Gaze Following as Goal Inference: A Bayesian Model

Abram L. Friesen and **Rajesh P. N. Rao**
{afriesen, rao}@cs.washington.edu
Department of Computer Science and Engineering
University of Washington, Box 352350
Seattle, WA 98195 USA

Abstract

The ability to follow the gaze of another human plays a critical role in cognitive development. Infants as young as 12 months old have been shown to follow the gaze of adults. Recent experimental results indicate that gaze following is not merely an imitation of head movement. We propose that children learn a probabilistic model of the consequences of their movements, and later use this learned model of self as a surrogate for another human. We introduce a Bayesian model where gaze following occurs as a consequence of goal inference in a learned probabilistic graphical model. Bayesian inference over this learned model provides both an estimate of another's fixation location and the appropriate action to follow their gaze. The model can be regarded as a probabilistic instantiation of Meltzoff's "Like Me" hypothesis. We present simulation results based on a nonparametric Gaussian process implementation of the model, and compare the model's performance to infant gaze following results.

Keywords: cognitive development; machine learning; artificial intelligence; goal inference; Bayesian modeling; gaze following.

Introduction

Gaze following plays an important role in cognitive development. Following the gaze of an adult, for example, allows a child to jointly attend to an object, learn its name and other properties, as well as learn useful actions to perform on the object through imitation. It has been shown that children as young as 12 months old can follow the gaze of an adult and engage in joint attention (Brooks & Meltzoff, 2002).

Recent results have shown that gaze following is not merely an imitation of head movement. For example, 14- and 18-month olds do not follow the gaze of an adult who is wearing a blindfold, although they follow gaze if the adult wears the same band as a headband. This suggests that these children do not follow gaze because they are aware of the consequences of wearing a blindfold (i.e., occlusion) and unlike 12-month olds, make the inference that the adult is not looking at an object. This observation is closely related to Meltzoff's "Like me" hypothesis (Meltzoff, 2005) which states that self-experience plays an important role in making inferences about the internal states of others. In particular, in the case of the blindfold experiment, self-experience with own eye closure and occluders may influence gaze following behavior. To test this hypothesis, Meltzoff and Brooks provided one group of 12-month olds with self-experience with an opaque blindfold while two other groups either had no self-experience or had self-experience with a windowed blindfold. On seeing an adult with a blindfold turn towards an object, most of the children who had had self-experience with blindfolds did not turn to the object while the other two groups did (Meltzoff & Brooks, 2008). These results suggest that (a) gaze following involves an inference of the underlying intention or goal of the head movement, and (b) self-experience plays a major role in learning the consequences of intentions and related actions.

In this paper, we propose a new model for gaze following and joint attention that can be viewed as a probabilistic instantiation of the "Like me" hypothesis. The model itself is general and can be applied to modeling other forms of goal-based imitation, but we focus here on gaze following. In the following section, we derive our framework for gaze following based on probabilistic graphical models. We describe how a child could learn a probabilistic model of the consequences of their own head movements, and later use this learned model to interpret the actions of another person. Bayesian inference over the learned graphical model provides both an estimate of another's fixation location and the appropriate action to move one's own gaze for joint attention. For the simulations, a model based on Gaussian process regression was used to learn the mapping between goals, actions, and their sensory consequences. We present preliminary results comparing the model to infant gaze following results and discuss the applicability of the proposed framework for understanding other forms of goal-based imitation and sensorimotor planning.

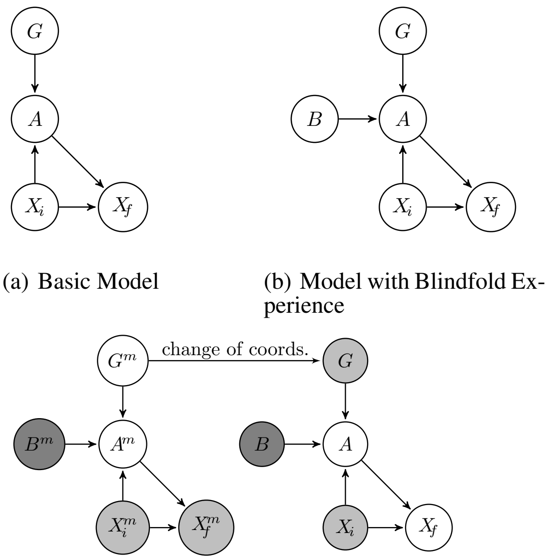
A Bayesian Model for Gaze-Following

In the following section, we develop and explain our model for gaze-following as goal inference. We begin with a simplified explanation of our graphical model, and then give an overview of the computational components.

