The Art of (Recursive Bayesian) Persuasion

Samuel A. Barnett (samuelab@princeton.edu)
Department of Computer Science, 35 Olden Street
Princeton, NJ 08540 USA

Abstract

Include no author information in the initial submission, to facilitate blind review. The abstract should be one paragraph, indented 1/8 inch on both sides, in 9 point font with single spacing. The heading "**Abstract**" should be 10 point, bold, centered, with one line of space below it. This one-paragraph abstract section is required only for standard six page proceedings papers. Following the abstract should be a blank line, followed by the header "**Keywords:**" and a list of descriptive keywords separated by semicolons, all in 9 point font, as shown below.

Keywords: AI; quantum; bitcoin; Elon Musk

Introduction

Modeling Evidence Effects

World Model: The Stick Task

The stick task consists of a judge, and a set of speakers indexed by I. A sample of N sticks whose lengths are given by

$$\mathcal{S}_N = \{S_1, S_2, \ldots, S_N\} \tag{1}$$

are drawn i.i.d. from the Uniform$[0,1]$ distribution and is fixed throughout the task.[1]

Each speaker observes the full sample before the task, and at each time step t a speaker i (one per time step, taking it in turn) chooses one stick from the sample to reveal to the judge. The speakers are not permitted to reveal a stick that has previously been shown to the judge. Notationally, the agent chooses action

$$a_t^{(i)} \in \{s_1, s_2, \ldots, s_N\} \setminus \mathcal{A}_{t-1}, \tag{2}$$

where s_i is the realization of random variable S_i and $\mathcal{A}_{t-1}$ denotes the first $t-1$ actions chosen. For simplicity, we let $\mathcal{A}_0 = \emptyset$.

At each time step t, the judge reasons about the sample mean of the sticks, $\bar{S} = \frac{1}{N}\sum_{n=1}^{N} S_n$.[2] In particular, the judge evaluates his posterior over whether the sample is 'long' or 'short', i.e., whether or not $\bar{S}_N \geq 0.5$. Hence, the relevant posterior for the judge at time step t is

$$p(\bar{S}_N \geq 0.5 \mid \mathcal{A}_t). \tag{3}$$

[1] In practice, we use a discrete approximation to this distribution by modeling the sticks as being drawn uniformly from the set $\{0.025, 0.075, \ldots, 0.975\}$.

[2] It is assumed that the judge knows N, but only observes the stick values through the speakers' actions.

Each speaker will have an incentive to select evidence that persuades the judge that the sample is either long or short, or the speaker will be indifferent towards the outcome.

The task runs for a total of $T \leq N$ time steps.

Agent Models

In the manner of Rational Speech Act (RSA) models, we model the agents as performing recursive Bayesian inference about one another's beliefs. To capture both evidence effects, we require four layers, which we divide into the 'naïve' layers and the 'pragmatic' layers. The model is implemented in WebPPL, a probabilistic programming language that allows for fast hierarchical inference in low-dimensional domains such as this one.

Naïve Judge and Speaker The first layer, J_0, describes the naïve judge - this judge is naïve in the sense that he does not model the speakers as having any incentives, and instead assumes that the actions are selected uniformly from the available sample. Relabeling without loss of generality, the posterior of J_0 at time t is given by

$$p_{J_0}(\bar{S}_N \geq 0.5 \mid S_1 = s_1, S_2 = s_2, \ldots, S_t = s_t). \tag{4}$$

The second layer, S_1, describes the naïve speaker, whose choice about which stick to show at time t is represented as a posterior over the available sticks defined in reference to p_{J0}. Importantly, each speaker has a bias β, where in our simulations we model the biases as being in the range $\beta \in \{-10, -5, -2, 0, 2, 5, 10\}$. This bias represents the incentive regarding the judge's inference: a positive (resp. negative) bias entails that the speaker is incentivized to show sticks to the judge that steer the judge towards the belief that the sample is long (resp. short).

To produce this behavior, the speaker samples a stick based on the soft-maximization of biased informativity of that stick, taking into account the previous sticks shown:

$$p_{S_1}(a_t \mid \mathcal{A}_{t-1},\ \mathcal{S}_N) \propto \exp\big(\beta \cdot a_t \cdot p_{J_0}(\bar{S}_N \geq 0.5 \mid \mathcal{A}_{t-1} \cup \{a_t\}\big). \tag{5}$$

Observe that, if $\beta = 0$, this entails that the speaker will sample the sticks uniformly based on the available evidence, which is equivalent to the speaker being indifferent to the outcome.

