Modeling Bounded Rationality of Agents During Interactions

Qing Guo and Piotr Gmytrasiewicz
Department of Computer Science
University of Illinois at Chicago
Chicago, IL 60607
qguo, piotr@cs.uic.edu

Abstract

Frequently, it is advantageous for an agent to model other agents in order to predict their behavior during an interaction. Modeling others as rational has a long tradition in AI and game theory, but modeling other agents' departures from rationality is difficult and controversial. This paper proposes that bounded rationality be modeled as errors the agent being modeled is making while deciding on its action. We are motivated by the work on quantal response equilibria in behavioral game theory which uses Nash equilibria as the solution concept. In contrast, we use decision-theoretic maximization of expected utility. Quantal response assumes that a decision maker is rational, i.e., is maximizing his expected utility, but only approximately so, with an error rate characterized by a single error parameter. Another agent's error rate may be unknown and needs to be estimated during an interaction. We show that the error rate of the quantal response can be estimated using Bayesian update of a suitable conjugate prior, and that it has a finitely dimensional sufficient statistic under strong simplifying assumptions. However, if the simplifying assumptions are relaxed, the quantal response does not admit a finite sufficient statistic and a more complex update is needed. This confirms the difficulty of using simple models of bounded rationality in general settings.

1 Introduction

In AI, an agent's (perfect) rationality is defined as the agent's ability to execute actions that, at every instant, maximize the agent's expected utility, given the information it has acquired from the environment (Russell and Norvig 2010). Let us note two aspects of this definition. First, the fact that the acquired information may be limited does not preclude perfect rationality. In other words, an agent may have very limited information but still be perfectly rational. Second, the above definition does not specify any particular procedure an agent is to use to decide which action to execute. Further, the definition is completely independent of any details of the implementation of any such decision making procedure.

The notion of bounded rationality received a lot of attention in economics and psychology. Simon (1955) coined the term and suggested it as an alternative to rationality. Simon pointed out that perfectly rational decision making is often difficult in practice due to limited cognitive and/or computational resources. He proposed that humans are *satisficers*, as opposed to perfect optimizers, and that they use heuristics to make decisions, rather than optimization rules. Gigerenzer (Gigerenzer and Goldstein 1996; Gigerenzer 2000) argued that simple heuristics could actually lead to better decisions than theoretically optimal procedures. The use of heuristics was also studied by Kahneman (Kahneman and Tversky 1982), who proposed his own alternative to perfect rationality called prospect theory. Rubinstein (1998) proposed that one needs to model an agent's decision-making procedures explicitly in order to model the agent's bounded rationality adequately.

This paper builds on an approach to modeling bounded rationality called quantal response (Camerer 2003; McKelvey and Palfrey 1995; 1998). It is a simple model which uses a single error parameter. Quantal response is simple in that it does not attempt to model the procedures, and their possible limitations, the agent may use to decide on its action. The great advantage of this model is that, first, there exist a myriad of procedural mechanisms by which perfect rationality could be implemented, heuristics which could be used, and possible ways in which any of these could have its functionality limited by the specific computational or cognitive architecture of the agent in question. Second, none of these implementation details and architectural limitations are observable by the external observer who is doing the modeling. In other words, quantal response abstracts away the unobservable parameters specific to implementation and treats them as *noise* which produces non-systematic departures from perfect rationality.

To make room for bounded rationality of the other agents we define a notion of approximately intentional agent model. It is analogous to perfectly rational agent model but with a noise factor inversely proportional to an error parameter, λ. According to quantal response (Camerer 2003; McKelvey and Palfrey 1995; 1998), probabilities of actions are given by the logit function of the actions' expected utilities. Thus actions that are suboptimal are possible, but their probabilities increase with their expected utilities.

Quantal response specifies the probabilities of an agent's actions given their expected utilities and the agent's error

parameter, λ. An additional complication is that an agent's error parameter is not directly observable. Instead, it must be inferred based on the agent's observed behavior. We take a Bayesian approach to this and propose that the modeling agent maintain a probability distribution over possible values of λ for the modeled agent, and that this probability be updated when new actions are observed. Intuitively, if an agent is observed acting rationally, then over time the error rate attributed to this agent should decrease (and, since λ is an inverse error, larger values of λ should become more likely). If, on the other hand, the modeled agent is frequently observed acting in ways that depart from perfect rationality, then the error rate attributed to it should increase (and smaller values of λ should become more likely).

Below we show how the update of the error parameter modeling bounded rationality of another agent can be performed. We also show that, in simple special cases, when the interaction is *episodic*, the error rate admits a sufficient statistic. We then derive a distribution over λ that is a member of a family of conjugate priors. That means that the update of the distribution over λ is particularly simple and that it results in another distribution in the same family of parametrized distributions. We further show that if the simplifying assumptions are relaxed then there is no sufficient statistic of finite dimension and no conjugate prior over λ.

