Modeling Delay Discounting using Gaussian Process with Active Learning

Jorge Chang[1], Jiseob Kim[2], Byoung-Tak Zhang[2], Mark A. Pitt[1], Jay I. Myung[1]
[1]Department of Psychology, The Ohio State University, Columbus, OH 43210, USA
[2]School of Computer Science and Engineering, Seoul National University, Seoul 151-742, KOREA

{*changcheng.1,pitt.2,myung.1*}*@osu.edu*, {*jkim,btzhang*}*@bi.snu.ac.kr*

Abstract

We explore a nonparametric approach to cognitive modeling. Traditionally, models in cognitive science have been parametric. As such, the model relies on the assumption that the data distribution can be defined by a finite set of parameters. However, there is no guarantee that such an assumption will hold, and it may introduce undesirable biases. For these reasons, a nonparametric approach to model building is appealing. We propose a novel framework that combines Gaussian Processes with active learning (GPAL), and evaluate it in the context of delay discounting (DD), a well-studied task in decision making. We evaluate GPAL in a simulation and a behavioral experiment, and compare it against a traditional parametric model. The results show that GPAL is a suitable modeling framework that is robust, reliable, and efficient, exhibiting high sensitivity to individual differences.

Keywords: Gaussian processes; optimal experimental design; delay discounting; nonparametric modeling; Bayesian inference

Introduction

Models of human cognition are built by designing an explanatory or descriptive model that fits data generated in a behavioral experiment. Although a model's parameterization is motivated by assumptions about the cognitive process under study, the empirical data strongly influence model design. Because of this, the design of the behavioral experiment from which the data were generated (e.g., which stimuli were presented to participants) can introduce bias into the model. This can occur, for example, by not sampling the stimulus space adequately, which can then lead to an incomplete or imprecise model. Two ways to reduce such bias are to not commit in advance to which stimuli should be sampled and to make as few assumptions about the cognitive model as possible, such as parameterization and functional form. In other words, make model-building and data collection as data-driven as possible, at least in the initial stage of model development. Gaussian processes (GP) provide a means of achieving these two goals, functioning as a nonparametric framework for experimentation. We evaluated the viability of a GP-based approach for cognitive modeling in humans.

Researchers in psychology have explored the use of GP to model human behavior (e.g., Cox, Kachergis, & Shiffrin, 2012; Griffiths, Lucas, Williams, & Kalish, 2009; Schulz, Speekenbrink, & Krause, 2018; Song, Sukesan, & Barbour, 2018) but it has yet to be a wide-spread approach. Here, we propose a flexible framework for cognitive modeling by combining GP with active learning (GPAL). GPAL extends traditional GP regression by including appealing features for cognitive science tasks. GPAL is capable of simultaneously modeling the data with minimal assumptions and optimizing the experimental design to find the underlying function efficiently. By virtue of being nonparametric, GPAL shows high sensitivity to individual differences and is able to capture a wider array of patterns compared to parametric approaches. This sensitivity should provide high-fidelity models. Optimization is desirable to minimize the length of a testing session, such as when experimentation is expensive (neuroimaging research) or time-constrained (clinical or special populations). While models produced by the nonparametric framework may not provide interpretable parameters, inferences about cognitive functioning can still be made by examining mathematical properties of the function, such as gradient, curvature or area under the curve.

In our study here, we examined the efficiency, reliability, robustness, and sensitivity of GPAL in the context of modeling delay discounting (DD). Data were collected from 30 participants in a delayed discounting task (e.g., "Do you prefer $10 today or $40 dollars in two weeks?"), which measures an individual's ability to delay gratification. This is a common task in decision-making research, and performance shows a strong correlation with other psychological phenomena, including impulsivity and addiction (Green & Myerson, 2004). The one-parameter hyperbolic model is a popular model that assumes future rewards decline in value hyperbolically with the length of the delay. Recent work from Cavagnaro, Aranovich, McClure, Pitt, and Myung (2016) used adaptive design optimization (ADO) to estimate the parameters of the function in an active learning fashion. One of the conclusions from that study is that none of the six models tested were able to capture the full range of behavioral patterns participants displayed in the task. Thus, DD provides a good test-bed in which to evaluate GPAL. The present investigation represents the first step toward validating GPAL as a premier modeling tool for cognitive science research.