Pragmatic Judge and Speaker The pragmatic judge, J_1, performs a similar inference as J_0, though with one crucial difference: the judge models each stick as having been sampled by a naïve speaker. To see how this is possible, observe that we can express the action of speaker i at time t by the random variable

$$A_t^{(i)} \sim p_{S_1}(\cdot \mid \mathcal{A}_{t-1},\, \mathcal{S}_N), \tag{6}$$

which is a categorical distribution computed by Equation 5. For ease of notation, we write the set of observations of S_1 speakers

$$\mathcal{A}'_T = \{A_1^{(i_1)} = a_1^{(i_1)}, A_2^{(i_2)} = a_2^{(i_2)}, \ldots, A_T^{(i_T)} = a_T^{(i_T)}\}, \tag{7}$$

and let $\mathcal{A}'_0 = \emptyset$.

The posterior of J_1 can now be computed via Bayes' rule:

$$\begin{aligned} p_{J_1}(\mathcal{S}_N \mid \mathcal{A}'_T) &\propto p(\mathcal{A}'_T \mid \mathcal{S}_N)p(\mathcal{S}_N) \\ &= p(\mathcal{S}_N)\prod_{t=1}^{T} p_{S_1}(A_t^{(i_t)} = a_t^{(i_t)} \mid \mathcal{A}'_{t-1}) \end{aligned}$$

Using the sum rule, we arrive at the relevant posterior, which in the case of discretized stick lengths is given by

$$p_{J_1}(\bar{S}_N \geq 0.5 \mid \mathcal{A}'_T) = \sum_{\mathcal{S}_N:\, \bar{S}_N \geq 0.5} p_{J_1}(\mathcal{S}_N \mid \mathcal{A}'_T). \tag{8}$$

We consider two distinct versions of the pragmatic judge: one in which he knows the biases of each agent in advance, and one in which these biases are drawn, i.i.d., from a categorical prior over the aforementioned range of bias values. This is necessary as the weak and strong evidence effects demand the former and latter versions, respectively.

We further divided the latter case into two by considering two possible priors: a flat prior over the bias range, and a 'V-shaped' prior which down-weights the likelihood of neutrality.[3] We considered these two priors as, although the V-shaped prior better describes the context of the stick task, we wished to show that strong evidence effects were robust to our choice of prior. Factoring in uncertainty over the bias is then a matter of marginalizing over the possible bias settings using the sum rule.

The final layer, the pragmatic speaker S_2, is nearly identical to the naïve speaker S_1, except for the fact that the pragmatic speaker performs soft-maximization based on J_1's judgment, as opposed to J_0's judgment:

$$p_{S_2}(a_t \mid \mathcal{A}_{t-1},\, \mathcal{S}_N) \propto \exp\big(\beta \cdot a_t \cdot p_{J_1}(\bar{S}_N \geq 0.5 \mid \mathcal{A}_{t-1} \cup \{a_t\}\big). \tag{9}$$

This layer is only required to capture the strong evidence effects.

[3]Concretely, the V-shaped prior is given by the normalized probability vector $\frac{1}{29}[8,4,2,1,2,4,8]$.

Simulations

Simulation 1: Weak Evidence Effect

Set-Up In the first simulation, we captured the pragmatic judge's ability to capture the *weak evidence effect*. We present the judge with one stick in turn from two speakers, whose biases are known to the judge and fixed at $\beta_1 \in \{2,5,10\}, \beta_2 = -\beta_1$. We then contrast the naïve and pragmatic judge's posteriors over whether the sample is long, having seen both sticks.

In this context, the weak evidence effect occurs when the naïve judge *decreases* his belief that the sample is long between observing the first and the second stick, while the pragmatic judge *increases* his belief. To see why this is the case, recall that the weak evidence effect is a result of a conditional probability being judged lower than the marginal, while the cause is probability-raising. In the stick task, we evaluate whether the cause is probability-raising with respect to a naïve judge, who effectively regards the sticks as being sampled i.i.d., whereas the 'conditional' and 'marginal' in question (that is, the conditional probabilities after the first and second stick observation) are evaluated with respect to the pragmatic judge with a more nuanced sampling assumption.

For this simulation, we vary the total sample size N between 3 and 5, and we consider the full range of possible stick pairs to be presented.

CoGNiTiVe ScIeNcE

Figure 1: This is a figure.

Results

Simulation 2: Strong Evidence Effect for Speakers

Set-Up In the second simulation, we investigate the strong evidence effect for *speakers*, and in particular, whether adding the pragmatic speaker is truly necessary to capture this effect. We consider a simple speaker with a bias $\beta \in \{2,5,10\}$. The stick task is then run for $T = 2$ steps, with $N = 5$ available sticks. In contrast to Simulation 1, the pragmatic judge (as modelled by the pragmatic speaker) no longer knows the speaker bias, and instead has either a flat or a V-shaped prior over possible bias values.

For each setting of bias value and bias prior shape, we consider 150 draws of initial stick samples (recording the true sample mean) for the speaker to choose from. We then look at the distribution over possible ordered pairs of sticks that the speaker chooses to show the judge, which we refer to as 'strategies'. For computational reasons, this distribution is inferred via a Monte Carlo estimate with 100 samples, as opposed to using the enumeration method that governs the lower-order inferences.