2 Logit Quantal Response

For simplicity, we assume that a modeling agent, called i, is considering the behavior of one other agent, j. The logit quantal response is defined as follows (Camerer 2003; McKelvey and Palfrey 1995; 1998):

$$P(a_j) = \frac{e^{\lambda u_{a_j}}}{\sum_{l=1}^{m} e^{\lambda u_{a_l}}}, \tag{1}$$

where $\{a_l : l = 1, 2, 3, ..., m\}$ is a set of all possible actions of the agent. $P(a_j)$ is the probability of the agent j taking the action a_j. $u_{a_j} \in \mathbb{R}$ is the expected utility of action a_j to agent j and $\lambda \geq 0$ is the (inverse) error rate of agent j. λ represents how rational agent j is: greater λ makes it more likely that j takes actions which have higher utilities. When $\lambda \to +\infty, P(a_j) = 1$ for the action which has the highest expected utility[1] and $P(a_j) = 0$ for all other actions. This means agent j is perfectly rational because he always chooses an action with the best expected utility. When $\lambda = 0$, $P(a_j) = 1/m$, $\forall j = 1, 2, 3, ..., m$, which means agent j chooses actions at random.

It is likely that the error rate λ of agent j is not directly observable to agent i. Bayesian approach allows agent i to learn this rate during interactions. To do this agent i needs a prior distribution, $f(\lambda)$, which represents i's current knowledge about agent j's error rate, and to observe agent j's action, a_j at the current step. The updated distribution is:

$$f(\lambda|a_j) = \frac{P(a_j|\lambda)f(\lambda)}{\int_0^\infty P(a_j|\lambda')f(\lambda')\,d\lambda'}. \tag{2}$$

[1]If there are many, say h, optimal actions with the same expected utilities, then $P(a_j) = 1/h$ for each of them.

Using the above equation, agent i can maintain his knowledge about agent j's bounded rationality by repeatedly updating $f(\lambda)$ during interaction.

Equation (2) may not be easy to apply because after updating the $f(\lambda)$ several times, it becomes more and more complicated. To overcome this it is convenient to look for a conjugate prior family. In Bayesian probability, if the posterior distribution is in the same family as the prior distribution, then this prior is called a *conjugate prior* (DeGroot 2004; Fink 1997). Conjugate priors are convenient because they make the updating process tractable; one just needs to update the parameters of the conjugate prior distribution (hyperparameters) to realize the Bayesian update.

3 Static Episodic Environments with Perfect Observability

In this section we consider the simplest case, when agent j's expected utilities u_{a_l} for all actions are *known* to agent i and remain the same during the interaction. In other words, agent j is not updating his beliefs since the environment is static and episodic (Russell and Norvig 2010) and i is observing j acting in the same decision-making situation repeatedly. The derivation and proof below follow techniques in (DeGroot 2004; Fink 1997).

Consider the following family of distributions over λ:

$$f(\lambda; u, n) = \frac{e^{\lambda u}/(\sum_{l=1}^{m} e^{\lambda u_{a_l}})^n}{\int_0^\infty e^{\lambda' u}/(\sum_{l=1}^{m} e^{\lambda' u_{a_l}})^n\,d\lambda'}, \tag{3}$$

where n and u are hyperparameters. Here n is a natural number including zero, and u is restricted by following: $u < n \max_l u_{a_l}$. One can verify (3) is a probability density function since $\int_0^\infty f(\lambda; u, n)\,d\lambda = 1$. See Appendix A for proof.

Proposition 1: The family of distributions $f(\lambda; u, n)$ in (3) is a conjugate family of distributions over λ in static episodic environments with known utilities of actions.

Proof: To verify that $f(\lambda; u, n)$ is a conjugate prior in this case we use (1), (2) and (3) to get the following (denote $1/\int_0^\infty e^{\lambda' u}/(\sum_{l=1}^{m} e^{\lambda' u_{a_l}})^n\,d\lambda' = c(u, n)$ for simplicity, where $c(u, n) > 0$ is a constant for fixed u and n, and assume the new observed action of agent j is a_j):

$$\begin{aligned}
f(\lambda|a_j) &= \frac{P(a_j|\lambda)f(\lambda; u, n)}{\int_0^\infty P(a_j|\lambda')f(\lambda'; u, n)\,d\lambda'} \\
&= \frac{\frac{e^{\lambda u_{a_j}}}{\sum_{l=1}^{m} e^{\lambda u_{a_l}}} \cdot \frac{e^{\lambda u}}{(\sum_{l=1}^{m} e^{\lambda u_{a_l}})^n} \cdot c(u, n)}{\int_0^\infty \frac{e^{\lambda' u_{a_j}}}{\sum_{l=1}^{m} e^{\lambda' u_{a_l}}} \cdot \frac{e^{\lambda' u}}{(\sum_{l=1}^{m} e^{\lambda' u_{a_l}})^n} \cdot c(u, n)\,d\lambda'} \\
&= \frac{e^{\lambda(u+u_{a_j})}/(\sum_{l=1}^{m} e^{\lambda u_{a_l}})^{n+1}}{\int_0^\infty e^{\lambda'(u+u_{a_j})}/(\sum_{l=1}^{m} e^{\lambda' u_{a_l}})^{n+1}\,d\lambda'} \\
&= f(\lambda; u + u_{a_j}, n + 1).
\end{aligned}$$

□

The proof above also established how to update the hyperparameters of our conjugate prior after observing that agent j executed his action a_j, with expected utility u_{a_j} :

$$f(\lambda; u, n) \xrightarrow{a_j} f(\lambda; u + u_{a_j}, n + 1). \quad (4)$$

Note that the integral in the denominator of $f(\lambda; u, n)$ does not always have an analytical solution, so we have to use numerical methods to calculate its value.