Gaussian Process with Active Learning (GPAL)

This section provides background on each component of the proposed GPAL framework. Figure 1 shows a schematic representation of it.

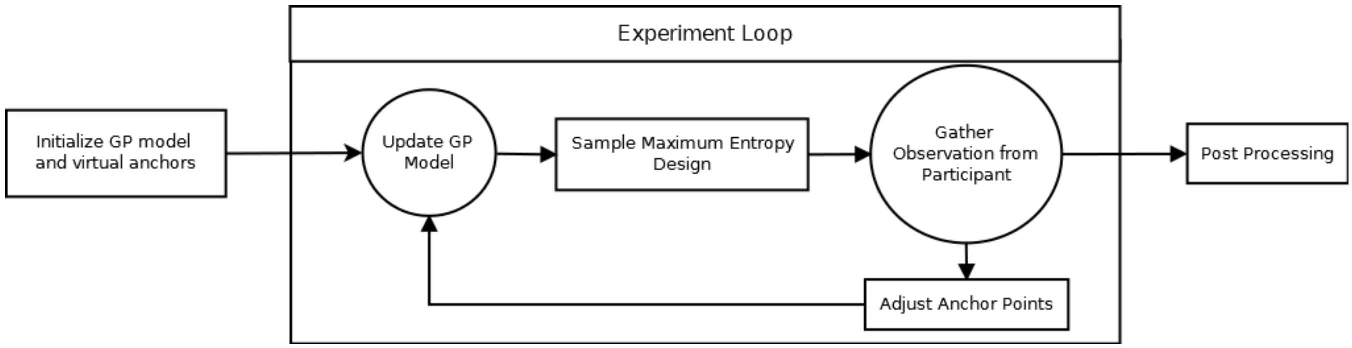

Figure 1: Schematic representation of the GPAL framework. The task is formulated as an active learning based classification task. Virtual anchors are used to restrict the sampling of the design space. On each trial in the experiment loop, an optimal design is picked from the restricted design space according to the maximum entropy criterion, an observation is made, and the GP model and virtual anchors are updated. After the looping, a post processing step may be used to refine the final GP model.

Gaussian Processes

Gaussian processes (GP) are tools for nonparametric Bayesian modeling that establish priors over functions and are a popular approach in machine learning for regression and classification tasks (Rasmussen & Williams, 2006). Formally, GP is a stochastic (random) process where any subset of random variables forms a Gaussian distribution. For a set of observed value pairs (X, f) and a set of unobserved pairs $(\widetilde{X}, \widetilde{f})$, the joint posterior distribution under GP is given by

$$\begin{bmatrix} f \\ \widetilde{f} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu \\ \widetilde{\mu} \end{bmatrix}, \begin{bmatrix} K_{f,f} & K_{f,\widetilde{f}} \\ K_{\widetilde{f},f} & K_{\widetilde{f},\widetilde{f}} \end{bmatrix} \right) \tag{1}$$

where K is a kernel function that defines the covariance between two function values. The kernel function used in this study is the squared exponential kernel which is defined by a length scale parameter l that controls the smoothness and the variance parameter σ^2, which is a measure of the average distance to the mean. This kernel function is a popular choice that has several desirable properties and is known to work well with very smooth functions. The posterior in Eq. 1 can then be used to model $\widetilde{F}$ using the conditional of the multivariate normal distribution.

Many tasks in cognitive science such as DD are not able to observe f directly due to the nature of human experiments. Instead, it is common to give participant choices resulting in multinomial observations. In the case of DD, participants are given two choices on each trial, thus resulting in binomial observations which can then be modeled as a GP binary classification task. This is commonly done by applying a sigmoid transformation function (e.g., probit in our case) to restrict the predicted values to a unit interval. As a consequence of this transformation, the likelihood is no longer Gaussian and requires the use of approximate methods to be estimated, such as expectation propagation as done here. We direct the readers to Rasmussen and Williams (2006), and Vanhatalo et al. (2012) for a comprehensive tutorial on GP and related techniques.

