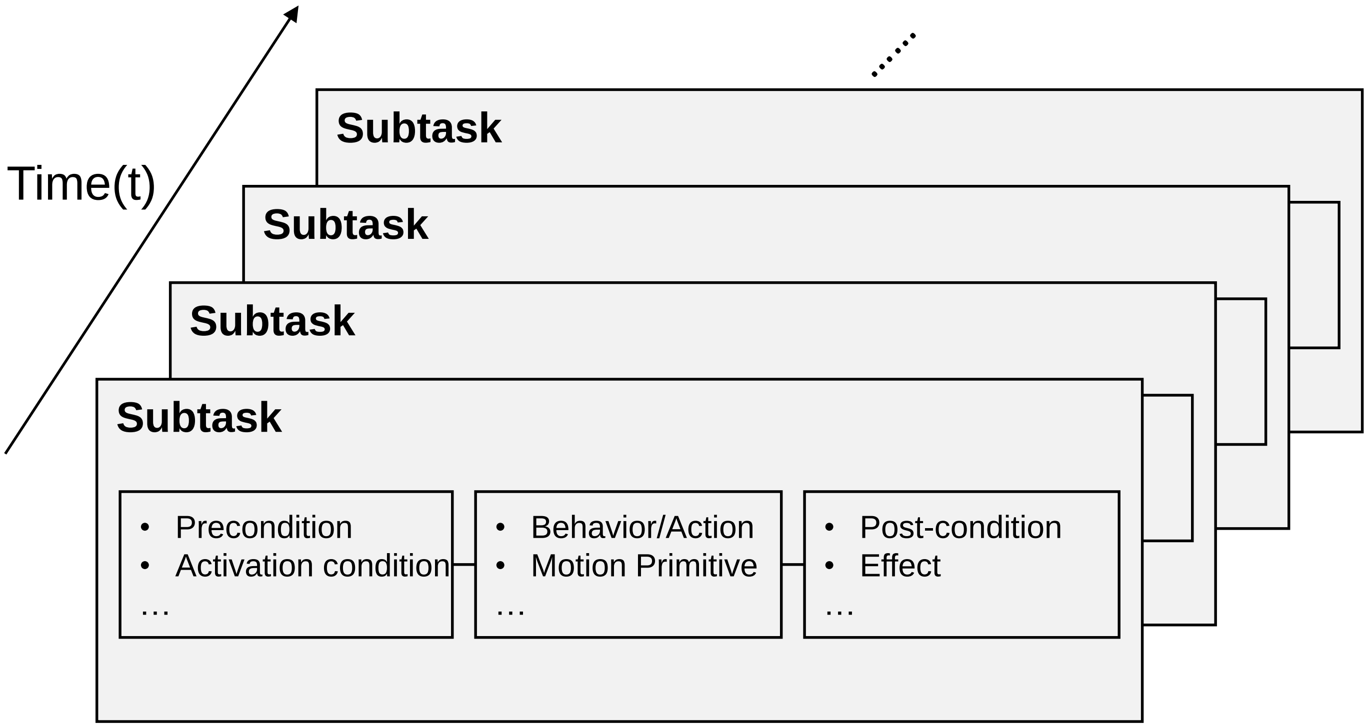
- Trajectory Learning
- Motion Optimization/ Generalization
- Law-level Learning

