Feeling the Force: Integrating Force and Pose for Fluent Discovery through Imitation Learning to Open Medicine Bottles

Mark Edmonds[1*] Feng Gao[1*] Xu Xie[1] Hangxin Liu[1] Siyuan Qi[1] Yixin Zhu[1] Brandon Rothrock[2] and Song-Chun Zhu[1]
* Equal Contributors

1. Center for Vision, Cognition, Learning, and Autonomy, UCLA
2. Jet Propulsion Laboratory, Caltech

Motivation

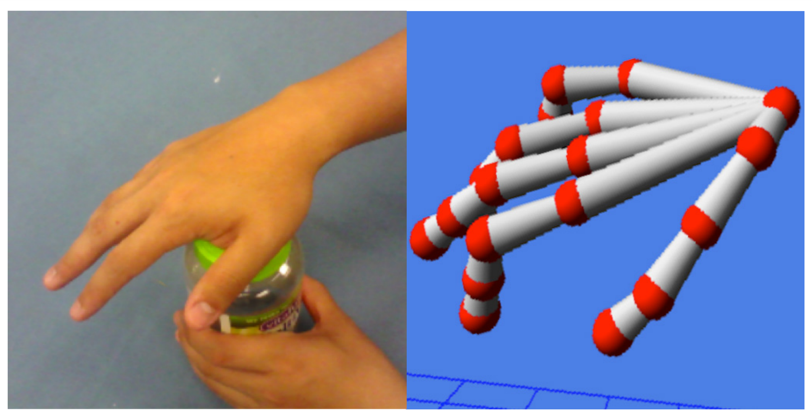


Fig. 1. Data used in traditional Hand-Object interaction study.

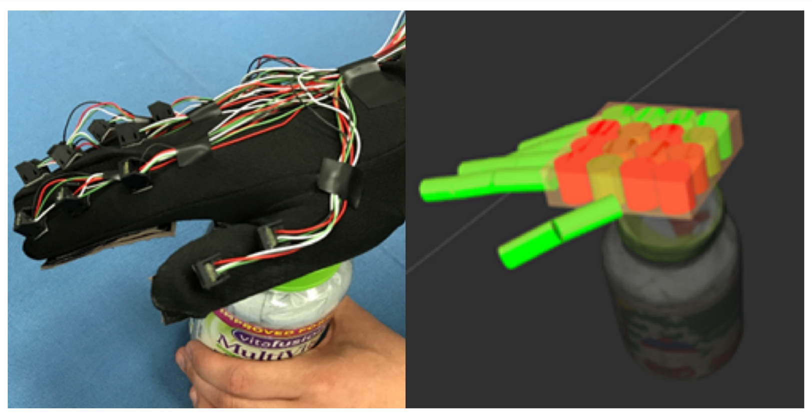


Fig. 2. Data used in our approach.

Consider the task of opening medicine bottles that have child-safety locking mechanisms. These bottles require the user to push or squeeze in various places to unlock the cap. By design, attempts to open these bottles using a standard procedure will result in failure. In this paper, we learn an action planner through both an And-Or graph (AOG) to represent the compositional nature of the task sequence and a bottom-up discriminative model from the observed poses and forces. We present a method for transferring this human-specific knowledge onto a robot platform and demonstrate that the robot can perform successful manipulations of unseen objects with similar task structure.