Active Learning

Behavioral experiments can be expensive in terms of both money and time, and the more time an experiment takes, the greater the chance that data quality will suffer due to fatigue or boredom. Systems that incorporate active learning are appealing because they mitigate these problems by optimizing efficiency through identifying highly informative design points based on previous observations (e.g., Cohn, Ghahramani, & Jordan, 1996). It is possible to incorporate active learning in GP based system by deriving a measure of information from the GP and then finding the design point that maximizes this objective function. For our experiment, we used entropy as an information theoretic objective function. We use the derivation of entropy in Houlsby, Huszr, Ghahramani, and Lengyel (2011) which approximates the entropy for GP classification.

Like many tasks in cognitive science, design points in DD are sampled from a discrete space. This space needs to be sparse enough to allow human subjects to make meaningful and differentiable choices. Thus, the search space for optimizing experiments is significantly smaller than in other areas that would use this kind of approach, thereby making grid search a better choice to maximize the entropy function than the proposed method in Houlsby et al. (2011).

Constrained Gaussian Process

Models of natural phenomena are often constrained by prior knowledge or experimental design. For example, when studying natural organisms, the range of the model can be constrained by the physical limitations of such organism. Similarly, a model can be constrained by the experimental design. For example, researchers often design experiments such that some of the outcomes are trivial and well-anticipated. Traditionally, these factors are incorporated in the model design and the range of the parameters and design space. This is a bit more difficult to do in GPAL since it is built to be a fully general modeling tool. It is, however, possible to impose some ”reasonable” constraints on the bounds, derivatives, and convexity of a GP. For a more detailed explanation of these

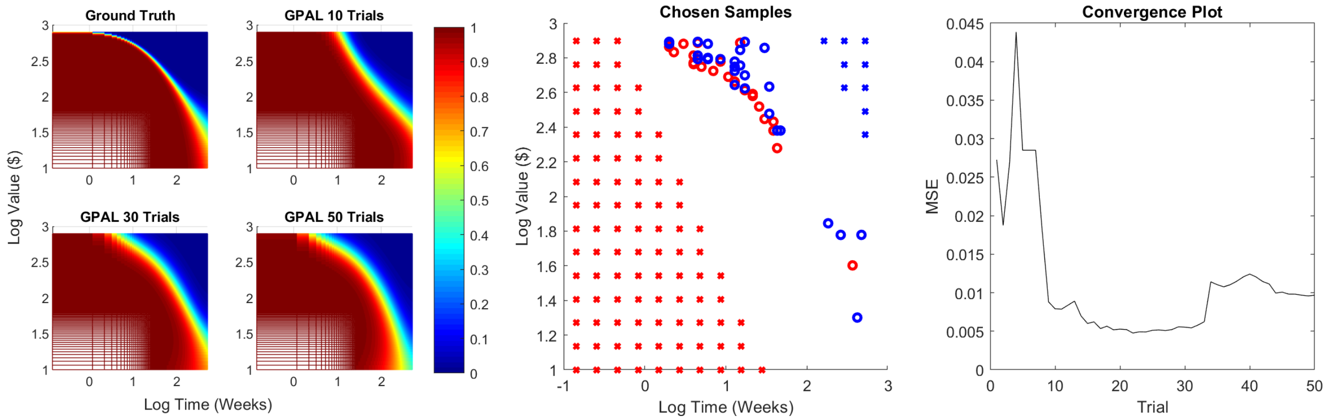

Figure 2: Example of a GPAL simulation using a hyperbolic model as the ground truth. The left panel show GPAL estimated models at four different trials in a top-down view where the z-axis represents the probability of choosing the delayed reward as a function of the two-dimensional, logarithmically scaled design space of (t_{LL}, A_{SS}). The x-shaped and circle-shaped samples in the middle panel represent virtual anchor points and design points selected by GPAL, respectively. The blue and red circles represent choosing the immediate and delayed reward, respectively. The convergence plot in the right panel represents the mean squared error (MSE) between the model at each trial and the ground truth.

methods, see Da Veiga and Marrel (2012). For our delay discounting experiment, we focus on two properties that are often desired in psychological tasks: (a) monotonicity; and (b) local constraints.