…

- Sequential behaviors
- Serial order in behavior
- High-level Learning

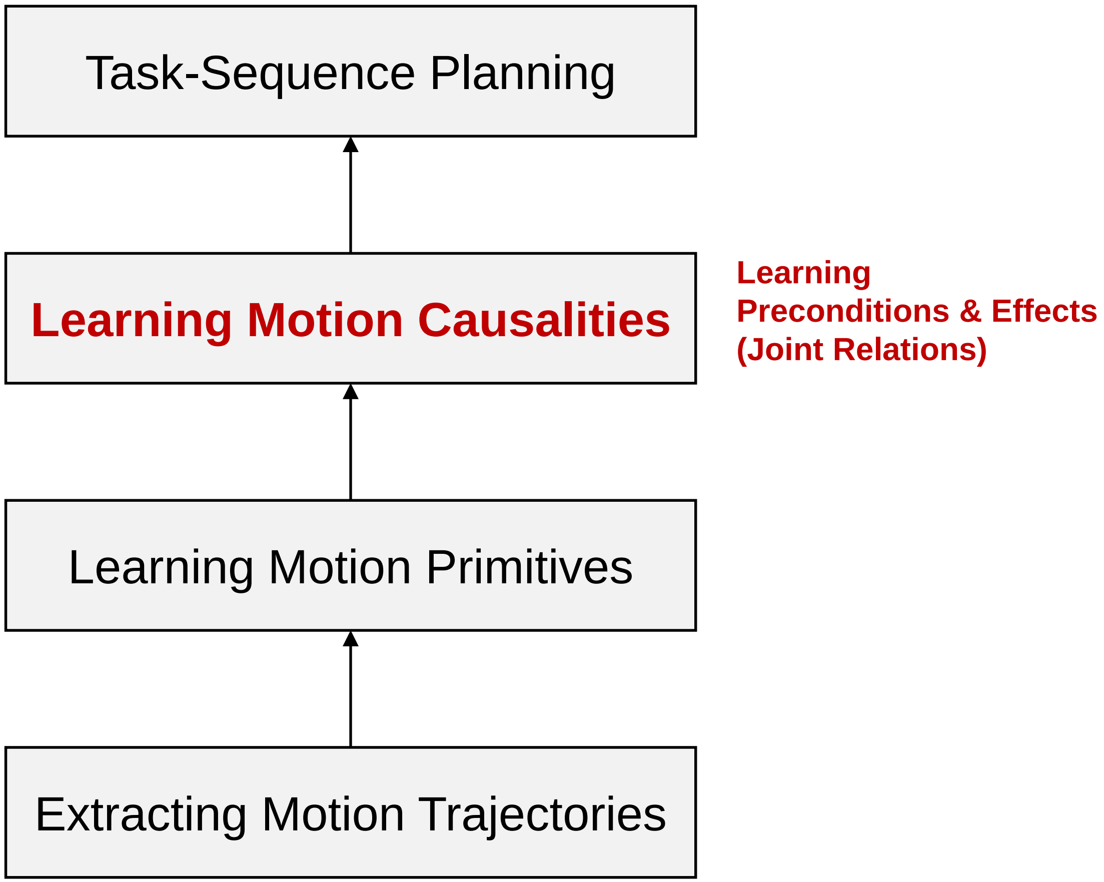
…

Task-sequence Learning
: Learning Preconditions&Effects

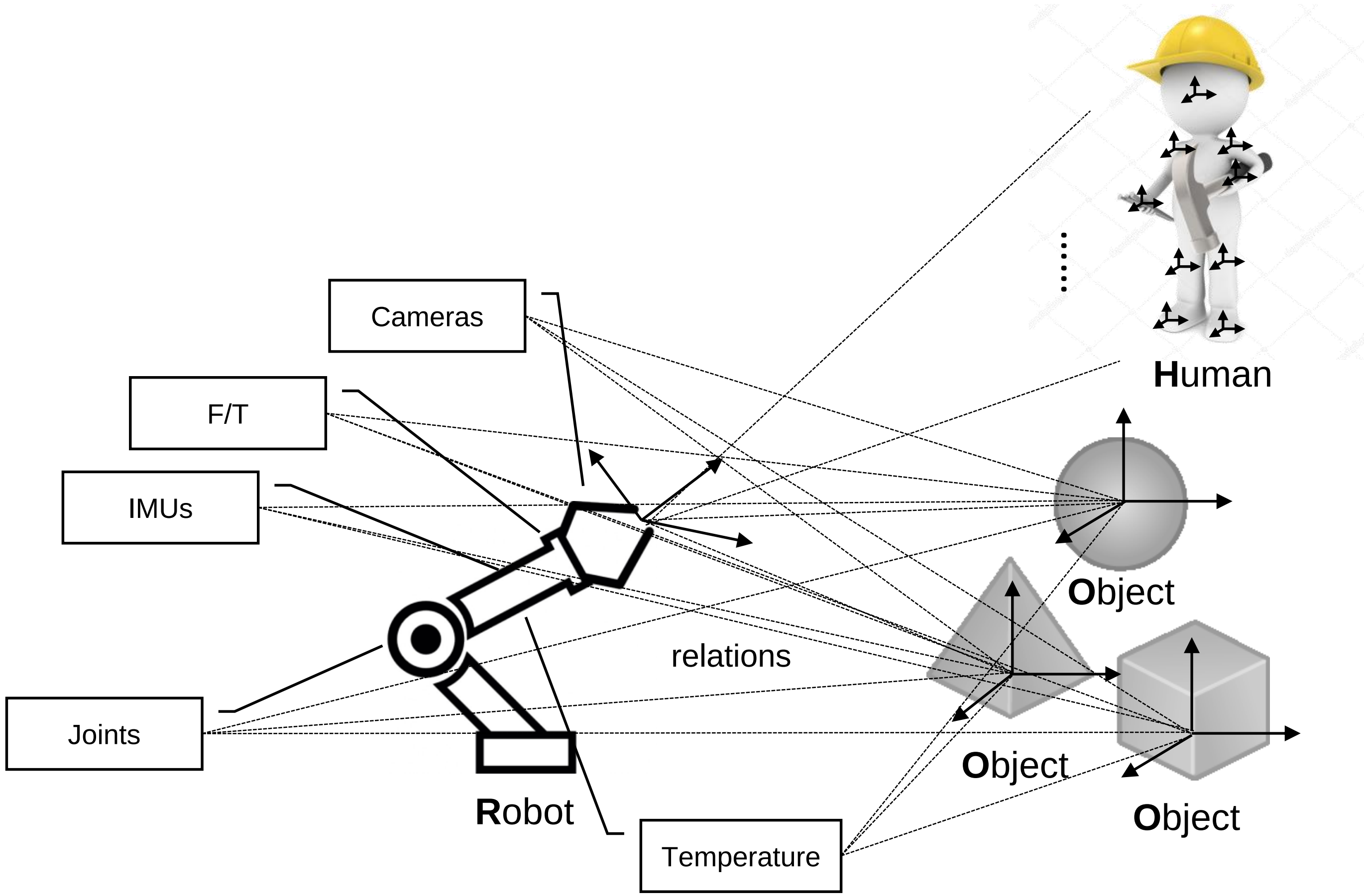


Task-Sequence Learning/Planning

Conceptual Process for Task-Sequence Planning in PbD

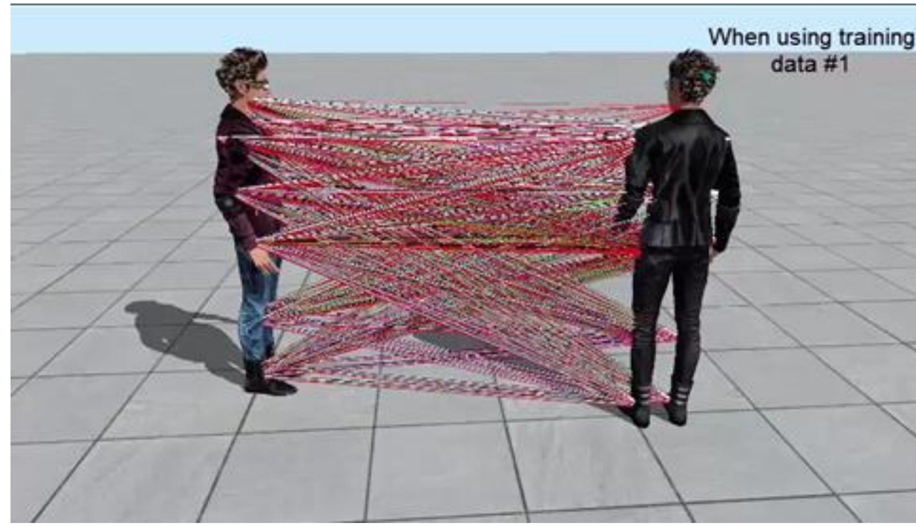


Joint Motion Significance: To Find Significant Joint Relations



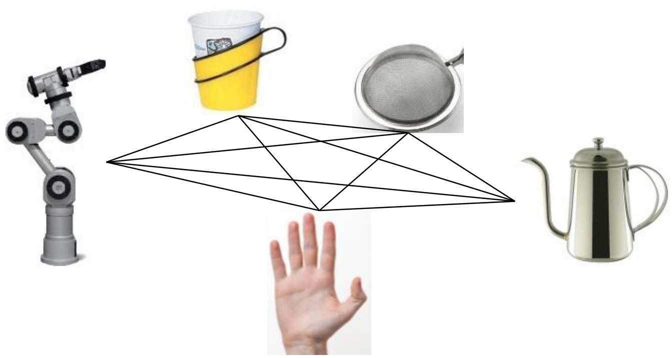
To find significant joint relations from tons of joint relations

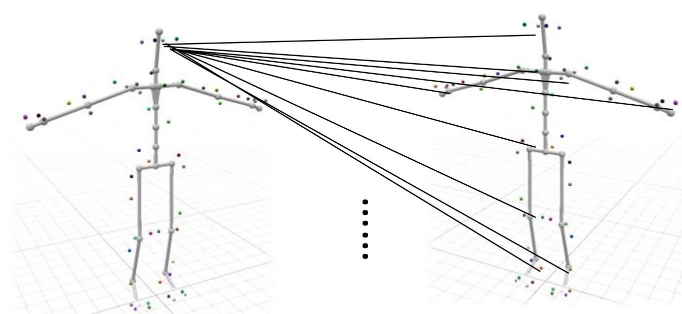
How to Find Significant Joint Relations



(3x3x3x3x3)x2 =**486 joint relations**
(3D positions and 3D rotations per object)

19x19x6 = **2,166 joint relations**
(19 joints x 6dimensions per human)