Hypothesis

Our hypothesis is the following: humans learn a goal-directed mechanism for planning gaze movements. A goal location, provided by either internal or external stimuli, combined with the current state, determines an action. This action, again in conjunction with the current state, determines the final state. We represent this mechanism with the graphical model shown in Figure 1(a) where G is the goal, A is the action, X_i is the current state, and X_f is the final state. In the context of gaze following, the goal is a desired fixation location, the action is a vector of motor commands, and the state represents head position and orientation.

With our proposed model, an artificial agent can both plan future movements given desired fixation points and determine fixation points given observed head poses. Both of these correspond to performing inference over the graphical model.



(a) Basic Model

(b) Model with Blindfold Experience

(c) Gaze following model

Figure 1: Graphical models for controlling gaze: (a) contains the basic model which relates the different variables, (b) demonstrates the influence of blindfold experience on the model, and (c) shows the combined graphical models for following the gaze of a mentor. Shaded variables demonstrate observed variables. The darker shading indicates that B is an observed discrete variable, while the rest of the nodes are continuous.

Furthermore, through the "Like me" hypothesis, a human can apply their model to another person in order to infer where they are looking and then translate that location back to their own model in order to fixate on the same location, as is necessary for gaze following (Figure 1(c)).

Computational Model

We begin by showing how an agent can learn the mapping between goals, states, and actions in order to plan how to fixate on a specific goal location. The agent first learns a *transition* model, f, (e.g., through exploration or "body babbling") which translates an initial position, x_i, and an action, a, to a final head position, x_f,

$$x_f = f(x_i, a) + \nu, \quad \nu \sim \mathcal{N}(0, \Sigma_\nu), \tag{1}$$

where $\mathcal{N}(\mu, \Sigma)$ signifies the normal distribution with mean μ and covariance matrix Σ. The learned transition model can in turn be used to learn a *policy* function, π, which maps the initial head position and a goal location g to an action

$$a = \pi(x_i, g) + \upsilon, \quad \upsilon \sim \mathcal{N}(0, \Sigma_\upsilon). \tag{2}$$

This function essentially determines the rotation required to turn the head from its current position to face the goal.

We use two separate Gaussian processes (GPs) (Rasmussen & Williams, 2006), GP_π and GP_f, to learn these nonlinear functions. GPs are commonly used in machine learning to infer a latent function $h(\cdot)$ from noisy observations $y_i = h(\mathbf{x}_i) + \nu,\ \nu \sim \mathcal{N}(0, \sigma_\nu^2)$ because they are both flexible and robust. First, they are nonparametric so they do not limit the set of functions that h can be chosen from. Second, they estimate a probability distribution over functions instead of choosing a single most-likely candidate function, allowing all plausible functions to be incorporated into the estimation process and reducing model bias. For these reasons, GPs are effective with small numbers of training samples, increasing their biological plausibility. We now show how we use these GPs for planning, goal inference, and gaze-following.

Goal-directed Planning GPs are trained via supervised learning: given a training dataset with noisy labels, the GP learns the functional mapping between the input data and the labels (output). The training data itself could be obtained, for instance, through a reinforcement-based paradigm that combines exploration of the goal-action-state space for training the transition GP with selection of data from successful trials for training the policy GP (see (Verma & Rao, 2006) for related ideas). After training, it is simple for the agent to fixate on a goal location. We assume that the agent knows its current state x_i and the goal g. Given these, the agent uses its learned distribution over functions, GP_π, to compute a probability distribution over actions, $p(a) = \mathcal{N}(\mu_a, \Sigma_a)$. This distribution can then be passed through GP_f to estimate the probability of the resulting state, $p(x_f) \approx \mathcal{N}(\mu_{x_f}, \Sigma_{x_f})$.[1] Thus, our model provides us both with the final position and an estimate of the uncertainty in the prediction.