One can verify that once there is a valid prior, all the posteriors are always valid. The question also arises as to what is an appropriate prior agent i should choose before any observations. Often one looks for an uninformed prior. In our case $f(\lambda; -\epsilon, 0)$, where $\epsilon > 0$ is a small positive value, is such an uninformed prior; it is almost flat over the positive real values of λ, as we show in the example below.

3.1 Example

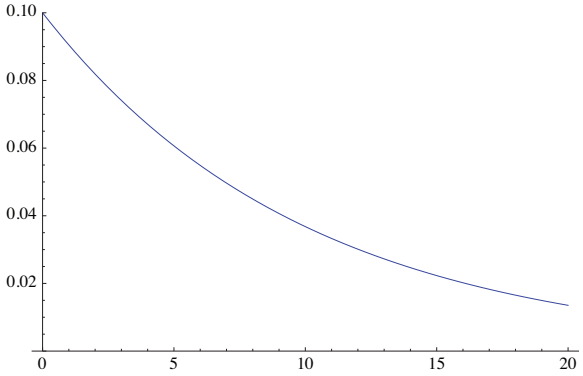

Figure 1: Example conjugate prior: $f(\lambda; -0.1, 0)$

Let us assume that agent j chooses from among three ($m = 3$) actions, with following expected utilities: $u_{a_1} = 0, u_{a_2} = 2, u_{a_3} = 10$. As we mentioned, we assume that the expected utilities of agent j are known to agent i, and that they do not change. Let the prior be $f(\lambda; -0.1, 0)$. Let us first compute the expected value of the error parameter i attributes to j under this distribution: $E(\Lambda) = \int_0^\infty \lambda f(\lambda)\, d\lambda = 10.0$. Using the formula of total probability for each action of j we get: $P(a_j) = \int_0^\infty P(a_j|\lambda) f(\lambda)\, d\lambda$. Thus the prior probabilities i attributes to each of j's actions are: $P(a_1) = 0.00524, P(a_2) = 0.00699, P(a_3) = 0.98777$. Figure 1 shows the initial prior. Note that this uninformative prior assigns relatively high probabilistic weight to high values of λ and hence high degree of j's rationality.

λ's Distribution	$E(\Lambda)$	$P(a_1)$	$P(a_2)$	$P(a_3)$
$f(\lambda; -0.1, 0)$	10.0000	0.00524	0.00699	0.98777
$f(\lambda; 29.9, 3)$	10.2477	0.00138	0.00229	0.99633
$f(\lambda; 299.9, 30)$	10.5309	0.00010	0.00029	0.99961
$f(\lambda; 11.9, 3)$	0.1069	0.20663	0.24147	0.55190
$f(\lambda; 119.9, 30)$	0.0328	0.28959	0.30785	0.40256

Table 1: Probabilities of agent j's actions derived from various distributions over error parameter λ in a static episodic environment.

Assume agent j acts rationally and always chooses his best action, a_3. Then Figure 2 and Figure 3 show the posteriors after three observations ($f(\lambda; 29.9, 3)$) and after 30 observations ($f(\lambda; 299.9, 30)$) of j's action a_3. We can see that higher values of λ become more likely if the agent always chooses the action with the best utility. We can also compute the probabilities of the three actions under these two posteriors, which are shown in Table 1.

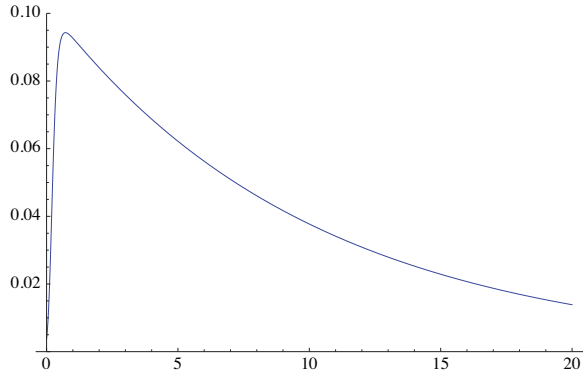

Figure 2: $f(\lambda; 29.9, 3)$, updated after observing three rational actions.