3~9 significant joint relations

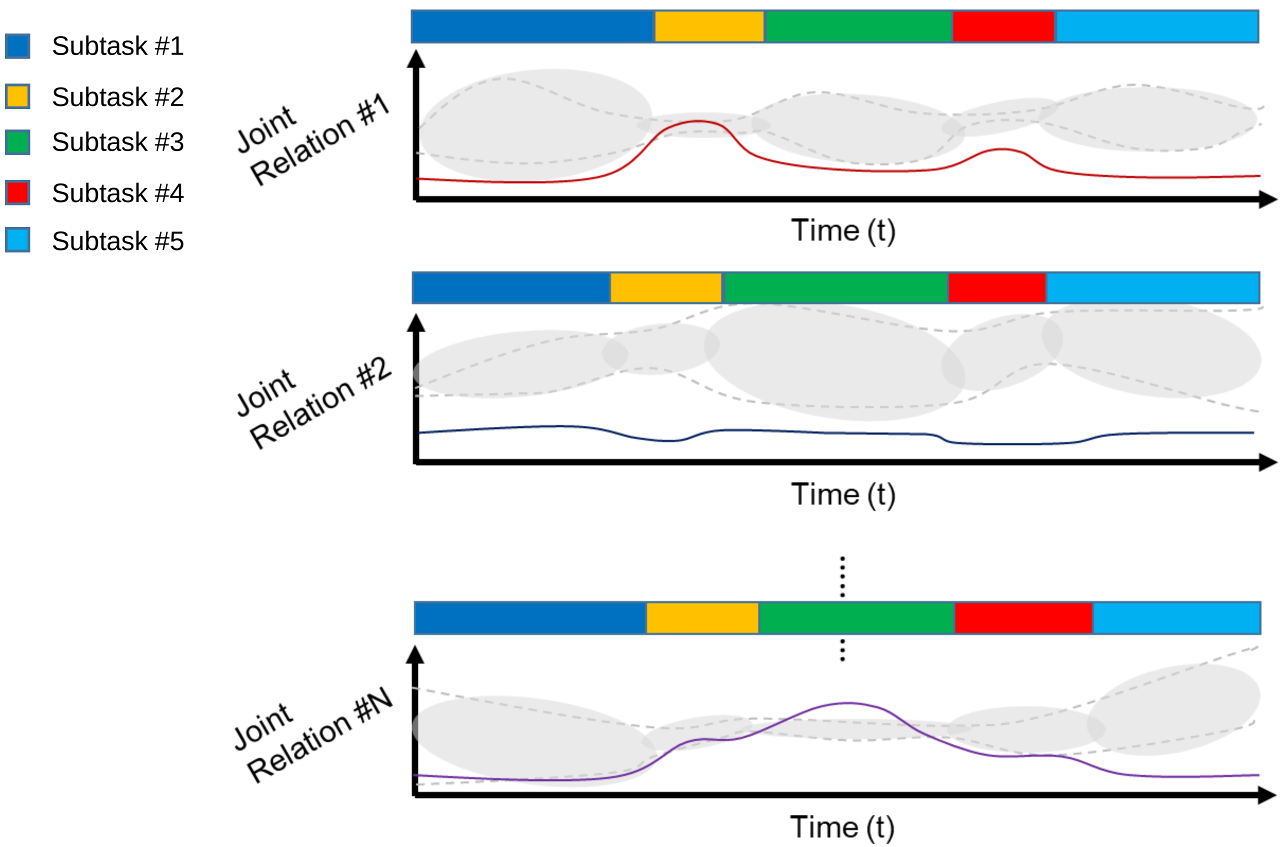
9~12 significant joint relations

By Joint Motion Complexity and Joint Motion Significance Measures

1. Calculate the joint significance and joint complexity measures of all individual joint relations
2. Segment a whole task into subtasks

3. Select Top *K*

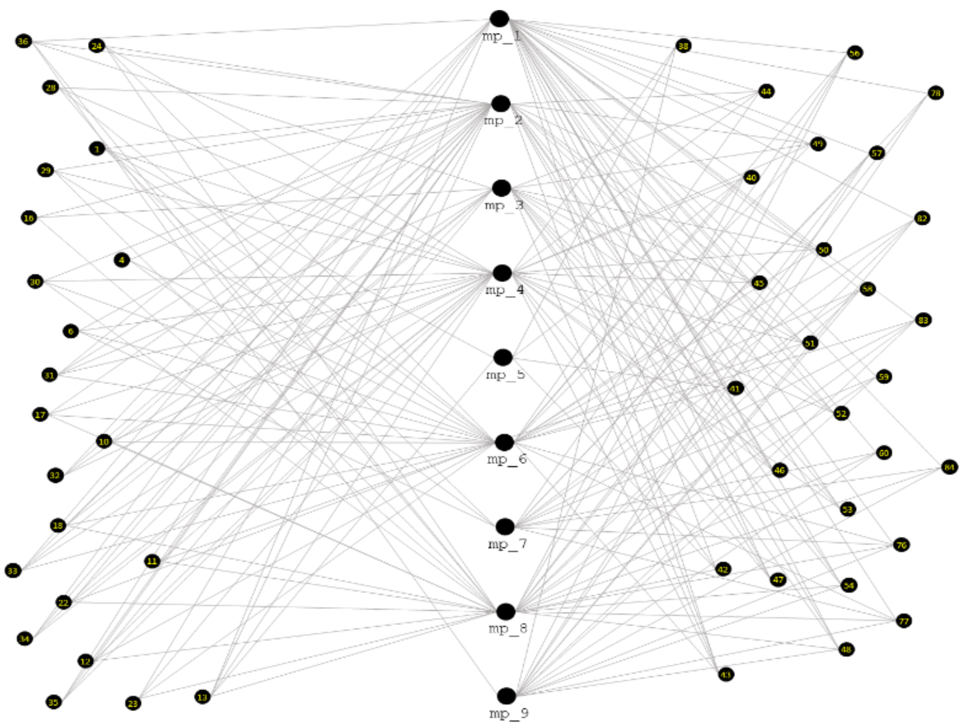
[Example]



Top *K* joint relations in every subtask

Representation of Joint Relations

preconditions   motion primitives   post-conditions



: a variable for motion primitives mp_*

: significant variables

By PDDL (Planning Domain Definition Language)

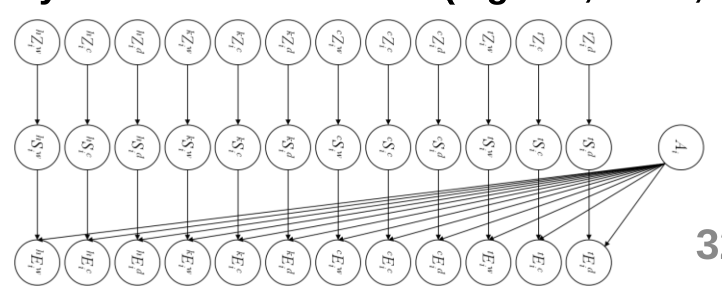
Problem file
; initial configuration
; goal configuration

```
PDDL Problem File
(define (problem ManipulationProblem)
(:domain Manipulation-Task)
(:objects)
(:init
(= (Pcbzig-Pcb1-XX) 0.2);
(= (Pcbzig-Pcb2-XX) 0.1568);
(= (Pcb1-Pcb2-XX) 0.0044);
(= (Pcb1-Pcb deck-XX) -0.0119);
(= (Pcb2-Pcb deck-XX) -0.0163);
(= (Pcb1-Pcb2-YY) -0.0432);
(= (Pcb1-Pcb deck-YY) -0.0393);
(= (Pcb2-Pcb deck-YY) 0.0046);
(= (Pcbzig-Pcb1-ZZ) 0.0836);
(= (Pcbzig-Pcb2-ZZ) 0.0792);
(= (Pcbzig-Pcb deck-ZZ) 0.0586);
(= (Pcb1-Pcb2-ZZ) -0.0013);
(= (Pcb1-Pcb deck-ZZ) -0.0229);
(= (Pcb2-Pcb deck-ZZ) -0.0214);
(= (Pcbzig-Pcb deck-RxRx) 0.0039);
(= (Pcb1-Pcb deck-RxRx) -0.0308);
(= (Pcb2-Pcb deck-RxRx) -0.0243);
(= (Pcbzig-Pcb1-RyRy) 0.039);
(= (Pcbzig-Pcb2-RyRy) 0.0302);
(= (Pcb1-Pcb2-RyRy) -0.0066);
(= (Pcb2-Pcb deck-RyRy) -0.019);
(= (Pcbzig-Pcb1-RzRz) -1.6332);
(= (Pcbzig-Pcb2-RzRz) -1.6012);
(= (Pcbzig-Pcb deck-RzRz) 0.0005);
(= (Pcb1-Pcb2-RzRz) 0.0326);
(= (Pcb2-Pcb deck-RzRz) 1.6041);
)
```