Goal Inference Through the use of our computational model, it is also possible to infer the goal of a head movement given observations of the starting and ending head poses, x_i and x_f, respectively. To accomplish this, the agent must be able to recover the inputs to each GP given the outputs. Fortunately, results from (Rasmussen & Ghahramani, 2003) allow us to estimate a distribution over the inputs given the outputs. As such, we can infer a distribution over actions given x_i and x_f and then use this to estimate a distribution over goals.

Gaze Following Goal-directed gaze following is accomplished through our use of the "Like me" hypothesis. The agent learns a model of itself and assumes that a mentor uses this same model. The agent observes the starting and ending states (head poses) of a mentor and then infers the goal location of the mentor, g^m, by inferring what it would be looking at if it were in the mentor's position. After inferring the mentor's goal, the agent transforms that goal into its own coordinate frame and then infers how to fixate on that goal. For this paper, we assume the agent has acquired the ability to transform between coordinate frames through prior experience.

Blindfold Experiments

To demonstrate the robustness and plausibility of this model, we recreate an experiment from (Meltzoff & Brooks, 2008) and test our model on it. We incorporate a blindfold variable,

[1] This is an approximation because, in general, a Gaussian passed through a nonlinear function does not remain Gaussian.

$B \in \{0,1\}$, (Figure 1(b)) and allow the agent to learn the effects of being blindfolded. Our model learns a new Gaussian process ($GP_{\pi}^{b=1}$, to use in place of GP_{π}) for when the agent is blindfolded. When the blindfold is in place regardless of the goal chosen, no action leads to that goal location being fixated. Goals in this case are not causally linked to states (head poses) or actions. The agent can learn this and then apply this knowledge to a mentor agent to infer that the mentor's goal is random in relation to the observed head movement, if it is blindfolded. However, if this alternate Gaussian process is not learned, the agent does not know anything about the blindfold and follows the mentor's head movement even if the mentor is blindfolded.

Technical Details

Gaussian Processes

We briefly introduce the notation and theory of Gaussian processes (GPs). More detail can be found in (Rasmussen & Williams, 2006). Formally, a Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution. The random variables represent the value of the function $h(\mathbf{x})$ at location $\mathbf{x}$. A GP is fully specified by its mean function $m(\cdot)$ and covariance function $k(\cdot,\cdot)$.We write the Gaussian process as $h(\mathbf{x}) \sim GP_h(m_h(\mathbf{x}), k_h(\mathbf{x},\mathbf{x}'))$, to denote the fact that the GP learns a probability distribution over functions. GPs place a prior $p(h)$ directly on the space of functions. We use a prior mean function $m_h = 0$ and use the squared exponential (SE) kernel as our prior covariance function $k_h(\mathbf{x}_p,\mathbf{x}_q) = \alpha^2 \exp\left(-\frac{1}{2}(\mathbf{x}_p - \mathbf{x}_q)^\top \Lambda^{-1}(\mathbf{x}_p - \mathbf{x}_q)\right)$, where $\mathbf{x}_p, \mathbf{x}_q \in \mathbb{R}^D$, $\Lambda = \mathrm{diag}([\ell_1^2, ..., \ell_D^2])$ is a diagonal matrix of squared characteristic length-scales, and α^2 is the variance of the latent function h.[2] Thus, the posterior predictive distribution of the function value $h_* = h(\mathbf{x}_*)$, for an arbitrary test input $\mathbf{x}_*$, is Gaussian with mean and variance

$$m_h(\mathbf{x}_*) = \mathbb{E}_h[h_*] = \mathbf{k}_*^\top(\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1}\mathbf{y} = \mathbf{k}_*^\top \beta, \tag{3}$$

$$\sigma_h^2(\mathbf{x}_*) = \mathrm{var}_h[h_*] = k_{**} - \mathbf{k}_*^\top(\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1}\mathbf{k}_*, \tag{4}$$

respectively, with $\mathbf{k}_* = k(\mathbf{X},\mathbf{x}_*)$, $k_{**} = k(\mathbf{x}_*,\mathbf{x}_*)$, $\beta = (\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1}\mathbf{y}$, $\mathbf{K}$ is the kernel matrix with $K_{ij} = k(\mathbf{x}_i,\mathbf{x}_j)$, $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_n]$ are the training inputs, and $\mathbf{y} = [y_1, ..., y_n]$ are the training targets.