To study the effect, we compare this distribution of strategies to what we call the 'optimal strategy': to show the sticks in descending order. Given only a naïve judge, this is the strategy that maximizes the judge's posterior probability that the sample is long. Hence, as $\beta \to \infty$ we would expect the probability mass of this strategy to tend to 1 for the naïve speaker. We therefore compare the probability mass of this strategy for the naïve speaker to the pragmatic speaker, to see if the taking into account that the judge is modeling the speaker bias reduces the likelihood that the speaker will adopt this strategy, and instead leads to a different Maximum A Posteriori (MAP) strategy. Such a reduction would show that the pragmatic speaker demonstrates a strong evidence effect.

Results Plotting the probability mass on the optimal probability for the naïve speaker against the pragmatic speaker, we see that, in the majority of cases (**give a proportion**), the naïve speaker has a higher probability of choosing the optimal strategy than the pragmatic speaker. In particular, the probability of choosing the optimal strategy for the pragmatic speaker never exceeds ε higher than the probability for the naïve speaker, and can (and often is) much lower. This shows that the pragmatic speaker is able to capture the strong evidence effect.

Interestingly, we see little difference in whether the judge is modeled to have a flat prior or a V-shaped prior over speaker biases, although the strong evidence effect is slightly more prevalent for flat priors. This may be because, if the speaker believes the judge to already be disposed towards believing that the speaker is biased, then it may be true for the speaker that choosing her sticks non-optimally has less of an impact on the judge's belief about her bias, and so the optimal strategy may look slightly more favorable.

We also see that, for both the naïve and pragmatic speaker, the probability of choosing the optimal strategy *increases* as agent bias increases and *decreases* as sample mean increases. The former phenomenon is clearly true for the naïve speaker, whereas for the pragmatic speaker we can argue that increasing the bias begins to crowd out considerations about how the judge perceives that bias. The latter phenomenon can be explained as follows: in the domain in which the sample mean is high, the speaker can afford to choose a sub-optimal strategy and still get her desired outcome, though if the sample mean is low it becomes increasingly important to just show the longest sticks.

The strong evidence effect also appears by looking at the MAP strategies in Table 1, which shows that (with one exception), the number of inferred distributions for which the optimal strategy was the MAP strategy decreases as we move from a naïve to a pragmatic speaker. Moreover, this gap increases as we increase the bias of the agent.

CoGNiTiVe ScIeNcE

Figure 2: This is a figure.

Table 1: Sample table title.

Error type	Example
Take smaller	63 - 44 = 21
Always borrow	96 - 42 = 34
0 - N = N	70 - 47 = 37
0 - N = 0	70 - 47 = 30

Simulation 3: Strong Evidence Effect for Judges

Set-Up In this simulation, we investigate the strong evidence effect for *judges*, which predicts a relationship between the strength of the evidence presented by a speaker, and the judge's perception of the speaker's bias. We again consider one speaker, and look at the judge's posterior probability of the speaker's bias value after observing one stick, where this distribution is computed exactly via enumeration.[4] We also vary whether the judge begins with a flat prior over bias values, or a V-shaped prior.

Results Figure 3 shows, predictably, that an increase stick length leads to an increased posterior probability in higher bias values, whereas for lower bias values this increase eventually plateaus or slightly decreases. This is clearer in the case of a flat bias prior than a V-shaped prior, where the initial skew in belief towards greater bias values suppresses the belief that the speaker is positively biased for lower stick values. However, in both cases the overall shape and trend of the curves are the same.

CoGNiTiVe ScIeNcE

Figure 3: This is a figure.

Discussion and Future Work

References

Chalnick, A., & Billman, D. (1988). Unsupervised learning of correlational structure. In *Proceedings of the tenth annual conference of the cognitive science society* (pp. 510–516). Hillsdale, NJ: Lawrence Erlbaum Associates.

Feigenbaum, E. A. (1963). The simulation of verbal learning behavior. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought.* New York: McGraw-Hill.

Hill, J. A. C. (1983). A computational model of language acquisition in the two-year old. *Cognition and Brain Theory*, *6*, 287–317.

Matlock, T. (2001). *How real is fictive motion?* Doctoral dissertation, Psychology Department, University of California, Santa Cruz.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

[4] Clearly, the judge does not know *a priori* what the speaker's bias value is in this context, in contrast to Simulation 1.

Ohlsson, S., & Langley, P. (1985). *Identifying solution paths in cognitive diagnosis* (Tech. Rep. No. CMU-RI-TR-85-2). Pittsburgh, PA: Carnegie Mellon University, The Robotics Institute.

Shrager, J., & Langley, P. (Eds.). (1990). *Computational models of scientific discovery and theory formation*. San Mateo, CA: Morgan Kaufmann.