Deep Anticipatory Networks for Active Perception

Yash Satsangi

August 26, 2019

joint work with Sungsu Lim, Shimon Whiteson, Frans Oliehoek, Martha White

Multi-camera systems

- Scarce Resources
  - Manpower
  - Computational resources.
  - Bandwidth.
  - Energy.

Sensor selection problem

At each time step t:

- Given: probability distribution over state of the world.

$$b(s) = \Pr(s).$$

- Select: 1 out of n available sensors.

$$\mathcal{A} = \{1 \dots n\}; \text{select } a \in A.$$

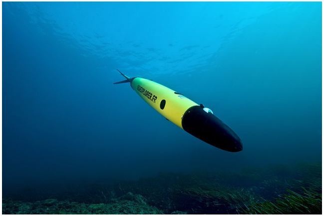
- Goal: to minimize future uncertainty about the state of the world.

$$b^a_{\mathbf{z}}(s) = \Pr(s|\mathbf{z}, a),$$

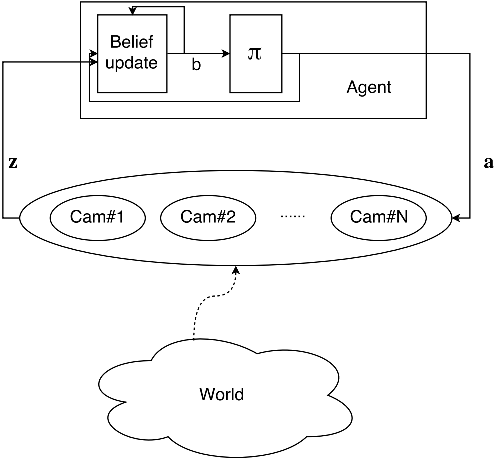
where $\mathbf{z} = \langle z_1, z_2, \dots, z_n \rangle$ is the observation vector.

Active perception

- Active perception is the ability of an agent to reason about its own limitations to come up with control strategies for information gathering.

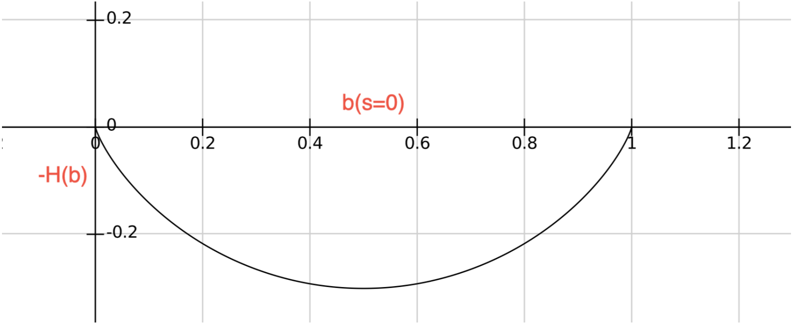


Active perception POMDP

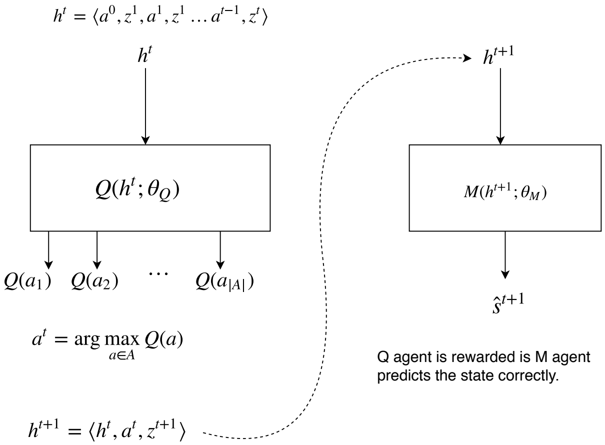


$$\rho(b) = -H(b) = \sum_{s \in S} b(s) log(b(s))$$

Entropy



- $\rho(b) = \sum_{s \in S} b(s) log(b(s))$
- If $\rho(b)$ is to be calculated then we need a definite representation of b.



Q agent is rewarded is M agent predicts the state correctly.

- Train Q and M simultaneously.

Analysis

- Approximate $\rho(b) = -H(b)$ with $\rho'(b) = \max_{p \in P}[\sum_s [b(s)R(s,p)]]$.

