Action-Based Models for Belief Space Planning  
Li Yang Ku, Shiraj Sen, Erik G. Learned-Miller, and Roderic A. Grupen  
University of Massachusetts Amherst

**Goal** - Address the dual problems of modeling and reasoning by employing an action-based model grounded in the robot’s own actions and perceptions.

**Aspect Transition Graph (ATG)**
- An ATG is a directed graph $G = (\mathcal{X}, \mathcal{U})$, composed of a set of aspect nodes $\mathcal{X}$ connected by a set of action edges $\mathcal{U}$.
- Each aspect node $x \in \mathcal{X}$ represents the properties of an object that are measurable given a set of sensor parameters.

$$p(O_T \notin S_{T-1} | z_{1:T}, a_{1:T}, M) = \sum_{o_i \in S_{T-1}} p(O_T = o_i | z_{1:T}, a_{1:T}, M) = \sum_{o_i \in S_{T-1}} \sum_{s_i \in S_{T-1}} p(x_i | z_{1:T}, a_{1:T}).$$

**Learning Models**
- An ATG is added to the robot memory $M$ only if the presented object has not been presented to the robot in the past.
- Given a sequence of observations $z_{1:t}$ and actions $a_{1:t}$ during trial $T$, the probability that the presented object $O_T$ during trial $T$ is novel can be calculated;

$$p(O_T \notin S_{T-1} | z_{1:T}, a_{1:T}, M) = \sum_{o_i \in S_{T-1}} p(O_T = o_i | z_{1:T}, a_{1:T}, M) = \sum_{o_i \in S_{T-1}} \sum_{s_i \in S_{T-1}} p(x_i | z_{1:T}, a_{1:T}).$$

**Minimum Entropy Planner**
- The action $a_t$ that minimizes the expected entropy over the random variable $O_T$ representing the object identity is selected.
- Future observation is estimated through object models created from past observation.

$$\arg \min_{a_t} E(H(O_T | z_{t+1}, a_t, z_{1:t}, a_{1:t-1})) = \arg \min_{a_t} \sum_{z_{t+1}} H(O_T | z_{t+1}, a_t, z_{1:T}, a_{1:t-1}) p(z_{t+1} | a_t, z_{1:T}, a_{1:t-1}).$$

**Simultaneous Object Modeling and Recognition**
- The robot builds a set of object models through interacting with random objects one at a time.
- Evaluated based on whether the robot can identify novel objects or recognize which object it corresponds to in memory.

**Experiment Results**
- Each test involves 100 trials and starts with an empty robot memory $M$.
- The efficiency of the planner is tested against a random policy.

<table>
<thead>
<tr>
<th>Test</th>
<th>Correct Identification</th>
<th>Correct Recognition</th>
<th>Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100/100</td>
<td>100/100</td>
<td>100%</td>
</tr>
<tr>
<td>2</td>
<td>99/100</td>
<td>100/100</td>
<td>99%</td>
</tr>
<tr>
<td>3</td>
<td>99/100</td>
<td>100/100</td>
<td>99%</td>
</tr>
<tr>
<td>4</td>
<td>99/100</td>
<td>100/100</td>
<td>99%</td>
</tr>
<tr>
<td>Mean</td>
<td>99/100</td>
<td>100/100</td>
<td>99%</td>
</tr>
</tbody>
</table>