It is also possible to predict with GPs when the test input $\mathbf{x}_* \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ is uncertain. This corresponds to seeking

$$p(h(\mathbf{x}_*)|\mu,\mathbf{\Sigma}) = \int p(h(\mathbf{x}_*)|\mathbf{x}_*)p(\mathbf{x}_*|\mu,\mathbf{\Sigma})\mathrm{d}\mathbf{x}_*. \tag{5}$$

For the SE kernel, we can compute the mean μ_* and variance σ_*^2 of equation 5 in closed form, following (Deisenroth, Huber, & Hanebeck, 2009; Quiñonero-Candela, Girard, Larsen, & Rasmussen, 2003).

$$\mu_* = \mathbb{E}_{\mathbf{x}_*}[\mathbb{E}_h[h(\mathbf{x}_*)|\mathbf{x}_*]|\mu,\mathbf{\Sigma}] = \mathbb{E}_{\mathbf{x}_*}[m_h(\mathbf{x}_*)|\mu,\mathbf{\Sigma}] \tag{6}$$

$$\sigma_*^2 = \mathbb{E}_{\mathbf{x}_*}[m_h(\mathbf{x}_*)^2|\mu,\mathbf{\Sigma}] + \mathbb{E}_{\mathbf{x}_*}[\sigma_h^2(\mathbf{x}_*)|\mu,\mathbf{\Sigma}] - \mu_*^2 \tag{7}$$

Note that this is a Gaussian approximation to the true distribution of $p(h(\mathbf{x}_*)|\mu,\mathbf{\Sigma}) \approx \mathcal{N}(\mu_*,\sigma_*^2)$ where the first two moments (mean and variance) are matched exactly. This computation can be extended to the multivariate case where $h : \mathbb{R}^D \to \mathbb{R}^E$. The only difference is that the target dimensions co-vary and the corresponding predictive covariance matrix is no longer diagonal (as it is when the input is deterministic in the multivariate case) (Deisenroth et al., 2009).

Inference in Graphical Models

In this paper, we use Bayesian networks, a type of probabilistic graphical model, to illustrate the relationships and conditional independencies between variables.

Forward Inference Using the above equations, we can compute $p(X_f|X_i = x_i, G = g)$ for the graphical model shown in Figure 1 as follows. Given a current state, x_i, and a goal, g, we first compute $p(A|x_i,g) = p(\pi(x_i,g) + \upsilon)$ as $p(A|x_i,g) = \mathcal{N}(m_\pi([x_i\ g]^\top), k_\pi([x_i\ g]^\top))$ from GP_π using (3) and (4) where we use $[x_i\ g]^\top$ as $\mathbf{x}_*$ and π as h. Alternatively, when gaze following we have a distribution over goals instead of a deterministic value. For this case, we compute $p(A|x_i,g) = \mathcal{N}(a|\mu_a,\sigma_a)$ using equations (6) and (7). This quantity can then be again substituted into (6) and (7),[3] where $[a\ x_i]^\top$ is now $\mathbf{x}_*$ and f is h, in order to compute $p(X_f|x_i,g) = p(f(x_i,a) + \nu|\mu_\mathbf{a},\mathbf{\Sigma_a}) \approx \mathcal{N}(\mu_{\mathbf{x_f}},\mathbf{\Sigma_{x_f}})$ from GP_f.