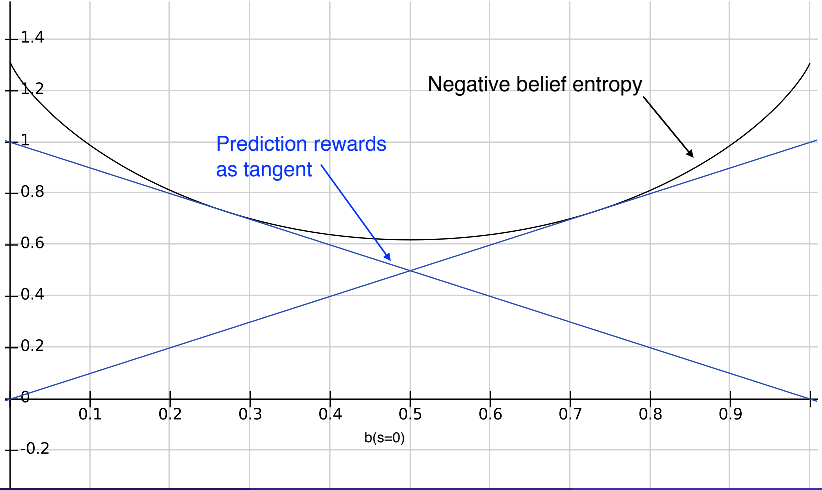
Theorem

Let $R(s,p) = 1$ *if* $s = p$ *and* 0 *otherwise, then*

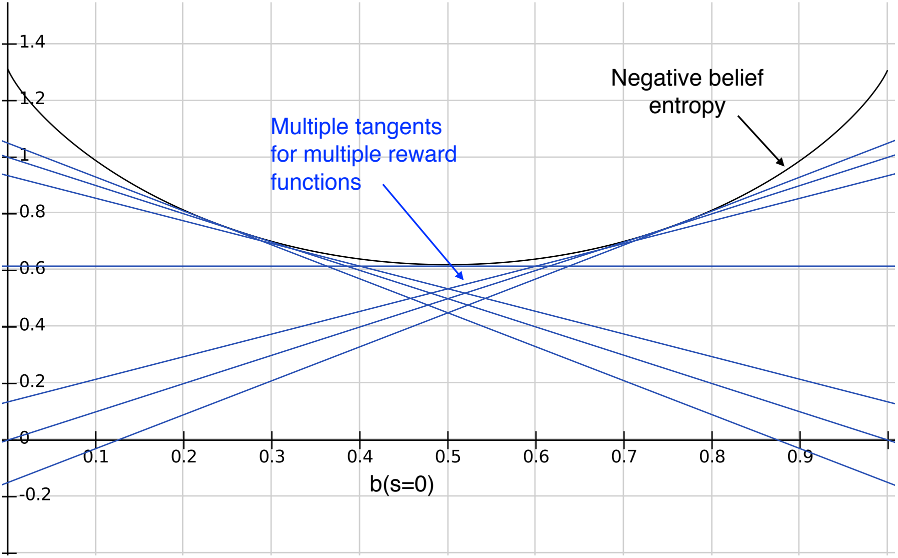
$$\rho(b) - \rho'(b) \leq 1.31$$

Analysis

$$\rho'(b) = \max_{p \in P}[b(s_0)R(s_0, p) + (1 - b(s_0))R(s_1, p)]$$



Analysis



1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
b(s=0)
Multiple tangents for multiple reward functions
Negative belief entropy

Analysis

Theorem

Let $R(s,p) = r'$ *if* $s = p$ *and* r'' *otherwise, then*

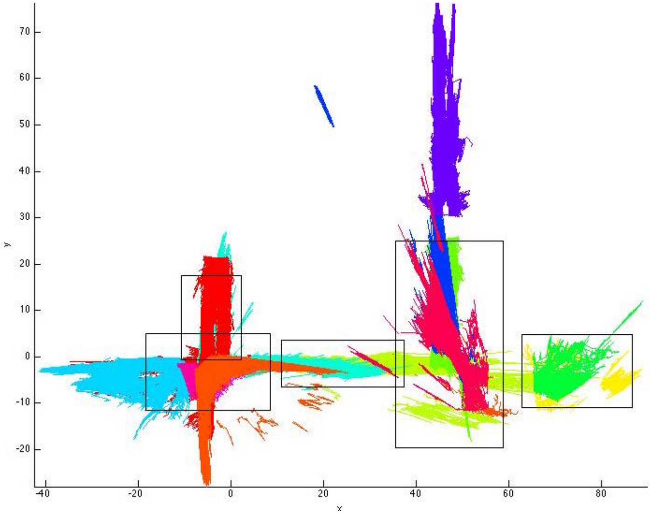
$$\rho(b) - \rho'(b) \leq \max\{\epsilon_1, \epsilon_2\}$$