Tasks based on subjective preference such as DD often assume that humans follow the axioms of rationality. This leads to researchers building monotonic models that predict a preference for a choice with a higher reward value (or a shorter delay) over another choice with a lower reward value (or a longer delay), if all other things are equal. One way to force monotonicity in GP models is by systematically adding virtual observations in areas where the constraint is violated (Riihimki & Vehtari, 2010). Specifically, this is done by building a joint model of the GP and its derivative. The derivative domain is then used to inject virtual observations into the model when the derivative of the GP violates the monotonicity constraint.

Regarding local constraints, experiments are often designed in a way such that the outcomes are trivial for some design points. For instance, the probability of preferring a choice of \$790 now to another choice of \$800 in 10 years should be virtually equally to 1. Ideally, we would like GPAL to avoid sampling such trivial design points. Again, inspired by Riihimki and Vehtari (2010), GPAL implements local constraints as follows: We can extend their idea of virtual noiseless observations to local constraints by placing them in trivial regions. We refer to these virtual observations as ”virtual anchors”. Virtual anchors act as a prior to the function being estimated by reducing its variance so as to avoid sampling in this region. They also have the benefit of removing the need for initial random sampling to start the active learning process. Further expanding on this idea, we can cover large areas of the design space with virtual anchors and systematically remove them using a moving margin that recedes when a design point is sampled nearby. In our simulations and experiment with human participants described below, we used a linear receding margin, though other schemes could also be used, depending upon the problem at hand. We would like to note that these are preliminary results and future work will focus on increasing the robustness of the model.

Models of Delay Discounting

Delay discounting (DD) is a preferential choice task that is often employed to measure impulsivity by quantifying the preference of an sooner-smaller reward (*SS*) against a later-larger reward (*LL*). This measure of impulsivity has been linked to various mental illnesses such as addiction, gambling, and ADHD (Koffarnus, Jarmolowicz, Mueller, & Bickel, 2013; Sharp et al., 2012; Reynolds, 2007). Models of DD typically start by defining the relation between the value of a reward A at time t as:

$$V = AD_t \tag{2}$$

where V represents the discounted value of A, and D_t the discounting factor. Under this framework, DD behavior is modeled by fitting choice data to a discounting curve that models D_t as a function of t. A popular model of choices is the 1-parameter hyperbolic model (Mazur [1987]):

$$D_t = \frac{1}{1+kt} \tag{3}$$

where $k(>0)$ is the parameter related to impulsivity in that high values of k are associated with high levels of impulsivity. Participant choices are fitted to this model by defining a sigmoid choice function for the probability of choosing the LL option over the SS option:

$$P(LL|k,\varepsilon) = \frac{1}{1+e^{\varepsilon(V_{SS}-V_{LL})}} \tag{4}$$

where V_{SS} and V_{LL} are the discounted values of the SS and LL choice options, respectively, and ε (>0) is a free parameter reflecting consistency of choice behavior. To aid participants in making more meaningful choices and ease visualization, we fix A_{LL} to \$800 and t_{SS} to 0 weeks (i.e., an immediate reward). Thus, the design space becomes a two-dimensional space of (t_{LL}, A_{SS}).

Cavagnaro et al. (2016) extended the hyperbolic model to include active learning by using an adaptive design optimization (ADO) framework. ADO is a parametric framework for Bayesian optimal adaptive experimentation that can be used to select the most informative design for parameter estimation as well as model discrimination (Cavagnaro, Myung, Pitt, & Kujala, 2010; Myung, Cavagnaro, & Pitt, 2013). For our experiments, we use ADO as baseline to compare the performance of GPAL against a parametric model.

Simulations of GPAL

We first tested the feasibility of GPAL in a simulation study in which GPAL was to recover a hyperbolic function used in Cavagnaro et al. (2010) with 10% noise in the observations. Virtual anchor points were incorporated by adding noiseless design points with a value of one or zero at the extreme values of the design space. Figure 2 (left panel) shows an example of the performance with a top-down view of the GPAL estimated DD models at four different trials. Each plot in the left panel represents the probability of choosing the delayed reward as a function of the "later-larger" time t_{LL} and the "sooner-smaller" value A_{SS}. Both dimensions are plotted in the log domain to highlight the difference between functions in our experiments. The decision boundary refers to the regions of interest where the probabilities are closer to 0.5.