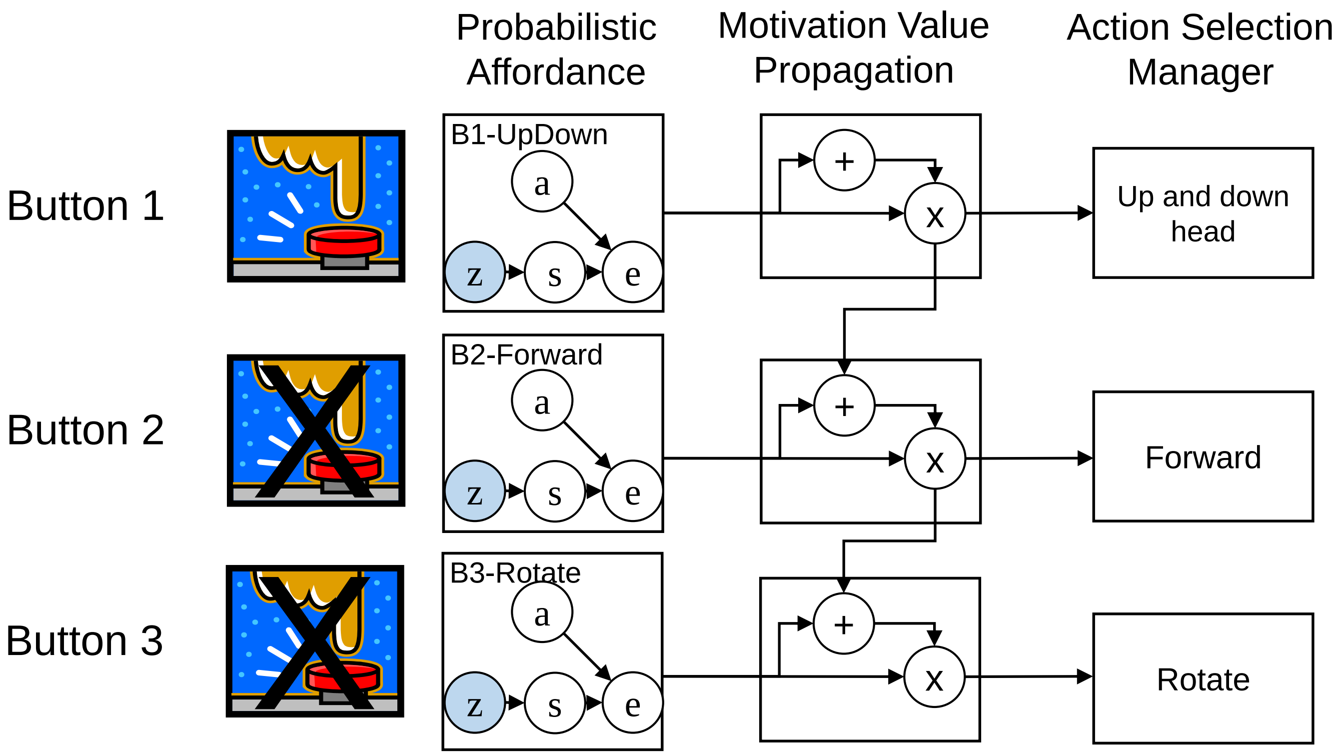
Domain file
; actions (preconditions, action label, effects)

```
PDDL Domain File
(Pcbzig-Pcb2-RzRz)
(Pcbzig-Pcb deck-RzRz)
(Pcb1-Pcb2-RzRz)
(Pcb1-Pcb deck-RzRz)
(Pcb2-Pcb deck-RzRz)
)
(:action primitive1
:parameters ()
:precondition (and
(>= (Pcbzig-Pcb1-XX) 0.1639) (<= (Pcbzig-Pcb1-XX) 0.21)
(>= (Pcbzig-Pcb2-XX) 0.1348) (<= (Pcbzig-Pcb2-XX) 0.1668)
(>= (Pcb1-Pcb2-XX) -0.0182) (<= (Pcb1-Pcb2-XX) 0.0434)
(>= (Pcb1-Pcb deck-XX) -0.0219) (<= (Pcb1-Pcb deck-XX) 0.0212)
(>= (Pcb2-Pcb deck-XX) -0.0318) (<= (Pcb2-Pcb deck-XX) 0.0131)
(>= (Pcb1-Pcb2-YY) -0.0532) (<= (Pcb1-Pcb2-YY) -0.0081)
(>= (Pcb1-Pcb deck-YY) -0.0689) (<= (Pcb1-Pcb deck-YY) 0.0048)
(>= (Pcb2-Pcb deck-YY) -0.0311) (<= (Pcb2-Pcb deck-YY) 0.0454)
(>= (Pcbzig-Pcb1-ZZ) 0.0701) (<= (Pcbzig-Pcb1-ZZ) 0.0962)
(>= (Pcbzig-Pcb2-ZZ) 0.0674) (<= (Pcbzig-Pcb2-ZZ) 0.0915)
(>= (Pcbzig-Pcb deck-ZZ) 0.0461) (<= (Pcbzig-Pcb deck-ZZ) 0.0713)
(>= (Pcb1-Pcb2-ZZ) -0.0306) (<= (Pcb1-Pcb2-ZZ) 0.0228)
(>= (Pcb1-Pcb deck-ZZ) -0.0514) (<= (Pcb1-Pcb deck-ZZ) -0.0129)
(>= (Pcb2-Pcb deck-ZZ) -0.0462) (<= (Pcb2-Pcb deck-ZZ) -0.0054)
(>= (Pcbzig-Pcb deck-RxRx) -0.0143) (<= (Pcbzig-Pcb deck-RxRx) 0.0359
(>= (Pcb1-Pcb deck-RxRx) -0.042) (<= (Pcb1-Pcb deck-RxRx) -0.0195)
(>= (Pcb2-Pcb deck-RxRx) -0.0419) (<= (Pcb2-Pcb deck-RxRx) -0.0143)
(>= (Pcbzig-Pcb1-RyRy) -0.0076) (<= (Pcbzig-Pcb1-RyRy) 0.0569)
(>= (Pcbzig-Pcb2-RyRy) 0.0108) (<= (Pcbzig-Pcb2-RyRy) 0.0437)
(>= (Pcb1-Pcb2-RyRy) -0.0166) (<= (Pcb1-Pcb2-RyRy) 0.0106)
(>= (Pcb2-Pcb deck-RyRy) -0.0392) (<= (Pcb2-Pcb deck-RyRy) -0.0072)
```