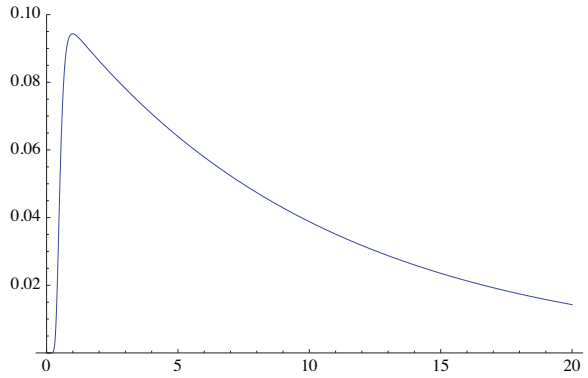

Figure 3: $f(\lambda; 299.9, 30)$, updated after observing 30 rational actions.

Now let us assume that agent j behaves randomly. Within the first three actions, he chooses each of his actions a_1, a_2 and a_3 once. The updated distribution over the error parameter is then $f(\lambda; 11.9, 3)$, which is shown in Figure 4. Further, if within j's 30 actions he chooses a_1 for ten times, a_2 ten times, and a_3 for ten times; then the posterior is $f(\lambda; 119.9, 30)$, which is shown in Figure 5. The results are intuitive. Thus, if agent j behaves randomly, lower values of λ, indicating stronger departure from perfect rationality, become more likely. Probabilities of the 3 actions under these two posteriors are also shown in Table 1.

4 Sequential Dynamic Environments with Perfect Observability of Finite Types

In this section, we extend our approach to more complex case of dynamic sequential environments. Again, we assume that expected utilities of j's actions are known to i, but now, since agent j may be updating his beliefs, the expected utilities of his actions do not remain constant but can take a finite

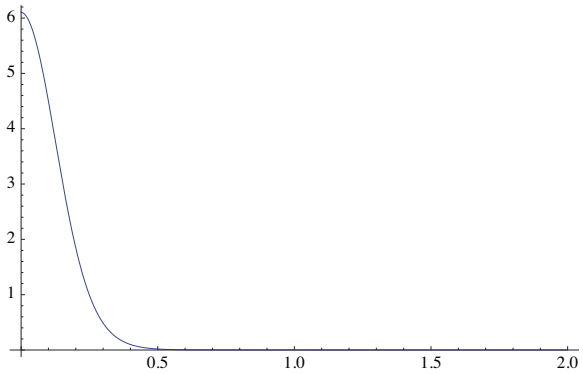

Figure 4: $f(\lambda; 11.9, 3)$, updated after observing three random actions.

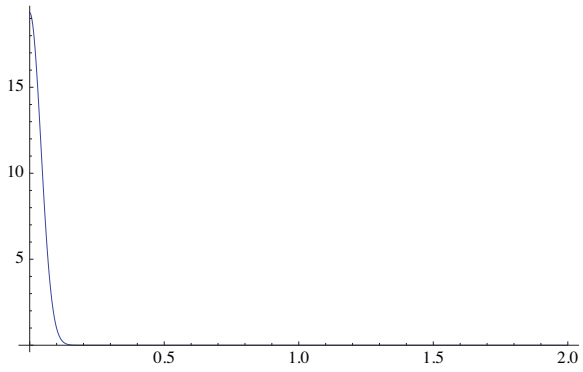

Figure 5: $f(\lambda; 119.9, 30)$, updated after observing 30 random actions.

number of values. We refer to each of the beliefs of agent j, together with his payoff function and other elements of his POMDP (Russell and Norvig 2010), as j's type, θ_j. Thus, the set of possible types of agent j, Θ_j, has K possible elements $1, 2, ..., K$. We denote $U(a_j|\theta_j = k) = u_{a_j,k}$, where $k = 1, 2, ..., K$, and assume that index k is observable (or computable) by agent i. Then the logit quantal response (1) for the probability of agent j taking action a_j given his kth type is:

$$P(a_j|k, \lambda) = \frac{e^{\lambda u_{a_j,k}}}{\sum_{l=1}^{m} e^{\lambda u_{a_l,k}}}. \tag{5}$$

Now Bayesian update, analogous to equation (2), becomes:

$$f(\lambda|a_j, k) = \frac{P(a_j|k, \lambda) f(\lambda)}{\int_0^\infty P(a_j|k, \lambda') f(\lambda')\, d\lambda'}. \tag{6}$$

We now have a proposition analogous to Proposition 1 in Section 3. Consider following family of distributions:

$$\begin{aligned} &f(\lambda; u, n_1, n_2, ..., n_K) \\ =&\frac{e^{\lambda u} / \prod_{k=1}^{K} (\sum_{l=1}^{m} e^{\lambda u_{a_l,k}})^{n_k}}{\int_0^\infty e^{\lambda' u} / \prod_{k=1}^{K} (\sum_{l=1}^{m} e^{\lambda' u_{a_l,k}})^{n_k}\, d\lambda'}, \end{aligned} \tag{7}$$

where $n_k = 0, 1, 2, ..., \forall k = 1, 2, ..., K$; and $u < \sum_{k=1}^{K} (n_k \max_l u_{a_l,k})$. One can verify that (7) is a valid probability density function since integral of the denominator converges if and only if $u < \sum_{k=1}^{K} (n_k \max_l u_{a_l,k})$. We skip the proof of this fact, which is similar to that in Appendix A.