The results suggest that GPAL can achieve reasonable convergence within the first 20 trials. Afterwards, as shown in the right panel, we see a decrease in performance (rise in MSE), likely due to Gaussian noise. The design points (red and blue circles) sampled by GPAL at the end of 50 trials, as shown in the middle panel of Figure 2, are reasonable choices that lie close to the decision boundary. The virtual anchors (red and blue x-shaped symbols) provide a reasonable starting point and leave enough distance in case the decision boundary needs to be pushed in one direction or the other. That is, the example shown on the left panel shows that by trial 10 the curvature hasn't been captured yet and the model is still heavily influenced by the anchors. By trial 30, we see a generally close match in shape, with a slight difference in very steep regions as expected due to Gaussian noise.

Modeling Delay Discounting Using GPAL

Experiment

We recruited 30 participants from a pool of undergraduate students at Ohio State University. Participants were asked to perform a DD task over two sessions, ADO and GPAL. The ADO session used ADO to fit the participant choices to the hyperbolic model. The GPAL session fit the data using the GPAL framework with virtual anchors. Both sessions

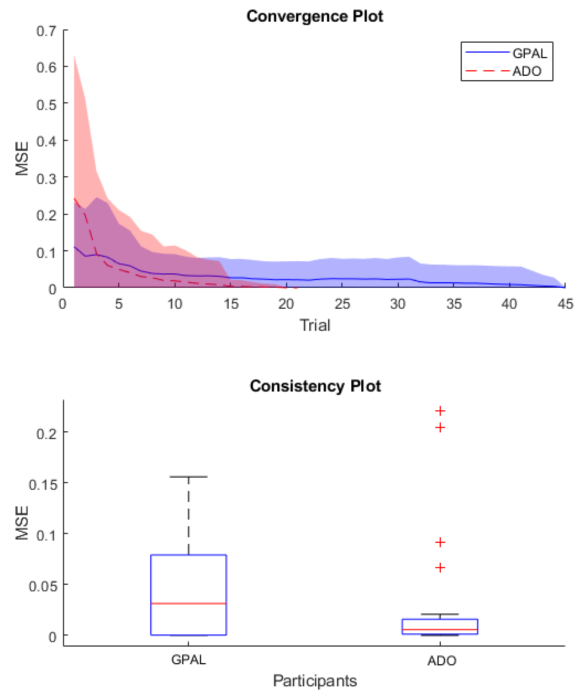

Figure 3: Aggregated results from the experiment. The top panel shows the MSE between the model at each trial and the last. The bottom panel shows the MSE between the first and second session for both experimental conditions.

were further divided into two identical and independent sub-sessions to test for reliability.

The sessions were presented in random order and participants were unaware of the identity of the session they were in. Each trial consists of a preference choice presented in the format "\$X now or \$800 in Y time in the future". The value of X (i.e., A_{SS}) ranged from \$10 to \$790 in multiples of 10 whereas the value of Y (i.e., t_{LL}) took on 48 values ranging from 1 day to 10 years spaced on a logarithmic scale. Each session started with 5 practice trials to familiarize participants with the task. This was followed by 20 trials for ADO and 50 trials for each of the two GPAL sub-sessions, for a total of 120 experimental trials. The two GPAL sessions were presented as a single testing block with no break between them. The number of trials was chosen based on previous ADO experiments and GPAL simulations. All the GPAL software was developed and implemented in MATLAB with the aid of *GPStuff* library for Gaussian processes (Vanhatalo et al., 2012).

