
Infomax Control Approaches for Social Robotics

Ian R. Fasel

Department of Computer Science
University of Arizona
ianfasel@cs.arizona.edu

Paul Ruvolo^{1,2}, Tingfan Wu^{1,2}, Javier Movellan^{2,3}

¹Department of Computer Science, ²CalIT2

³Institute for Neural Computation

University of California, San Diego

{ruvolo, ting, movellan}@mplab.ucsd.edu

Robotics problems often involve updating and acting on the basis of beliefs about the world state. In social robots, it is important to maintain beliefs not just about objects in the room and the locations of the walls, but also beliefs about peoples' locations, their head directions, who they are interacting with, and their emotional states. Each of these aspects of the social world may change rapidly and whose interactions may be much more complex than those of inanimate objects. Social robots, often designed with human-like affordances, are typically equipped with a set of orientation specific sensors such as cameras located in the eyes and microphones in the ears. Provided these sensors and the ability to move the neck to change the fields-of-view of these sensors, what strategy should a robot use for changing its head-gaze direction over time? Unlike special-purpose robots designed for a single task, the high-level goal of a social robot may be composed of many different subtasks that take varying degrees of priority depending upon the actions of its interactive human partners. Consider the case of a robot acting as a mobile tour guide in a museum, the robot may have to switch between tasks such as tracking specific people it is currently engaged with or looking for new people who might be interested in a tour. Some possible head-movement strategies for balancing these objectives might be to track any face that comes into view until it leaves some specified region, look for "interesting" events, fix the robot's gaze at the doorway of the museum, or shift gaze randomly.

In the absence of a specific prioritization over the relevant subtasks a reasonable strategy for a robot is to act in such a way that the uncertainty about the social world state is always kept as small as possible. In the current work, we formalize this idea by proposing the Information Maximization (InfoMax) POMDP, in which the reward function at each time step is the negative entropy of the probability distribution over the state of the social world (belief state). Because neg-entropy is an intrinsic reward derived from the belief state, it is not specific to any particular task and does not require a human give it one. InfoMax may therefore be a good "default" strategy in the absence of another, external task demand.

We explored this idea by implementing an InfoMax control policy on a robotic head during a three week "Neuromorphic Engineering" workshop in Telluride Colorado. The head, which is designed to look like Albert Einstein, complete with facial expressions, has directional cameras in its eyes which are fed to computer vision algorithms that locate human faces, detect their head pose, and recognize their facial expressions. The robot was also equipped with a microphone array which could rapidly localize sources of sound, and a wide-field silicon retina which can detect a variety of subtle motion cues. The robot was placed in a large room in which people could enter and leave randomly, and who sometimes would approach the robot and try to interact with it. There was no high-level "goal" for the robot – that is, Einstein didn't have a job. His intrinsic desire was to maintain as much certainty about the state of the social environment as possible.

We experimented with a variety of methods for combining the sensors to form probabilistic beliefs about the locations and facial expressions of people, and ultimately settled on a simple recurrent neural network-based belief model. We then used a natural actor-critic reinforcement algorithm to learn a policy on a simulated social world, using a convolutional radial-basis-function network as the controller to capture shift-invariances. We found that depending on the reliability of the sensors and the dynamics governing the people in the simulated social world, a number of different behaviors emerge from the reinforcement learning algorithm. These include person tracking, oc-

asionally shifting gaze between different people, and turning to look towards suddenly bursts of sound. Interestingly, the convolutional network typically learned center-surround-like weight vectors that operated on probabilistic belief maps.