Proposition 2: The family of distributions in (7), $f(\lambda; u, n_1, n_2, ..., n_K)$ is a conjugate family of distributions over λ in a sequential dynamic environment with perfect observability of finite number of types.

Proof: Analogous to that of Proposition 1.

Similarly to the simpler case of Proposition 1, the proof of Proposition 2 establishes the update of the hyperparameters of the conjugate prior based on the observed action, a_j, with expected utility $u_{a_j,k}$:

$$\begin{aligned} &f(\lambda; u, n_1, n_2, ..., n_K) \xrightarrow{a_j,k} \\ &f(\lambda; u + u_{a_j,k}, n_1, n_2, ..., n_{k-1}, n_k + 1, n_{k+1}, ..., n_K). \end{aligned} \tag{8}$$

Similarly to Section 3, once there is a valid prior, e.g. $f(\lambda; u, n_1, n_2, ..., n_K)$, all the posteriors are always valid. An uninformative prior agent i can choose before observing any of j's actions can be $f(\lambda; -\epsilon, 0, 0, ..., 0)$. Then after any number of observations the current u is the accumulated utility of all actions the agent has taken minus ϵ, and current n_k is the counter of occurrence of the kth type.

4.1 Example

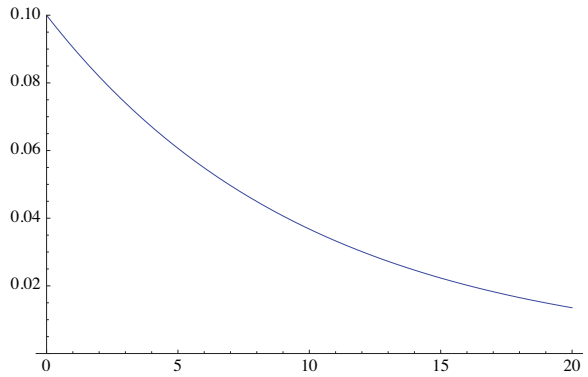

Figure 6: Example conjugate prior: $f(\lambda; -0.1, 0, 0)$

Similarly to Section 3.1, assume that agent j chooses from among three ($m = 3$) actions, and has two types $\theta_j = 1, 2$, with following expected utilities: $u_{a_1,1} = 0, u_{a_2,1} = 2, u_{a_3,1} = 10, u_{a_1,2} = 5, u_{a_2,2} = 15, u_{a_3,2} = 0$. We assume that the expected utilities and current type of agent j are known to agent i. Let the prior be $f(\lambda; -0.1, 0, 0)$. The expected value of the error parameter i attributes to j under this distribution and the prior probabilities i attributes to each of j's actions given his type (calculated by $P(a_j|k) = \int_0^\infty P(a_j|k, \lambda) f(\lambda)\, d\lambda$) are shown in Table 2. Figure 6 shows the initial prior.

Assume agent j acts rationally and always chooses his best action, a_3 in type 1 and a_2 in type 2. Then Figure

λ's Distribution	$E(\Lambda)$	$P(a_1\|\theta_j=1)$	$P(a_2\|\theta_j=1)$	$P(a_3\|\theta_j=1)$	$P(a_1\|\theta_j=2)$	$P(a_2\|\theta_j=2)$	$P(a_3\|\theta_j=2)$
$f(\lambda;-0.1,0,0)$	10.0000	0.005237	0.006986	0.987776	0.005754	0.990838	0.003409
$f(\lambda;74.9,3,3)$	10.2988	0.000823	0.001540	0.997637	0.000893	0.998853	0.000254
$f(\lambda;749.9,30,30)$	10.5597	0.000073	0.000222	0.999705	0.000074	0.999920	0.000006
$f(\lambda;31.9,3,3)$	0.0605	0.254935	0.282766	0.462299	0.275487	0.507722	0.216791
$f(\lambda;319.9,30,30)$	0.0189	0.308014	0.319443	0.372543	0.319654	0.388056	0.292290

Table 2: Probabilities of agent j's actions derived from various distributions over error parameter λ in a sequential dynamic environment with two types.

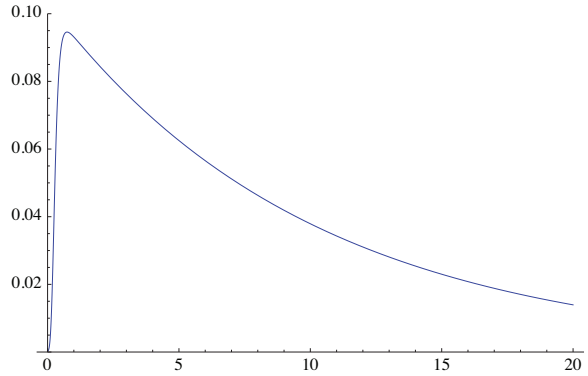

Figure 7: $f(\lambda;74.9,3,3)$, updated after observing three rational actions under type 1 and three rational actions under type 2.