where

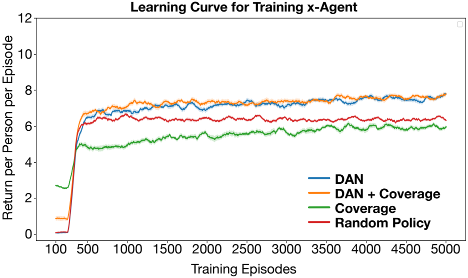
$$\epsilon_1 = \log(e^{r'} + (|S|-1)e^{r''}) - r',$$

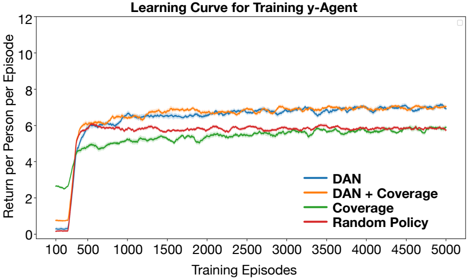
$$\epsilon_2 = \log(\frac{1}{|S|}) - (\frac{1}{|S|})(r' + (|S|-1)r'') + \log(e^{r'} + (|S|-1)e^{r''}).$$

Experiments - sensor selection

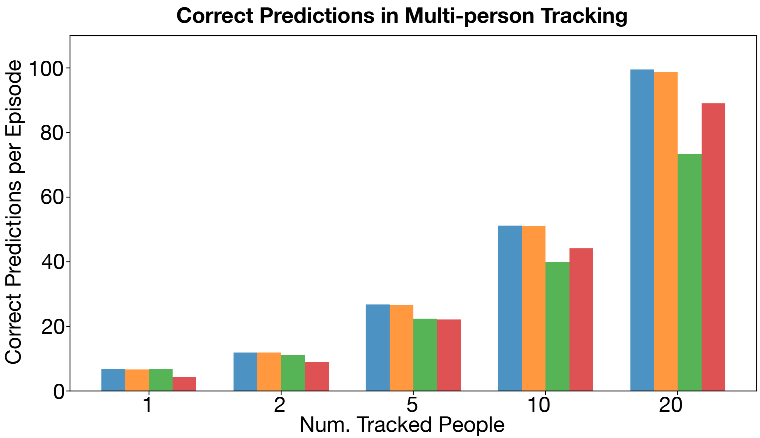


Experiments

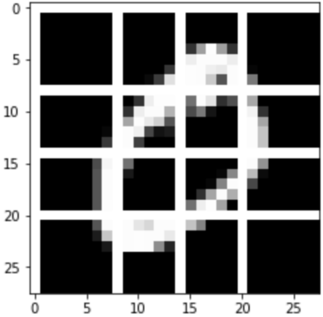




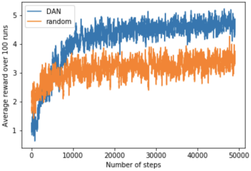
Experiments

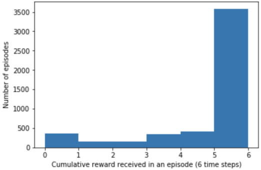


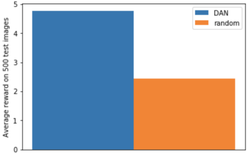
Discrete attention

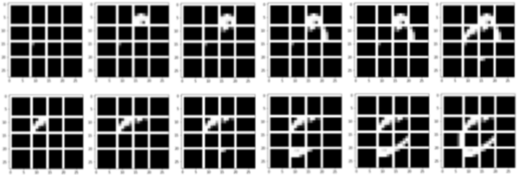


Experiments









Summary

- DANs - cooperative task between a Q network and an M network.
- Rewarding agent for correct predicting maximizes information.
- Application - sensor selection, discrete attention, decision trees.