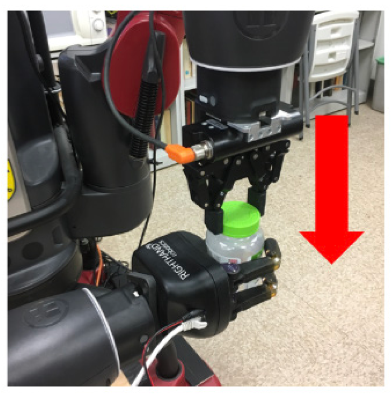


Fig. 3. Learned robot manipulation model using force

Data Collection

Fig. 4. Tactile glove used in data collection

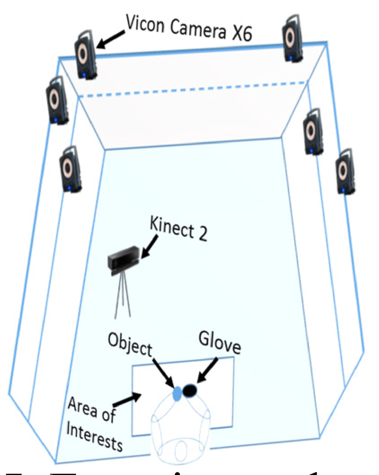


Fig. 5. Experimental setup

- To capture both pose and force in hand-object interactions, we utilize an open-source tactile glove. The tactile glove employs a network of 15 IMUs to measure the rotations between individual phalanxes. Hand pose is reconstructed using forward kinematics. The force is measured using Velostat, a piezoresistive material. Velostat sensors are placed proximal and distal link on each phalange and a 4x4 regions on palm. We utilize a Vicon motion capture system to obtain the relative poses between the wrist of hand and object parts. Six Vicon cameras are placed on top left and top right in front of the area of interests.
- Approximately 10 trials were collected for each grasping strategy.

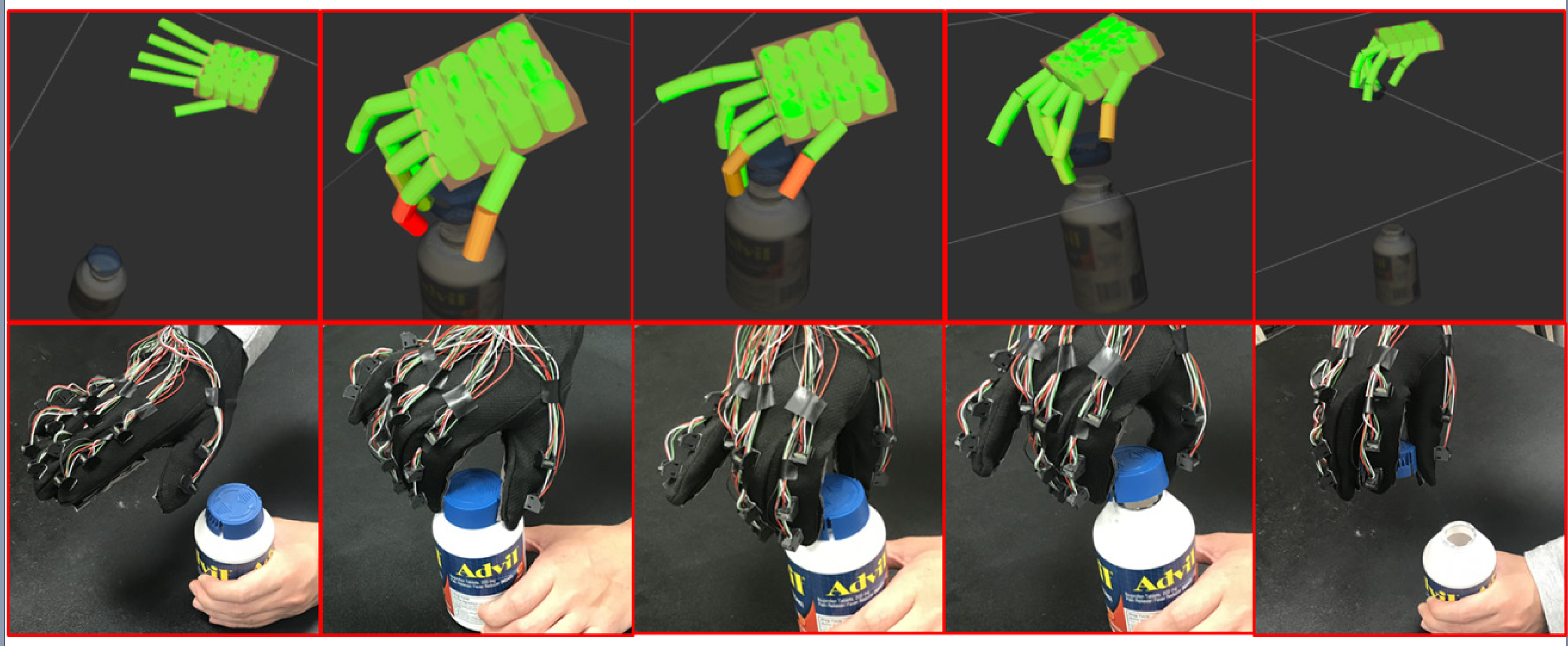
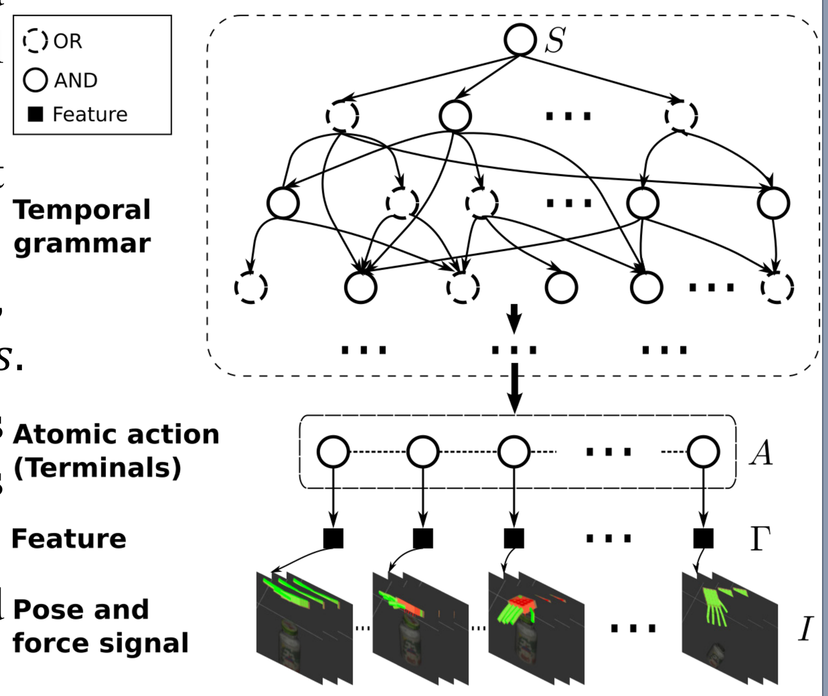