Reverse Inference More complicated is the reverse inference where we infer the goal from observations of the current and final state, $p(G|X_i = x_i, X_f = x_f)$. This computation is similar to the filtering and smoothing computations used in Gaussian dynamical systems (reviewed in (Deisenroth, 2010)). To begin, we set $X_i = x_i$ and perform forward inference using our prior $p(G = g)$. We obtain $p(A|x_i) = \mathcal{N}(a|\mu_a,\Sigma_a)$ and $p(X_f|x_i) = \mathcal{N}(x_f|\mu_{x_f},\Sigma_{x_f})$. Now, we approximate the joint $p(A,X_f|x_i)$ with a Gaussian.[4] The remaining terms, the cross terms of the joint covariance $\Sigma_{ax_f} = \mathbb{E}_{ax_f}[ax_f^\top] - \mu_a\mu_{x_f}^\top$, are computed as in (Deisenroth et al., 2009; Deisenroth, 2010). Now that we have the joint, we incorporate the measurement x_f by applying standard Gaussian conditioning formulas to the joint to get $p(A|x_i,x_f) = \mathcal{N}(a|\mu_a^p,\Sigma_a^p)$, where $\mu_a^p = \mu_a + \Sigma_{ax_f}(\Sigma_{x_f})^{-1}(x_f - \mu_{x_f})$, and $\Sigma_a^p = \Sigma_a - \Sigma_{ax_f}(\Sigma_{x_f})^{-1}\Sigma_{ax_f}^\top$.

However, we want a distribution over goals, not over actions. Rather than attempting to find a closed-form solution to the joint $p(G,X_f|x_i)$, we sample from $p(A|x_i,x_f)$ (we approximate the expectation with a finite sum) and use those as the "measurements" when we compute $p(G,A|x_i,x_f)$. These

[2]We determine the hyperparameters, θ, of each GP by maximizing the marginal likelihood $p(\mathbf{y}|\mathbf{X},\theta)$ with respect to the hyperparameters, which is equivalent to maximizing the posterior distribution over the hyperparameters for a flat hyperprior $p(\theta)$. This optimization was done using the gpml software, available at `http://www.gaussianprocess.org`.

[3]Technically, we use the formulas for multivariate inputs and outputs which can be found in (Deisenroth, 2010).

[4]This is a standard approximation used in Gaussian filters such as the UKF (Julier & Uhlmann, 2004) and the GP-ADF (Deisenroth et al., 2009).

"measurements" act in the same way as the measured x_f, above, and allow us to condition out on the observed $A = a$ in order to get our desired result.

Results

Implementation Details

To test our model, we randomly sample goal positions and then compute the action required to fixate on this goal. We add Gaussian noise to this action, compute the resulting gaze vector if this action were taken, and add Gaussian noise to this gaze vector. This method is equivalent to training the model with rejection sampling wherein the agent rejects all samples that do not result in successful fixation on the goal position. The Gaussian processes are trained on this randomly generated data and then tested on separate test data.

The default reference frame for both agent and mentor is at the origin gazing along the x-axis. Each agent has their own reference frame and we assume that we know the transformation from the mentor's reference frame to the agent's. This transformation is not learned by our model; however, we believe that this is a minor assumption; especially since we already assume the agent can observe the mentor's position and gaze angle for our model. The mentor and agent are positioned as shown in Figure 3. Goal locations for the training data were generated uniformly at random from the area between the agent and the mentor (within the rectangle formed by $x = [100, 500], y = [-500, 500]$, where the agent is at $(0, 0)$ and the mentor is at $(500, 0)$).

We used Gaussian noise with standard deviation of 3 degrees for angles and a standard deviation of 10 cm for locations and distances. For reverse inference, the prior goal state $p(G)$ is a Gaussian centered halfway between the two agents along the x-axis with a very high variance. While this prior is quite weak, a single observation of (x_i, x_f) is insufficient to overcome the prior in reverse inference. Instead, we use a single observation of x_i and five observations of x_f to get an accurate estimate of $p(G|x_i, x_f)$. More precisely, we run reverse inference with the observed values $(x_i, x_f^{(1)})$ to compute $p(G|x_i, x_f^{(1)})$, and then use this as the prior for a second run of reverse inference to compute $p(G|x_i, x_f^{(1:2)})$. We repeat this five times to compute $p(G|x_i, x_f^{(1:5)})$. We believe that this is a reasonable number of observations for a human to make in the short amount of time taken for gaze following.

