IJCAI 2017 Outline

End-to-end Situated Dynamics for Robot Task Learning from Human Demonstration

Overview

We model a dynamic system from low level motion equations to high level task understanding. An industrial robot learns a task from human demonstrations (such as how to fold clothes), using a spatial, temporal and causal And-Or graph representation. The robot abstracts the states of clothes into geometric descriptions and assigns a "value" to those abstract states. Then, when folding new clothes, the robot makes deductive plans to maximize the values, in contrast to the example-based inductive learning where the robot can only follow previous human examples.

Model

Definition 1. Environment: The world (or *environment*) is modeled by a generative grammar of objects, actions, and changes in conditions. We use the stochastic context-free And-Or graph (AOG), which explicitly models variations and compositions of spatial (S), temporal (T), and causal (C) concepts, called the STC-AOG.

Definition 2. State: A *state* is a configuration of the believed model of the world. In our case, a state is a parse-graph (pg) of the And-Or graph, representing a selection of parameters (Θ_{OR}) for each Or-node. The set of all parse-graphs is denoted Ω_{pg}.

Definition 3. Fluent: A *fluent* is a condition of a state that can change over time. Since the parameters (Θ_{AND}) of the And nodes of a parse-graph are conditions that may change over time, those parameters can be considered fluents. As shown in equation 1, a fluent (indexed by $i \in \mathbb{N}$) is a real-valued function on the state:

$$f_i : \Omega_{pg} \to \mathbb{R} \tag{1}$$

Definition 4. Fluent-vector: A *fluent-vector* F is a column-vector of fluents f. Equation 2 shows a fluent-vector of k fluents.

$$F = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_k \end{pmatrix} \tag{2}$$

Definition 5. Relevant fluent: A fluent-vector might contain irrelevant fluents for an action. The *relevant fluents* are a subset of fluents, described using an element-wise multiplication of a binary vector w by the fluent-vector.

$$w \circ F = \begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_k \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_k \end{pmatrix} = \mathrm{diag}(w)F \tag{3}$$

More generally, we relax the constraint that w must be a binary vector. The diagonal matrix is now an arbitrary matrix W, and the relevant fluent is denoted WF. W is most semantically meaningful when it is a diagonal matrix.

Definition 6. Action: An *action* a is a change in fluents given a precondition. The precondition of an action depends on relevant fluents W_a as well as a representative example pg_a. Let $\theta = (W_a, pg_a)$ denote these parameters. The precondition is a probability $P(X = pg \mid a;\ \theta_a)$ over a random variable $X \in \Omega_{pg}$. The energy of a pg is computed by comparing the difference between the relevant fluents of the observation Wpg to the relevant fluents of the action Wpg_a:

$$E(pg;\ W_a, pg_a) = ||W_a(pg - pg_a)|| \tag{4}$$

The probability density function is below (with a slack variable β, and partition function $Z(\beta)$):

$$P(X = pg \mid a;\ \theta_a) = \frac{1}{Z(\beta)} \exp(-\beta E(pg; \theta_a)) \tag{5}$$

Learning $\theta_a = (W_a, pg_a)$

Assume we're given supervised examples of a pair of fluent-changes $((F_i^{source}, F_i^{target}), (F_j^{source}, F_j^{target})) \in \mathcal{D}_a$ that were both caused by the same action a.

The parameter pg_a is modeled as the mean of all initial fluents for that action:

$$pg_a = \frac{1}{|\mathcal{D}_a|} \sum_{\mathcal{D}_a} F_i^{source} \tag{6}$$

Next, we'll solve for the other parameter W_a.