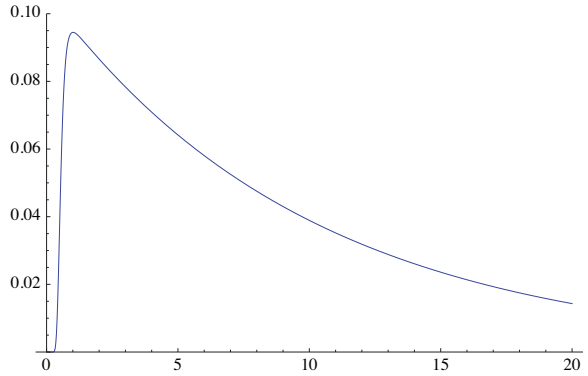

Figure 8: $f(\lambda;749.9,30,30)$, updated after observing 30 rational actions under type 1 and 30 rational actions under type 2.

7 shows the posterior after three observations of action a_3 under type 1 and three observations of a_2 under type 2 ($f(\lambda;74.9,3,3)$); and Figure 8 shows the posterior after 30 observations of action a_3 under type 1 and 30 observations of a_2 under type 2 ($f(\lambda;749.9,30,30)$) . Again higher values of λ become more likely if the agent always chooses the action with the best utility. The probabilities of the actions under the two different types with these two posteriors are shown in Table 2.

Assume that agent j behaves randomly. Within the first six actions, he chooses each of his actions a_1, a_2 and a_3 once under type 1 and each of the possible actions once under type 2 respectively. The updated distribution over i's error parameter is then $f(\lambda;31.9,3,3)$, which is shown in Figure 9. Further, if within j's 60 actions he chooses a_1 for ten times, a_2 ten times, and a_3 for ten times under type 1 and each of the possible actions ten times under type 2 respectively; then the posterior is $f(\lambda;319.9,30,30)$, which is shown in Figure 10. Again we see if agent j behaves randomly, lower values of λ become more likely. Probabilities of the three actions under the two different types with these two posteriors are also shown in Table 2.

5 Sequential Dynamic Environments with Perfect Observability of Continuous Types

Let us consider an even more general case, in which the expected utilities u_{a_l} are not limited to a finite number of values but can lie in some interval or even on the real line:

$$P(a_j|\boldsymbol{u},\lambda)=\frac{e^{\lambda u_{a_j}}}{\sum_{l=1}^{m}e^{\lambda u_{a_l}}}, \tag{9}$$

where $u_l < u_{a_l} < u_l{}', l=1,2,...,m$, $u_l \geq -\infty$ and $u_l{}' \leq \infty$ are lower and upper bounds of the expected utilities u_{a_l}, and where $\boldsymbol{u}$ is a vector of expected utilities of all m actions, $\boldsymbol{u}=(u_{a_1},u_{a_2},...,u_{a_m})$. Again assume u_{a_l} are known to agent i, and he observes agent j's action a_j.

Similarly to Section 4, the Bayesian update equation with continuous types is

$$f(\lambda|a_j,\boldsymbol{u})=\frac{P(a_j|\boldsymbol{u},\lambda)f(\lambda)}{\int_0^\infty P(a_j|\boldsymbol{u},\lambda')f(\lambda')\,d\lambda'}. \tag{10}$$

If we want to update the distribution of λ it would be convenient to find a conjugate prior of (9) for λ. However, forming a conjugate prior in this case is not easy, and may be impossible. The reason is that the construction of conjugate prior distributions (DeGroot 2004; Fink 1997) is based on the existence of sufficient statistics of fixed dimension for the given likelihood function (equation (9) in this case). However, under very weak conditions, the existence of fixed dimensional sufficient statistic restricts the likelihood function to the exponential family of distributions (Barndorff-Nielsen and Pedersen 1968; Fraser 1963). Unfortunately, (9) does not belong to the exponential family with continuous utilities $\boldsymbol{u}$ when $m \geq 2$. (see Appendix B for proof).

In other words, in this case, there is no known way of deriving a family of conjugate priors. Two ways of circumventing this difficulty present themselves. First is to discretize $\boldsymbol{u}$

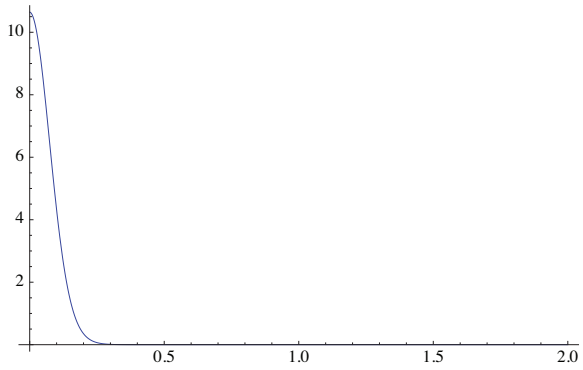

Figure 9: $f(\lambda; 31.9, 3, 3)$, updated after observing three random actions under type 1 and three random actions under type 2.