Model Performance

Overall, we found that the model performs quite well. It learns accurate transition and policy functions from small amounts of noisy training data ($n = 200$ data points were used in our accuracy tests) and the nonparametric nature of Gaussian processes ensures very little customization is required. The computational complexity of evaluating our model for gaze following is $\mathcal{O}(E^3) + \mathcal{O}(DE^2n^2)$, where D and E are the dimensionalities of the training inputs and the training targets, respectively, and n is the size of the training set. As such, our model may not scale well to high dimensions without additional approximations; however, it currently runs in sub-minute times for the dimensionality we are using. Additionally, the Gaussian approximations we are making have little effect because the data is generally unimodal and close to symmetric.

Figure 2 shows performance results for our model as it performs forward inference, reverse inference, and gaze following (combined reverse and forward inference). Other than a few outliers, the estimated values of the model are accurate and precise. The model is robust to noise and is able to provide strong gaze following results even though additional levels of uncertainty are introduced by the second level of inference.

Blindfold Self-Experience Task

In order to validate the cognitive plausibility of our model, we recreate a cognitive science experiment from (Meltzoff & Brooks, 2008), where infants' self-experience with a blindfold affects whether or not they follow the gaze of blindfolded adults. In their tests, one third of the 96 infants are given experience with a blindfold, another third are given experience with a "windowed" blindfold, and the remainder gain no experience with either. The children then interact with an adult experimenter for four trials where, in each trial, the experimenter stops playing with the infant, places the blindfold over her eyes, and then silently turns her head to align with one of two targets (placed between the experimenter and the infant but offset to the left and right). The trials are given a score of $+1$ if the infant looks in the direction of the target, -1 if the infant looks in the direction of the other target, and 0 if the infant looks elsewhere. The looking score is calculated as the sum of correct looks, incorrect looks, and no-looks and thus the possible looking score range is $[-4, +4]$.

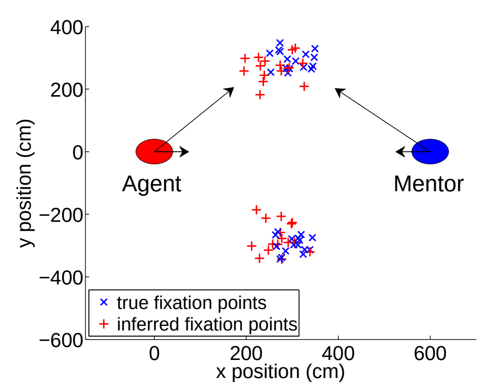
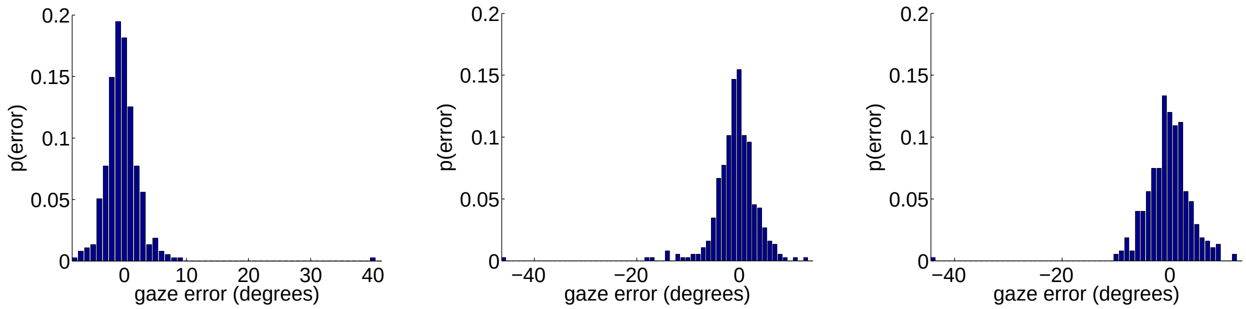


Figure 3: Experimental setup of a gaze following task. Inferred goal positions are shown (red) next to true goal positions (blue). Black arrows represent the initial and final gaze vectors of the agent and mentor for a single test data point.