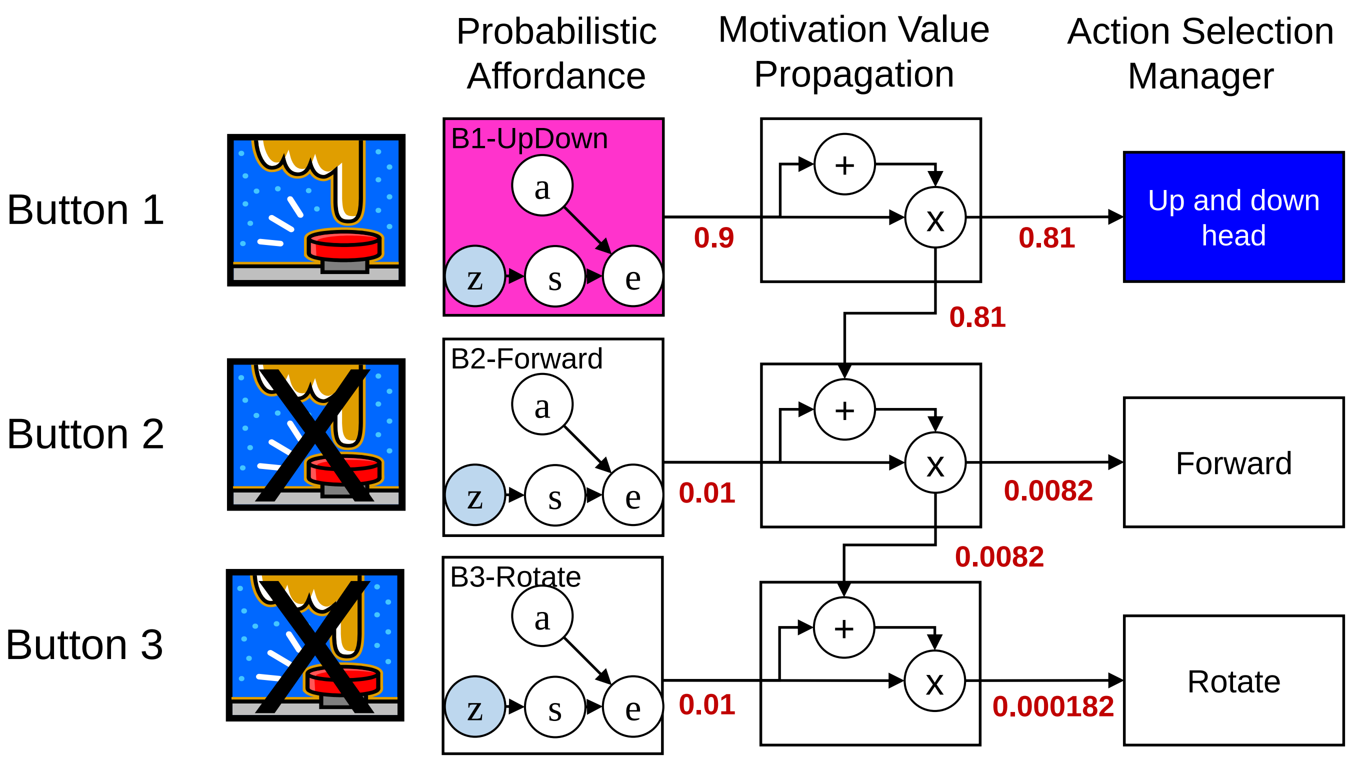
By Probabilistic Models (e.g. BN, HMM, etc.)



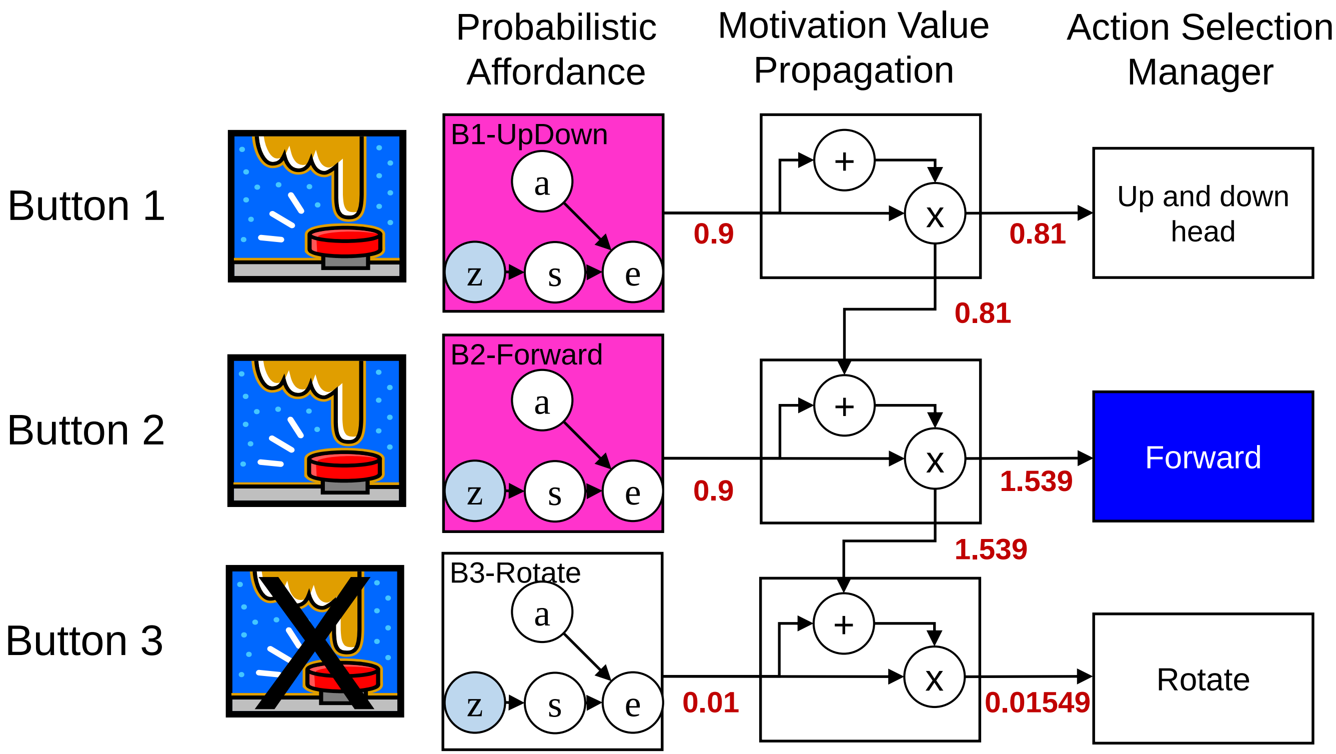
Action Selection for Goal-oriented Task-sequence Planning (1/4)