The change in relevant fluents must match for an ideal W:

$$W(F_i^{target} - F_i^{source}) = W(F_j^{target} - F_j^{source}) \tag{7}$$

$$\Rightarrow W(\triangle F_i) = W(\triangle F_j) \tag{8}$$

To select a couple good candidate W^α from a large dictionary of possible matrices, we employ the Minimax Entropy Principle. The statistic of the relevant fluents $W^\alpha \triangle F$ is $E_f[W^\alpha \triangle F]$, where $f(\triangle F)$ is estimated by the sample mean computed from training examples. Given a set of relevant fluents $S = \{W^{(\alpha)}, \alpha = 1, 2, ..., K\}$, a model $p(\triangle F)$ is constructed such that it reproduces the relevant fluent statistics as observed, following the maximum entropy principle (for all $\alpha = 1, 2, ..., K$):

$$E_p[W^{(\alpha)} \triangle F] = E_f[W^{(\alpha)} \triangle F] \tag{9}$$

Then using the minimum entropy principle, we select the relevent fluents S that best characterize $f(\triangle F)$ by minimizing $KL(f, p)$.

Alternative way to find relevant fluents (might be computationally easier?)

An approximate W minimizes the penalty:

$$\mathcal{P}_1 = \sum_{\mathcal{D}_a} ||W(\triangle F_i - \triangle F_j)|| \tag{10}$$

We enforce sparsity in W by introducing another penalty term:

$$\mathcal{P}_2 = ||W||_1 \tag{11}$$

And lastly, we enforce semantic meaning by introducing a penalty on how diagonal a matrix is:

$$\mathcal{P}_3 = DiagDist(W) \tag{12}$$

All together, we formulate the following optimization problem to obtain W_a:

$$W_a = \underset{W}{\operatorname{argmin}}\, \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 = \underset{W}{\operatorname{argmin}} \sum_{\mathcal{D}_a} ||W(\triangle F_i - \triangle F_j)|| + ||W||_1 + DiagDist(W) \tag{13}$$

Inference

A candidate action $\hat{a}$ is identified by sampling from the conditional probability:

$$\hat{a} \sim P(a \mid pg) \propto P(pg \mid a)P(a) \tag{14}$$

The likelihood $P(pg \mid a)$ is already defined by equation 5, and the prior $P(a)$ is solved through a kernel density estimation (KDE) over the training dataset.

Definition 7. Value: A *value* is a real-valued function defined on the space of fluent-vectors Ω_F to indicate a preference of being at a state:

$$V : \Omega_F \to \mathbb{R} \tag{15}$$

The Lagrangian function L is defined by the difference in kinetic K and potential V energy of the robot state q situated in the world $F \in \Omega_{pg}$:

$$K(q, \dot{q}) = \frac{1}{2}\dot{q}^\intercal M(q)\dot{q} \qquad V(q, F) = V(F) \qquad L = K - V \tag{16}$$

The constraints of motion given by the Euler-Lagrange equation yields:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = \frac{d}{dt}\frac{\partial K(q,\dot{q})}{\partial \dot{q}} - \frac{\partial K(q,\dot{q})}{\partial q} + \frac{\partial V}{\partial q} = 0 \tag{17}$$

The fluent-vector of the robot's configuration state is represented by $U \in \mathbb{R}^r$. The connection between values, fluents, and actions are represented by the following factoring of the potential function:

$$\frac{\partial V}{\partial q} = \frac{\partial V}{\partial f_i}\frac{\partial f_i}{\partial u_j}\frac{\partial u_j}{\partial q} \tag{18}$$

Appendix

U (frequency) + C (causality) + V (value)

not MLE

Notes from Prof Zhu:

```
human utility and values in the fluent space
(Humans prefer certain fluent states
under certain conditions and tasks,
how to define a consistent utility function
on high dimensional fluent space, is unknown.
It depends strong on boundary conditions).

In Chinese, they say:  "your hip decides your brain"
That is:   where you sit (i.e. conditions, situation, and task)
decides what you think and do.

How do we formulate this in the utility?
What type of field is the utility function, is it conservative?
and the gradients of the utility is a vector field,
but what type of vector field is it?
```

Physically feasible plan pg to reach the goal δF.

$$\{(pg, I) : \phi(pg) = \delta F\}$$

Thought: Grammar to landscape.

A value vector field that accepts languages.

a, b, a, b is a vortex