Similarly, in our experiment we train 60 separate agents with our model on randomly generated training data. One third of these agents are given additional experience with a blindfold wherein they train an additional GP_π for their model

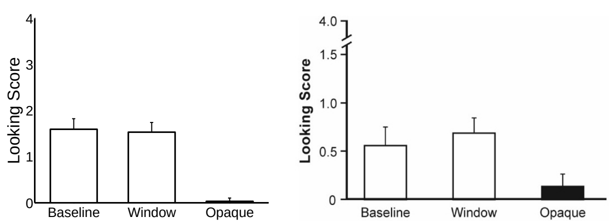


(a) Forward inference (b) Reverse inference (c) Gaze following

Figure 2: Histogram plots showing the probability of an error (in degrees) between the inferred and the true gaze vector. These probabilities were estimated from 375 test points spread uniformly over the test region. Note how the accuracy gracefully decays as more complicated inference is performed ((a) is the simplest, while (c) is the most complex).

so that they now have GP_π and GP_π^b, where GP_π is the original GP of the model and GP_π^b is the GP with blindfold experience. The other 40 agents are trained normally although one group is the baseline group and the other is the windowed group. However, in our formulation, the windowed blindfold in the model amounts to presenting the same training data as the no-blindfold case, so these two groups will be identical except for noise. Each agent is presented with 4 trials where it observes a mentor make a head turn to face either 45 degrees to the left or 45 degrees to the right (plus noise). Trials are scored as +1 if the agent turns its head at least 30 degrees in the direction of the correct target and −1 if it turns its head at least 30 degrees in the direction of the wrong target.

The agents with no blindfold experience use the basic model (which contains no blindfold knowledge) and thus assume that the mentor is fixating on an object to the left or to the right. Those with blindfold experience observe that the mentor is wearing a blindfold and use their learned GP_π^b for the reverse inference (applying their model to the mentor). For this group, the agent has (hopefully) learned that, when blindfolded, there is no correlation between the mentor's gaze and the mentor's goal position. The blindfold-experienced agent then uses GP_π for the forward inference (because the agent is not wearing a blindfold).



(a) Computational Model (b) Infant Data

Figure 4: Comparison of computational model to actual collected infant data from (Meltzoff & Brooks, 2008).

In order to simulate infants with little experience with blindfolds, we provided the agents in this experiment with very little training data ($n = 15$). If we use more training data, the agents perform almost perfectly in this task. Results are shown in Figure 4 and match those of Meltzoff and Brooks (Figure 4(a)). This is indicative that gaze following involves understanding the underlying intention or goal and that self-experience plays a major role in learning the consequences of intentions and related actions.

Related Work

Our model is closely related to the goal-based imitation model of Verma and Rao (Verma & Rao, 2006) and the inverse planning model of Baker et al. (Baker, Saxe, & Tenenbaum, 2009). Unlike these previous models, which assume discrete state and action spaces, our model is based on a nonparametric model that allows learning and inference in continuous state and action spaces. Also related to our approach are models for planning based on probabilistic inference (Toussaint & Storkey, 2006; Verma & Rao, 2007; Botvinick & An, 2009). Unlike these models, the gaze following model proposed herein involves inference over a single time-step, which simplifies the inference problem. We hope to explore the applicability of our model for multi-step inference problems in future work. Our model for gaze following joins the growing number of Bayesian models for cognition proposed in recent years and acknowledges the psychophysical and neurobiological evidence for Bayesian mechanisms in perception and action (e.g., (Rao, Olshausen, & Lewicki, 2002; Oaksford & Chater, 2007)).

Within the realm of gaze following models, one class of models posits that young infants watch an adult's head movement in space and are drawn to the correct hemi-field where they are attracted to a salient target object (Butterworth & Jarrett, 1991); over time, they learn to follow gaze to an object (Moore, 1999). A second class of models supports the nativist view that infants have a built-in module for interpreting eye gaze in terms of visual experience in others (Baron-Cohen, 1995). A third class of models adopts the developmental view that gaze following behavior emerges from self-experience (Meltzoff & Brooks, 2007). Our model can be regarded as a

Bayesian example of this last class of models.