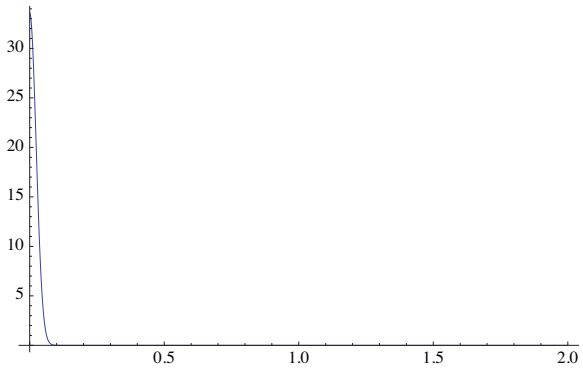

Figure 10: $f(\lambda; 319.9, 30, 30)$, updated after observing 30 random actions under type 1 and 30 random actions under type 2.

and approximate it by fitting its values into a finite number of types. The second one is to give up on conjugate priors altogether and use numerical approximation to update λ.

6 Conclusion

In this paper we postulated that bounded rationality of agents be modeled as noise, or error rate, that perturbs their rational action selection. Since error rates of other agents are not directly observable, we presented ways to learn these parameters during interactions. The learning uses Bayesian update, for which it is convenient to use a family of conjugate priors over the possible values of the error rate. We found the conjugate priors of logit quantal response functions for static and episodic environments, and for sequential dynamic environments with finite number of observable types. The existence of conjugate priors under these assumptions makes the task of learning another agent's error rate simple and tractable. However, we have also shown that if the space of types of the modeled agent is continuous, then the quantal response likelihood does not satisfy the precondition needed for construction of conjugate priors over the error rates. Discretizing their utilities to make continuous types fit into finite pre-specified types can be a way of solving this difficulty. Another method is to abandon the search for conjugate priors and use numerical approximations.

A Proof that the Conjugate Prior for the Static Episodic Environment is a Valid Probability Density Function

Here we prove (3) is a probability density function. First, we prove the proposition: $\int_0^\infty e^{\lambda' u}/(\sum_{l=1}^m e^{\lambda' u_{a_l}})^n \, d\lambda'$ converges if and only if $u < n \max_l u_{a_l}$.

It is trivial that when $n = 0$, $\int_0^\infty e^{\lambda' u} \, d\lambda'$ converges if and only if $u < 0$. So we just need to consider $n \geq 1$. We first prove the sufficiency of the condition. Given $u < n \max_l u_{a_l}$:

$$\begin{aligned}
&\int_0^\infty \frac{e^{\lambda' u}}{(\sum_{l=1}^m e^{\lambda' u_{a_l}})^n} \, d\lambda' \\
=&\int_0^\infty \frac{1}{(\sum_{l=1}^m e^{\lambda'(u_{a_l} - u/n)})^n} \, d\lambda' \\
\leq&\int_0^\infty \frac{1}{(e^{\lambda'(\max_l u_{a_l} - u/n)})^n} \, d\lambda' \\
=&\int_0^\infty e^{\lambda'(u - n \max_l u_{a_l})} \, d\lambda' \\
<&\infty,
\end{aligned}$$

namely $\int_0^\infty e^{\lambda' u}/(\sum_{l=1}^m e^{\lambda' u_{a_l}})^n \, d\lambda'$ converges.

Now prove the necessity of the condition by contradiction. Assume $\int_0^\infty e^{\lambda' u}/(\sum_{l=1}^m e^{\lambda' u_{a_l}})^n \, d\lambda'$ converges, but $u \geq n \max_l u_{a_l}$, then

$$\begin{aligned}
&\int_0^\infty \frac{e^{\lambda' u}}{(\sum_{l=1}^m e^{\lambda' u_{a_l}})^n} \, d\lambda' \\
=&\int_0^\infty \frac{1}{(\sum_{l=1}^m e^{\lambda'(u_{a_l} - u/n)})^n} \, d\lambda' \\
\geq&\int_0^\infty \frac{1}{(m \, e^{\lambda'(\max_l u_{a_l} - u/n)})^n} \, d\lambda' \\
=&m^{-n} \int_0^\infty e^{\lambda'(u - n \max_l u_{a_l})} \, d\lambda' \\
=&\infty.
\end{aligned}$$

This contradicts with our assumption that the integral converges. Therefore $u < n \max_l u_{a_l}$.

Now we have proven that $\int_0^\infty e^{\lambda' u}/(\sum_{l=1}^m e^{\lambda' u_{a_l}})^n \, d\lambda'$ converges given $u < n \max_l u_{a_l}$. Therefore (3) is a probability density function because $\int_0^\infty f(\lambda; u, n) \, d\lambda = 1$.

B Proof of Quantal Response for Continuous Types not Belonging to the Exponential Family

We prove (9) does not belong to the exponential family when $m \geq 2$. Since the factor $e^{\lambda u_{a_j}}$ is trivially in exponential form, we only need to prove the other factor $1/\sum_{l=1}^m e^{\lambda u_{a_l}}$ is not in exponential form (Klauer 1986), which is equivalent

to saying $\sum_{l=1}^{m} e^{\lambda u_{a_l}}$ is not in exponential form. To prove this, we only need to prove its logarithm

$$\ln\left(\sum_{l=1}^{m} e^{\lambda u_{a_l}}\right) \tag{11}$$

as a function of λ contains infinite number of linearly independent functions (Fraser 1963).