Results and Analysis

We tested GPAL on its efficiency, reliability, robustness and sensitivity. Efficiency was assessed by comparing the convergence speed between the two frameworks, ADO and GPAL. We expect efficient models to converge quickly to a final solution. We measured this by the speed at which they approach their final solution. Figure 3 (top panel) shows the MSE at

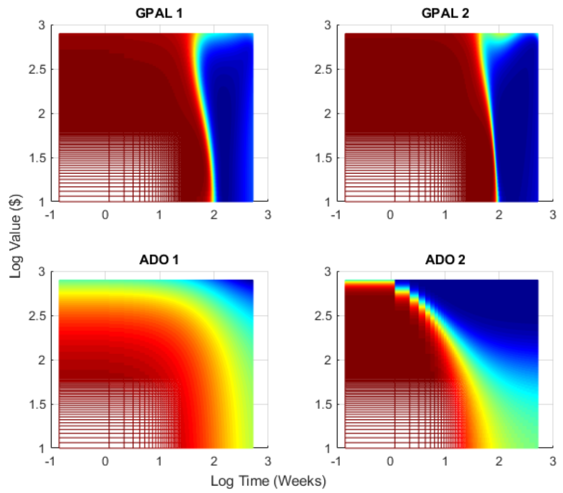

Figure 4: An example participant with consistent results between sessions in the GPAL condition but with inconsistent results in the ADO condition.

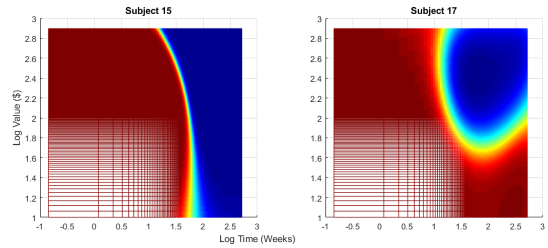

Figure 5: Two selected examples of non-monotonic patterns that consistently emerged during our experiments. The left panel is an example of a non-monotonic pattern with respect to the monetary reward dimension (i.e., y-axis). The right is an example of a non-monotonic pattern with respect to both dimensions.

each trial between the current estimated model and the last model of the experiment. Both models achieve reasonable convergence quickly with ADO flattening out at around 20 trials and GPAL at 30. GPAL starts with the advantage of having access to the virtual anchors, which act as priors. This means that the initial estimate is much closer to the its final solution compared to ADO which is reflected in the initial values of the results. However, ADO shows faster convergence, which can be seen in the rate and consistency at which the MSE decreases. This is an expected result since ADO assumes a hyperbolic model which allows it to make stronger inferences. We also expect efficient models to pick design points that lie close to the decision boundary of each participant, as these represent the most informative points.

One way to assess reliability is by comparing the GPAL function across the two testing sessions. GPAL, if reliable, should produce consistent results across the sessions, and this is what we find. Figure 3 (bottom panel) shows the MSE between sessions for all 30 participants in both conditions. Overall GPAL performance was good, with an average difference of 0.047, which is deemed quite small. For comparison the average MSE between ADO sessions was 0.026. Again, this is expected due to the added flexibility of GPAL.

As seen in Figure 3, the ADO condition had several outliers that were very inconsistent across sessions. Interestingly, these participants were much more consistent in the GPAL condition. Figure 4 shows an example in which the results for the GPAL condition are significantly more consistent than their ADO counterpart. This was the case for all the outliers in the ADO condition. We find that this phenomena tends to happen when GPAL predicts a function shape that is hard for the hyperbolic model to fit in ADO.

Further inspecting the GPAL results, we find that inconsistent samples are largely produced by a shift in the decision boundary in the extremes of the design space. Points in this region are also more influential for GPAL because they determine the concavity of the function whereas for ADO, this is determined by the form of the hyperbolic function. These results also suggest that participants tend to be less consistent for designs in this region. Methods to address this problem will be discussed in the next section.

Regarding robustness, we assessed this property by examining a model's ability to predict unseen data. Operationally, robustness was measured by turning an estimated model, whether ADO or GPAL, into a classifier by setting a decision threshold for the predicted probability to generate predicted outcomes. We then tested classification performance by performing cross validation between the observations of each session. In other words, the GPAL-estimated model was used to predict the designs picked by ADO and the ADO-estimated model was used to predict the designs picked by GPAL. Note that both datasets are comprised of data points that are considered to be hard by their respective framework, making them significantly harder to predict than a random sample. We found that ADO performed literally at the chance level of 49.99% accuracy whereas GPAL achieved a 56.53% accuracy. While this result is not particularly impressive, we take this result as evidence that GPAL is able to produce a better classifier or learn better from noisier data than ADO. This result can also be taken as evidence of higher sensitivity to individual differences, since we expect a sensitive model to produce a better and more robust classifier.