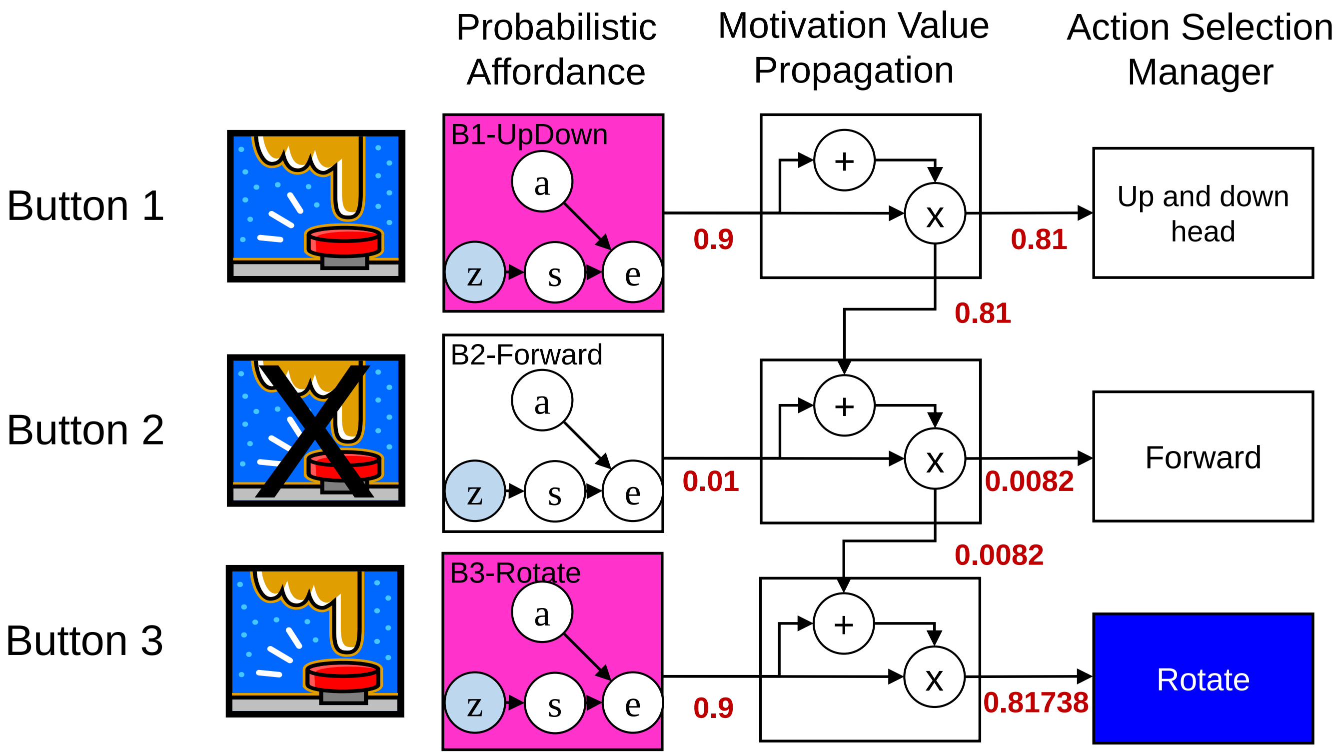
Action Selection for Goal-oriented Task-sequence Planning (2/4)

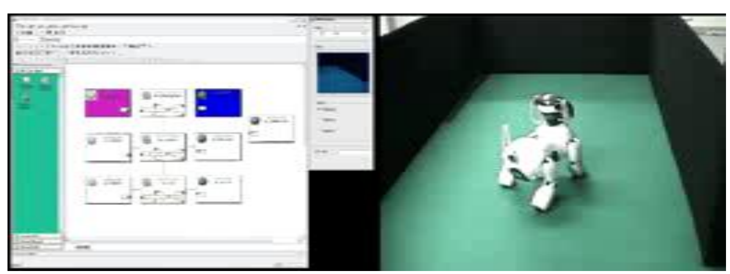


Action Selection for Goal-oriented Task-sequence Planning (3/4)



Action Selection for Goal-oriented Task-sequence Planning (4/4)





Tea-Service Task

Case I: a human snatches a teabag from the robot on the way to delivering it into a cup.

Case II: a human delivers a teabag into a cup while the robot is approaching the teabag for grasping it.

[00:00:18] x6

[00:00:14] x6

[00:00:11] x6

Case III: a human directly moves to a cup while the robot pours the water into the cup.

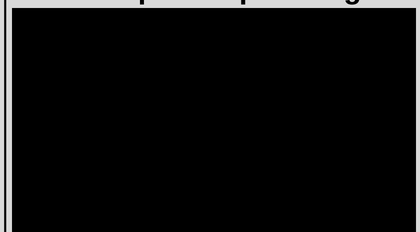
Human-Robot Interaction Game Task

Green Wheel:  Blue Wheel:  Steel Bar: 

Human-Human Interaction

without human interaction

Task-sequence planning with the other human



* this white-coated guy delivers a green wheel instead of the black-coated guy.

*this white-coated guy puts a green wheel back while the black-coated guy is approaching a blue wheel .

* this white-coated guy delivers a blue wheel instead of the black-coated guy.

Human-Robot Interaction

without human interaction

Task-sequence planning with human

* this guy delivers a green wheel instead of the robot.

*this guy puts a green wheel back while the robot is approaching a blue wheel .

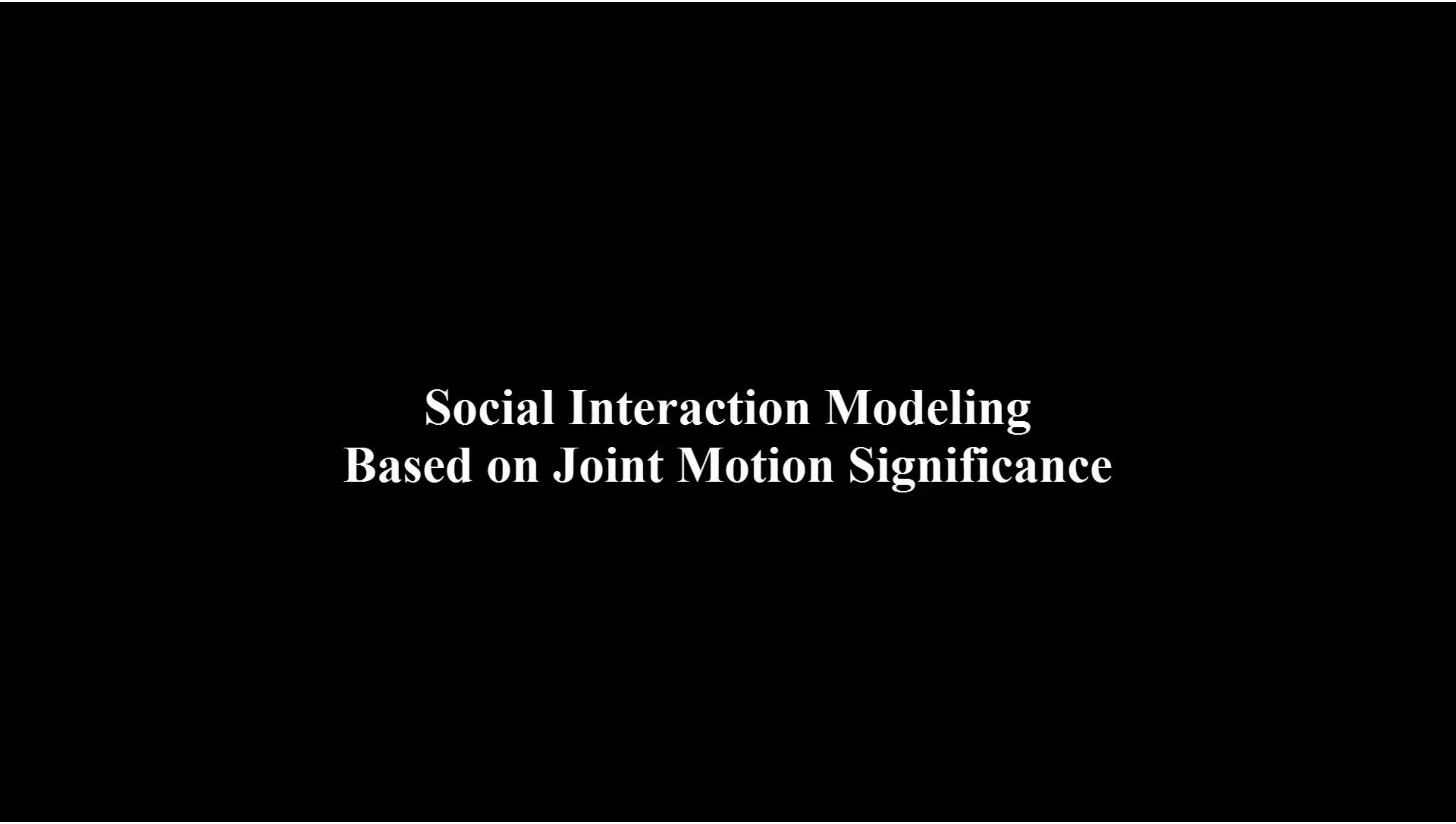
* this guy delivers a blue wheel instead of the robot.