Summary and Conclusion

This paper proposes a Bayesian framework for social interaction that postulates that (1) children learn a probabilistic model of the sensory consequences of their own movements through self-experience, and (2) they use this learned model to interpret the actions of others. Specifically, we show how gaze following can be modeled as goal inference within such a framework: probabilistic inference over the unknown variables in a learned graphical model allows an agent to infer another's gaze direction as well as the action to direct gaze to the same location. When given self-experience with blindfolds, the model learns the consequences of occlusion and subsequently does not follow the gaze of a blindfolded agent, replicating infant gaze following results.

The proposed framework raises several interesting questions: (1) Can the model be extended to other goal-based imitation tasks, e.g., goal-directed reaching? (2) This paper explored a nonparametric implementation based on Gaussian processes but what are possible ways of neurally implementing the learned graphical model? (3) The paper assumes that states in the environment are known (corresponding to the case of MDPs or Markov decision processes) – how does the model extend to the more realistic case where only observations of states are available (partially observable MDPs or POMDPs) and where learning involves reward-based mechanisms? We plan to explore these issues in future work.

Acknowledgments

This research was supported by an NSERC fellowship to A. Friesen and ONR Cognitive Science program grant no. N000140910097. We thank Marc Deisenroth, Andrew Meltzoff, and L. Elisa Celis for discussions. We also thank Marc for sharing his code.

References

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action Understanding as Inverse Planning. *Cognition*, *113*, 329 - 349.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Botvinick, M., & An, J. (2009). Goal-directed decision making in prefrontal cortex : A computational framework. *Advances in Neural Information Processing Systems (NIPS)*.

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology*, *38*(6), 958–966.

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*, 55 - 72.

Deisenroth, M. P. (2010). *Efficient reinforcement learning using gaussian processes*. Karlsruhe: KIT Scientific Publishing.

Deisenroth, M. P., Huber, M. F., & Hanebeck, U. D. (2009). Analytic moment-based gaussian process filtering. In *Proceedings of the 26th annual international conference on machine learning* (pp. 225–232). New York, NY, USA: ACM.

Julier, S. J., & Uhlmann, J. K. (2004). Unscented Filtering and Nonlinear Estimation. *IEEE Review*, *92*(92), 401 - 422.

Meltzoff, A. N. (2005). Imitation and other minds: The "like me" hypothesis. In S. H. . N. Chater (Ed.), *Perspectives on imitation: From cognitive neuroscience to social science* (pp. 55–77). Cambridge, MA: MIT Press.

Meltzoff, A. N., & Brooks, R. (2007). Eyes wide shut: The importance of eyes in infant gaze following and understanding other minds. In R. Flom, K. Lee, & D. Muir (Eds.), *Gaze-following: Its development and significance* (p. 217 - 241). Mahwah, NJ: Erlbaum.

Meltzoff, A. N., & Brooks, R. (2008). Self-experience as a mechanism for learning about others: A training study in social cognition. *Developmental Psychology*, *44*(5), 1257 - 1265.

Moore, C. (1999). Gaze following and the control of attention. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life* (p. 241 - 256). Mahwah, NJ: Erlbaum.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press.

Quiñonero-Candela, J., Girard, A., Larsen, J., & Rasmussen, C. (2003). Propagation of Uncertainty in Bayesian Kernel Models-Application to Multiple- Step Ahead Forecasting. *IEEE International Con- ference on Acoustics, Speech and Signal Processing*, 701 - 704.

Rao, R. P. N., Olshausen, B. A., & Lewicki, M. S. (2002). *Probabilistic Models of the Brain: Perception and Neural Function*. Cambridge, MA: MIT Press.

Rasmussen, C. E., & Ghahramani, Z. (2003). Bayesian Monte Carlo. *Advances in Neural Information Processing Systems*, *15*(15), 489 - 496.

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press.

Toussaint, M., & Storkey, A. (2006). Probabilistic inference for solving discrete and continuous state Markov Decision Processes. *23rd International Conference on Machine Learning*.

Verma, D., & Rao, R. (2006). Goal-based imitation as probabilistic inference over graphical models. *Advances in NIPS*, *18*.

Verma, D., & Rao, R. P. N. (2007). Imitation Learning Using Graphical Models. *ECML*, 757 - 764.