Discussion and Conclusion

How does one build a model of human cognition? We introduced GPAL as a data-driven (bottom-up), nonparametric approach with the aim of overcoming biases in parametric modeling approaches for model development and inference. The diversity of data patterns in our experiments illustrates these features of GPAL. GPAL can uncover concave, convex and approximately linear shapes, and do so quickly, providing the modeler with a higher fidelity description of performance. We envision researchers using this information in one of two ways. The straightforward way is to use GPAL as an exploratory tool for providing an unbiased picture of the raw data to aid the formulation of a parametric model. This gives traditional models a stronger justification in which to ground their assumptions. A second way to utilize GPAL is to replace parametric models altogether. While this second approach requires a paradigm shift in the way models are interpreted, it comes with the potential benefit of providing more accurate measurements. Below we illustrate these ideas by discussing the benefits of GPAL in the context of DD.

Previous models of DD have assumed a monotonic function in both dimensions of money and time. This is a reasonable assumption to make since participants are expected to prefer larger sums of money and shorter time spans. However, Figure 5 shows a few instances that violate this "rationality" assumption. It might be enticing to think that non-monotonic functions in GPAL are a product of noise or model biases. If this is the case, GPAL can be adapted to produce monotonic function using the approach in Riihimki and Vehtari (2010). However, we believe that these non-monotonic patterns are not caused by artifacts. To show this, we focus on the two patterns exemplified in Figure 5. These patterns can be seen across several different participants which we do not show in the interest of space. Since these patterns are repeatable and present across several participants, we find it is unlikely that they are a product of random noise. To support this hypothesis, participants that showed non-monotonic behavior were given five additional trials in which GPAL picked designs using gradient information to identify key regions of non-monotonocity. This process gave a chance for participants to correct their choices to reflect a more conventional function. However, only about a fourth of the participants corrected themselves with the rest confirming their previous behavior. We hypothesize that this non-monotonic, irrational behavior is caused by the interaction between non-linearities in the perceived value of money and delay time. Using the two non-monotonic patterns shown in Figure 5, we provide two possible explanations that would produce this outcome. The first pattern on the left could be caused by a "soft" threshold at which the value of money rapidly decreases, making the time component less relevant. Similarly, the pattern on the right could be caused by a threshold in time at which the delayed reward becomes significantly less appealing. In short, we think that being able to observe these kinds of patterns using GPAL can be a powerful tool to justify choices in parametric models.

A more radical idea is to use GPAL as the primary modeling tool. One of the main benefits of a parametric approach is the ability to formulate theories based on a small set of parameter values. In the case of DD, the k parameter of the hyperbolic model is of particular interest because it is thought to be related to an individual's impulsivity. An analysis such as this is not possible when using a nonparametric model like GPAL since the number of parameters is not constant. However, one could still extract meaningfully information from a GPAL-estimated model. One approach that has been explored in the literature is to interpret the hyperparameters of the kernel function (e.g. Wu, Schulz, Speekenbrink, Nelson, & Meder, 2018). In our case however, the hyperparameters of the square exponential kernel do not relate well to the k parameter. A different alternative could be to use the (parameter-free) estimate of the area under the curve (AUC) of the GPAL function across the input space as an alternative to the k parameter. When this was done for the data from our experiment, the AUC shows a positive correlation to the k parameter. This suggest that the AUC could be used as a measure of impulsivity in a fully nonparametric model but more work needs to be done in this regard. More generally, it is possible to attribute meaning to mathematical properties of GPAL models, which would allow GPAL to function as a primary modeling tool. One benefit of this approach is that the increased sensitivity of the framework might produce more accurate measurements compared to their more constrained counterpart. Additionally, these measurements come from mathematical properties which can be applied to other types of models allowing for easier comparison between models.

Future work will also focus on evaluating the performance of GPAL in a wider array of behavioral tasks. This will allow us to show additional techniques that were not applicable to the DD task. We must also address issues that come from combining GP with active learning. We found that GPAL can be overly sensitive when observations were sparse. While our data suggest that the model is likely to converge within 30 trials, we need to develop the means of ensuring model fidelity while not sacrificing efficiency. The source of this problem is likely due to the greedy nature of active learning. One way to address this problem is to extend active learning to include a bias towards region that are hard for human subjects.