Human-Virtual Avatar Interaction : Social Interaction (1/5)

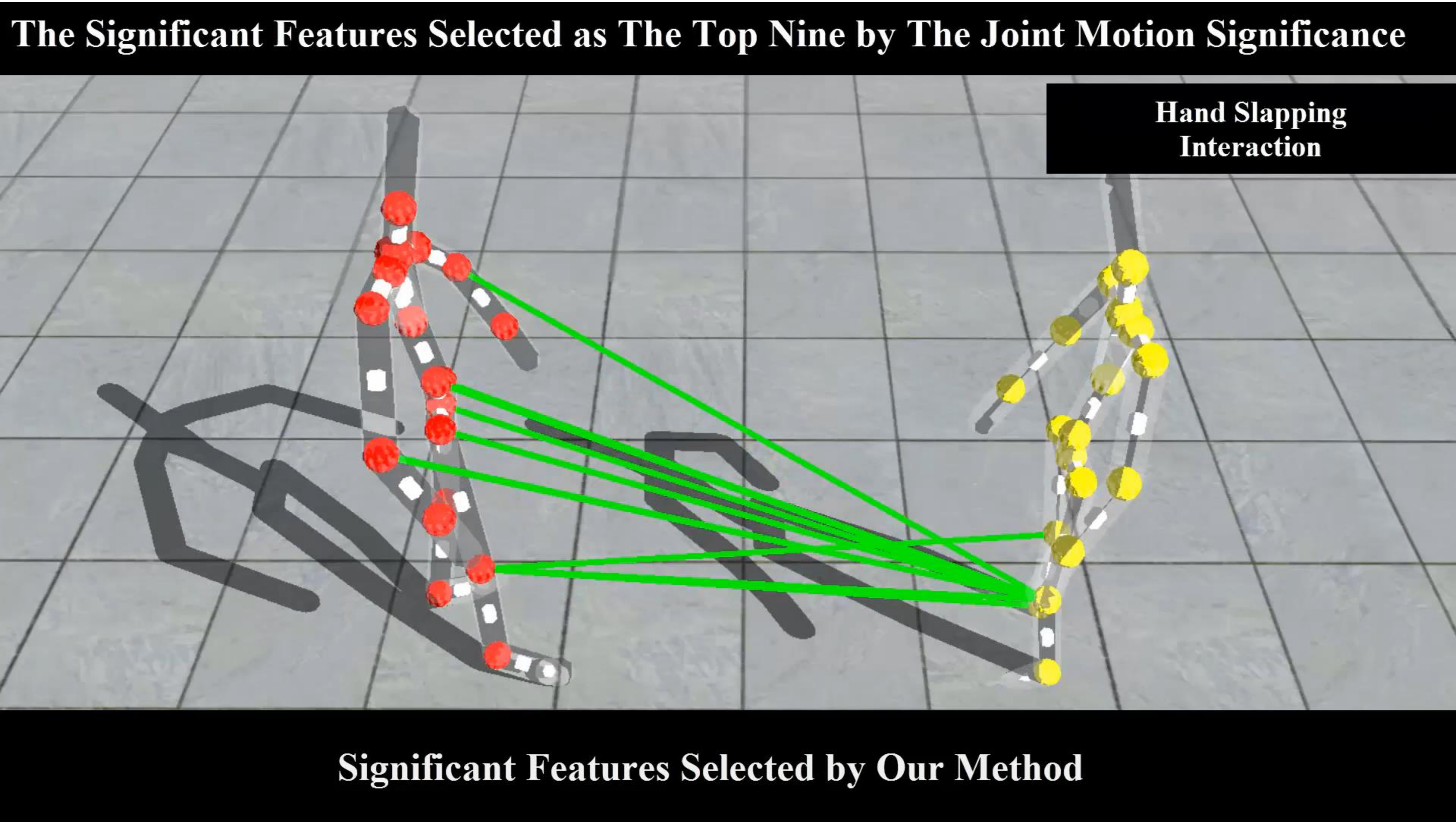
Training Data of Five Social Interaction:

Hand Slapping, Hand Shaking, Shoulder Holding, Object Passing, and Target Kicking

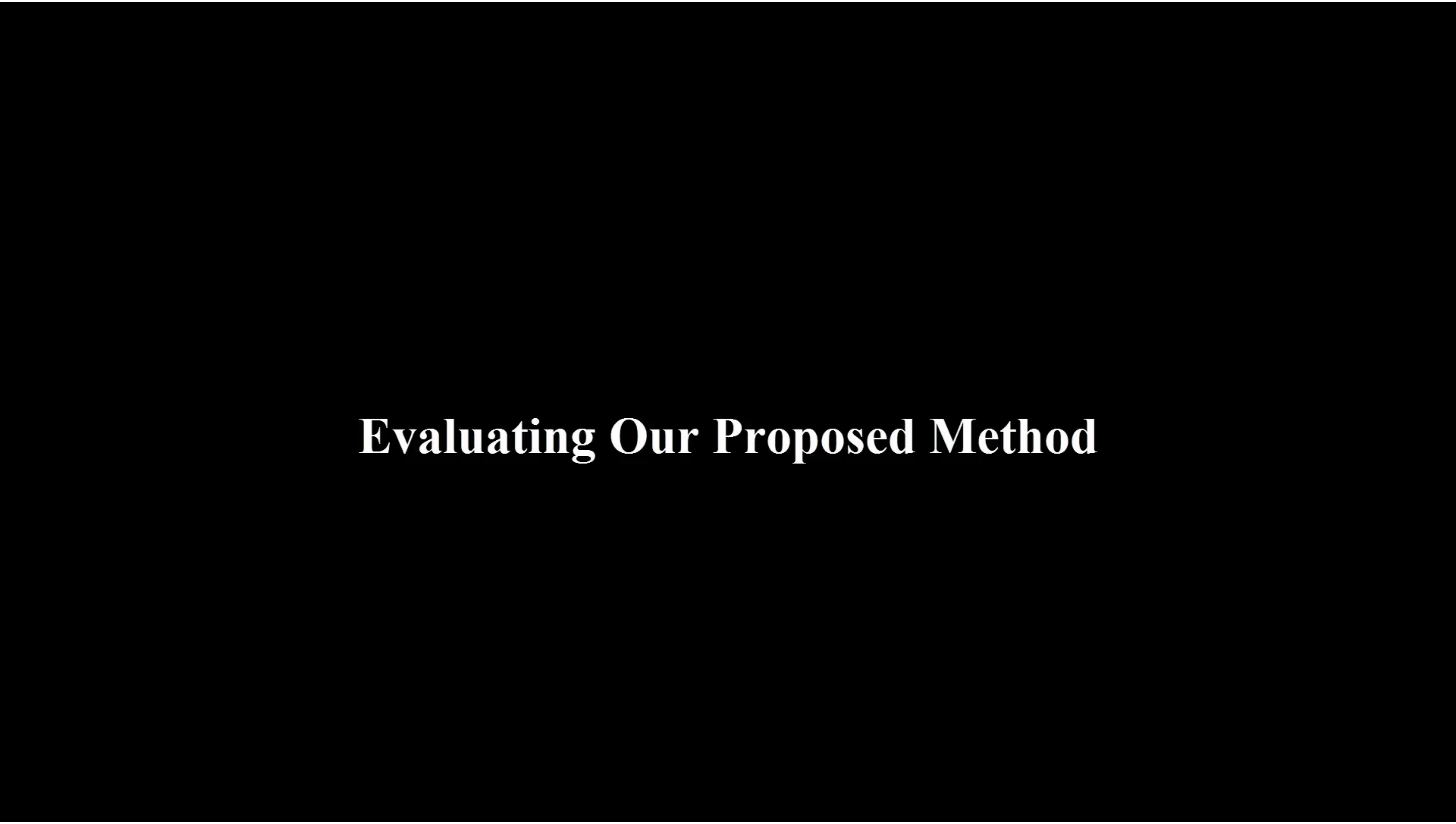
Human-Virtual Avatar Interaction : Social Interaction (2/5)