Let us assume $m = 2$ for simplicity (the proof for $m > 2$ is the same). Since $u_{a_1} \in (u_1, u_1')$ and $u_{a_2} \in (u_2, u_2')$, there exist countable infinite rational numbers in (u_1, u_1') and (u_2, u_2'). Let us pick up n different rational numbers in each interval for u_{a_1} and u_{a_2}: $u_{a_1} = p_1', p_2', ..., p_n'$ and $u_{a_2} = q_1', q_2', ..., q_n'$. We get n functions from (11):

$$f_1(\lambda) = \ln(e^{p_1'\lambda} + e^{q_1'\lambda}),$$
$$...$$
$$f_n(\lambda) = \ln(e^{p_n'\lambda} + e^{q_n'\lambda}).$$

We now prove $f_1(\lambda), f_2(\lambda), ..., f_n(\lambda)$ are linearly independent. Let $c_1, c_2, ..., c_n \in \mathbb{R}$. Consider

$$c_1 f_1(\lambda) + c_2 f_2(\lambda) + \cdots + c_n f_n(\lambda) = 0. \tag{12}$$

Let L be the lowest common denominator of the fractional forms of $p_1', ..., p_n', q_1', ..., q_n'$, then $p_1 = Lp_1', ..., p_n = Lp_n', q_1 = Lq_1', ..., q_n = Lq_n'$ are integers. Let $x = e^{\frac{\lambda}{L}}$ (note that this is a bijection from $\lambda \geq 0$ to $x \geq 1$). Then we get:

$$c_1 \ln(x^{p_1} + x^{q_1}) + \cdots + c_n \ln(x^{p_n} + x^{q_n}) = 0.$$

Taking its derivative, we get:

$$c_1 \frac{p_1 x^{p_1-1} + q_1 x^{q_1-1}}{x^{p_1} + x^{q_1}} + \cdots + c_n \frac{p_n x^{p_n-1} + q_n x^{q_n-1}}{x^{p_n} + x^{q_n}} = 0.$$

Multiplying both sides by $(x^{p_1} + x^{q_1}) \cdots (x^{p_n} + x^{q_n})$, we will get a new equation with a polynomial on the left hand side whose degree is less than $2(p_1 + q_1 + \cdots + p_n + q_n)$. Therefore the new equation has less than $2(p_1 + q_1 + \cdots + p_n + q_n)$ roots on $\mathbb{R}$, and the number of roots on $[1, +\infty)$ is at most as that on $\mathbb{R}$. We can always choose proper $p_1', ..., p_n', q_1', ..., q_n'$ so that none of $c_1, c_2, ..., c_n$ can be canceled out. In order to make the new equation hold $\forall x \geq 1$, it has to hold that $c_1 = c_2 = \cdots = c_n = 0$.

So far we have proven that (12) implies $c_1 = c_2 = \cdots = c_n = 0$, which means the system $f_1(\lambda), f_2(\lambda), ..., f_n(\lambda)$ are linearly independent. Adding $f_{n+1}(\lambda) = \ln(e^{p_{n+1}'\lambda} + e^{q_{n+1}'\lambda})$ (with proper p_{n+1}' and q_{n+1}') into the system makes a new linearly independent system of $n+1$ functions. By repeating this process, we will get a system of infinite linearly independent functions. Now we have proven (11) contains infinite number of linearly independent functions, which means (9) does not belong to the exponential family.

References

Barndorff-Nielsen, O., and Pedersen, K. 1968. Sufficient data reduction and exponential families. *Math. Scand.* 22:197–202.

Camerer, C. F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press.

DeGroot, M. H. 2004. *Optimal Statistical Decisions (Wiley Classics Library)*. Wiley-Interscience.

Fink, D. 1997. A Compendium of Conjugate Priors. Technical report.

Fraser, D. A. S. 1963. On sufficiency and the exponential family. *Journal of the Royal Statistical Society. Series B (Methodological)* 25(1):115–123.

Gigerenzer, G., and Goldstein, D. G. 1996. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review* 103(4):650–669.

Gigerenzer, G., ed. 2000. *Adaptive thinking: Rationality in the real world*. Oxford University Press.

Kahneman, D., S. P., and Tversky, A. 1982. *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.

Klauer, K. C. 1986. Non-exponential families of distributions. *Metrika* 33:299–305.

McKelvey, R., and Palfrey, T. 1995. Quantal response equilibria for normal form games. *Games and Economic Behavior* 10:6–38.

McKelvey, R., and Palfrey, T. 1998. Quantal response equilibria for extensive form games. *Experimental Economics* 1:9–41.

Rubinstein, A. 1998. *Modeling Bounded Rationality*. MIT Press.

Russell, S., and Norvig, P. 2010. *Artificial Intelligence: A Modern Approach (Third Edition)*. Prentice Hall.

Simon, H. 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* 69:99–118.