Fig. 6. Example of captured data

Representation

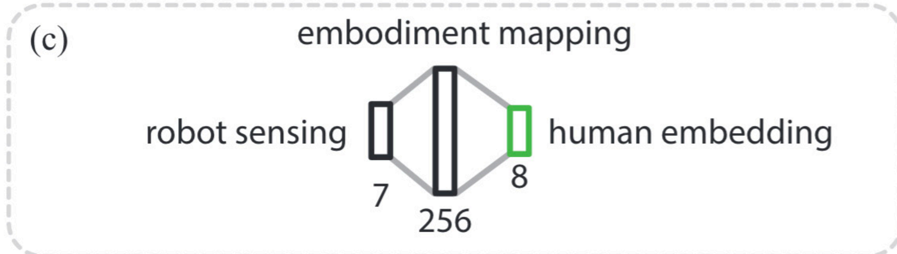
An AOG $G = (S, V, R, P, \sum)$ is a directed graph providing a hierarchical and compositional representation for entities.

- S is a start symbol that represents an event category (*e.g.* opening a bottle).
- V is a set of nodes $V = V^{AND} \cup V^{OR} \cup V^{T}$, $AND\ nodes, OR\ nodes, and\ Terminal\ nodes.$
- $R = \{r: \alpha \rightarrow \beta\}$ is a set of production rules that represent the top-down sampling process from a parent node α to its child nodes β.
- $P: p(r) = p(\beta|\alpha)$ is the probability associated with each production rule.
- $\sum$ is the language defined by the grammar.