Human-Virtual Avatar Interaction : Social Interaction (3/5)

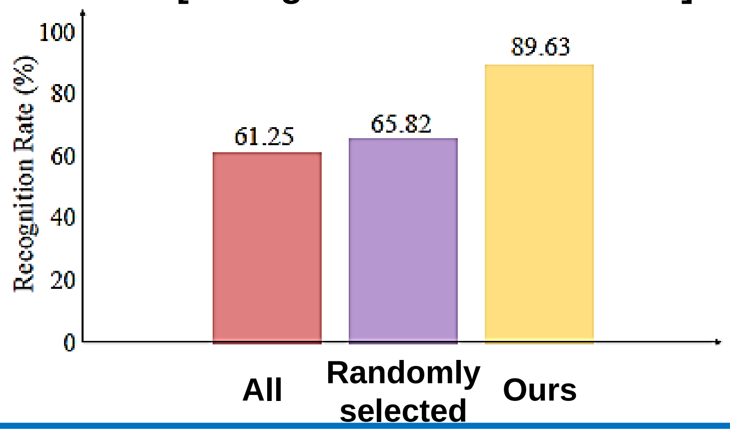


Human-Virtual Avatar Interaction : Social Interaction (4/5)

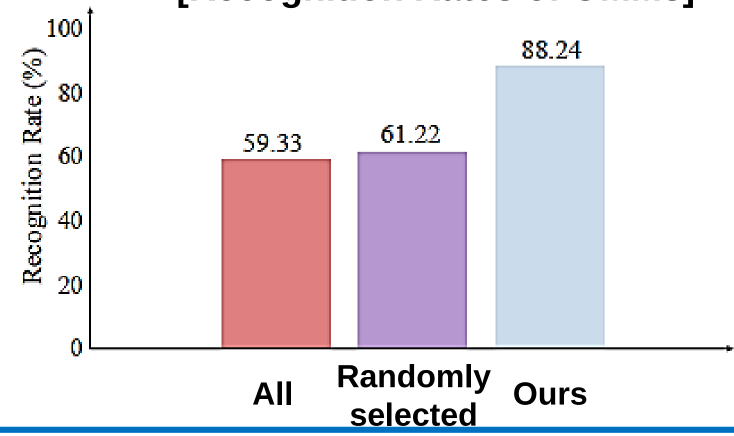


Human-Virtual Avatar Interaction : Social Interaction (5/5)

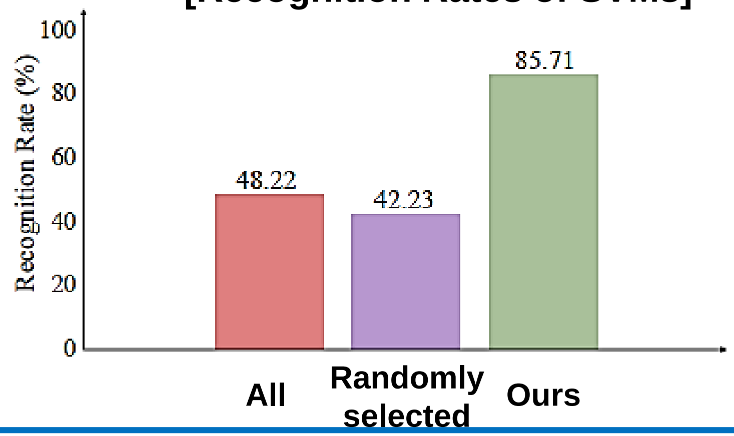
[Recognition Rates of HMMs]



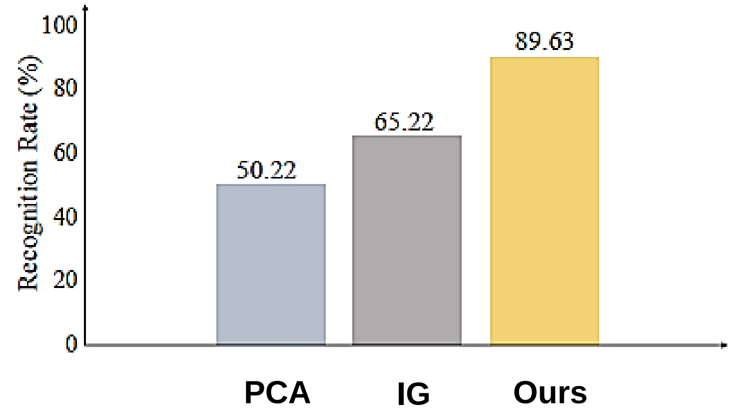
[Recognition Rates of GMMs]



[Recognition Rates of SVMs]



[Averaged Recognition Rates of HMMs, GMMs, SVMs]



Thank You for Your Attention!

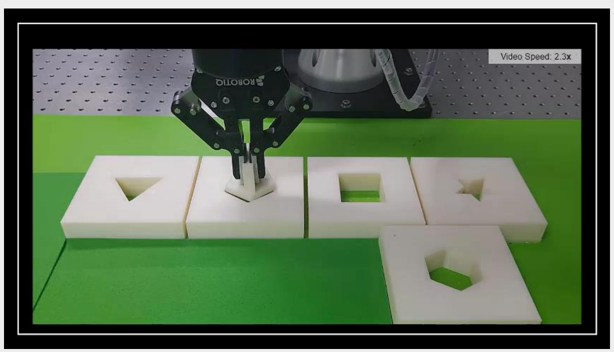
Ongoing Works

Collaborators

Motion Complexity & Deep Fitting

Sang Hyoung Lee
Korea Institute of Industrial Technology

Nam Jun Cho
Hanyang University



Deep Grasping

Young-Bin Park
Hanyang University

Byung Wan Kim
Hanyang University

Jong Soon Won
Hanyang University