In conclusion, the work in this paper represents a first step towards the development of a novel modeling framework in cognitive science. We propose the use of a nonparametric, model-free approach for cognitive modeling based on GP. This framework serves as a middle ground between raw-data, which are hard to visualize, and parametric models, which rely on strong assumptions. The experiments in the DD task showed that GPAL is a practical framework that yields consistent results efficiently. GPAL showed a high degree of sensitivity to individual differences that were able to uncover non-trivial patterns. This is exemplified by the presence of non-monotonic discounting functions that are present in several participants. These characteristics make GPAL a promising tool for constructing unbiased and sensitive models of cognition.

Acknowledgments

The work was supported by grant FA9550-16-1-0053 from the Air Force Office of Scientific Research.

References

Cavagnaro, D. R., Aranovich, G. J., McClure, S. M., Pitt, M. A., & Myung, J. I. (2016, Jun 01). On the functional form of temporal discounting: An optimized adaptive test. *Journal of Risk and Uncertainty*, *52*(3), 233–254.

Cavagnaro, D. R., Myung, J. I., Pitt, M. A., & Kujala, J. V. (2010). Adaptive design optimization: A mutual information-based approach to model discrimination in cognitive science. *Neural Computation*, *22*(4), 887-905.

Cohn, D. A., Ghahramani, Z., & Jordan, M. I. (1996). Active learning with statistical models. *Journal of Artificial Intelligence Research*, *4*, 129-145.

Cox, G. E., Kachergis, G., & Shiffrin, R. M. (2012). Gaussian process regression for trajectory analysis. *Proceedings of the 34th annual conference of the Cognitive Science Society*, 1440-1445.

Da Veiga, S., & Marrel, A. (2012). Gaussian process modeling with inequality constraints. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, *Ser. 6, 21*(3), 529-555.

Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, *130*(5), 769–792.

Griffiths, T. L., Lucas, C., Williams, J., & Kalish, M. L. (2009). Modeling human function learning with gaussian processes. In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.), *Advances in neural information processing systems 21* (p. 553-560). Curran Associates, Inc.

Houlsby, N., Huszr, F., Ghahramani, Z., & Lengyel, M. (2011, 12). Bayesian active learning for classification and preference learning. *arXiv preprint:1112.5745*.

Koffarnus, M. N., Jarmolowicz, D. P., Mueller, E. T., & Bickel, W. K. (2013). Changing delay discounting in the light of the competing neurobehavioral decision systems theory: a review. *Journal of the Experimental Analysis of Behavior*, *99*(1), 32-57.

Myung, J. I., Cavagnaro, D., & Pitt, M. A. (2013, 06). A tutorial on adaptive design optimization. *Journal of Mathematical Psychology*, *57*, 53-67.

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning (adaptive computation and machine learning)*. The MIT Press.

Reynolds, B. (2007, 01). A review of delay-discounting research with humans: Relations to drug use and gambling. *Behavioural pharmacology*, *17*, 651-67.

Riihimki, J., & Vehtari, A. (2010). Gaussian processes with monotonicity information. In Y. W. Teh & M. Titterington (Eds.), *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (Vol. 9, p. 645-652). PMLR.

Schulz, E., Speekenbrink, M., & Krause, A. (2018). A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, *85*, 1 - 16.

Sharp, C., Barr, G., Ross, D., Bhimani, R., Ha, C., & Vuchinich, R. (2012). Social discounting and externalizing behavior problems in boys. *Journal of Behavioral Decision Making*, *25*(3), 239-247.

Song, X. D., Sukesan, K. A., & Barbour, D. L. (2018, Apr 01). Bayesian active probabilistic classification for psychometric field estimation. *Attention, Perception, & Psychophysics*, *80*(3), 798-812.

Vanhatalo, J., Riihimki, J., Hartikainen, J., Jylanki, P., Tolvanen, V., & Vehtari, A. (2012). Bayesian modeling with gaussian processes using the gpstuff toolbox. *arXiv preprint:1206.5754*.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, *2*, 915-924.