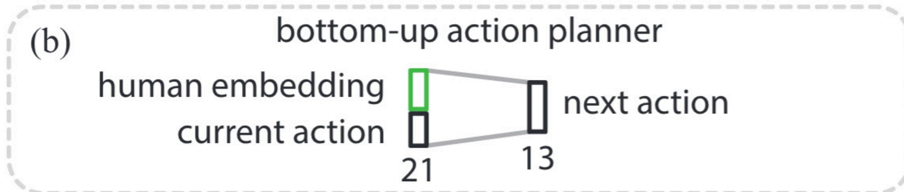


Learning

Embodiment Mapping: The embodiment mapping seeks a function, $s_h = \varphi(s_r)$ to map from robot sensing states to equivalent human states during demonstrations. We train this function using a neural network. The training data consists of successful robot actions sampled from the AOG and human demonstration data.



Bottom-up Term: The bottom up term is learned using a neural network. The input to the network is a low-dimensional embedding of the human demonstration and the current action. The output of the network is the probability of each possible next action.



Top-down Term: The top-down term is learned using an unstructured grammar induction method. The grammar induced represents the task structure as an And-Or graph.

$$p(G|X) \propto p(G)p(X|G) = \frac{1}{Z}e^{-\alpha||G||}\prod_{pg_i \in X} p(pg_i|G)$$

Planning

Problem Definition: The planning objective is to find the best next action a^*_{k+1} given the observed partial parse graph $pg_k(a_0, \dots, a_k)$. We select the next partial parse graph by adding each action to the parse graph and selecting the parse graph that minimizes the following energy term:

$$p(pg_{k+1}|pg_k, f_k) = \frac{1}{Z}\exp\{-\mathcal{E}(pg_{k+1}|pg_k, f_k)\}$$

where $\mathcal{E}(pg_{k+1}|pg_k, f_k) = -\log\left[p(pg_{k+1}|pg_k)\right] - \log\left[p(a_{k+1}|a_k, f_k)\right]$

Top-down Term: $p(pg_{k+1}|pg_k)$ plans the next action given the sequence of previous actions. It encodes a *long-term* relation between the previous action and the next action. The top-down term is parsed from the AOG using the Earley Parser.

Bottom-up Term: $p(a_{k+1}|a_k, f_k)$ plans the next action using both the current action label and the observed fluent. It encodes a *short-term* relation using the current pose and haptic feedback (fluent) and a single previous action.

Results

We use a dual-armed 7-DoF Baxter robot from Rethink Robotics mounted on a DataSpeed Mobility Base as our robot platform. While there may be multiple ways to open each bottle, not all methods are considered equivalent. We evaluate our method using three different configurations of the system: (1) using top-down planning only, (2) using bottom-up only planning, and (3) using the top-down and bottom-up planning.

TABLE I: Baseline 1, top-down only planning

Evaluation	bot. 1	bot. 2	bot. 3	bot. 4	bot. 5
Success	8.7%	5.6%	4.4%	8.7%	26.1%
Success (extra/wrong)	21.7%	5.6%	34.8%	47.8%	39.1%
Failure (action)	69.6%	77.7%	60.8%	34.8%	30.4%
Failure (execution)	0%	11.1%	0%	8.7%	4.4%

TABLE II: Baseline 2, bottom-up only planning

Evaluation	bot. 1	bot. 2	bot. 3	bot. 4	bot. 5
Success	4.4%	0%	4.4%	0%	4.4%
Success (extra/wrong)	13%	11.8%	30.4%	42.9%	17.4%
Failure (action)	82.6%	76.4%	65.2%	57.1%	78.2%
Failure (execution)	0%	11.8%	0%	0%	0%

TABLE III: Proposed, top-down and bottom-up planning

Evaluation	bot. 1	bot. 2	bot. 3	bot. 4	bot. 5
Success	8.7%	17.6%	17.4%	20%	60.9%
Success (extra/wrong)	52.2%	17.6%	65.2%	73.3%	17.4%
Failure (action)	39.1%	64.8%	13%	6.7%	21.7%
Failure (execution)	0%	0%	4.4%	0%	0%

Our results show a novel method of naturally capturing visually hidden states of a task and transferring them to the robot through human demonstrations using a tactile glove. The proposed method (Table III) integrates both the task structure and haptic feedback and vastly outperforms the methods that solely utilize task structure or haptic feedback.

The work reported herein is supported by DARPA XAI grant N66001-17-2-4029 and ONR MURI grant N00014-16-1-2007.