Learning Structured Preferences

Learning Structured Preferences

Leon Bergen[1]**, Owain R. Evans, Joshua B. Tenenbaum**
{bergen, owain, jbt}@mit.edu
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract

Learning the preferences of other people is crucial for predicting future behavior. Both children and adults make inferences about others' preferences from sparse data and in situations where the preferences have complex internal structures. We present a computational model of learning structured preferences which integrates Bayesian inference and utility-based models of preference from economics. We experimentally test this model with adult participants, and compare the model to alternative heuristic models.

Keywords: Theory of Mind; Bayesian Modeling; Preference Learning

Introduction

Children and adults are highly adept at inferring hidden mental states such as preferences from observed behavior. Work in developmental psychology has shown that by 18 months, children have learned that others have preferences which are different from their own; moreover, they can infer these preferences on the basis of observing several choices (Repacholi & Gopnik, 1997). Further work has shown that children's inferences are sensitive to the statistical properties of observed choices, so that choices which are less likely a priori are more likely to reflect strong preferences (Kushnir, Xu, & Wellman, 2008).

Despite the ease with which people appear to reason about others' preferences, such reasoning is a highly challenging computational task. An individual's preferences are never fully determined by their past choices. Moreover, people have to infer preferences on the basis of sparse data, perhaps only a few observed choices.

Our goal in this paper is to develop a computational model of preference learning on the basis of sparse observations. In the spirit of Marr's levels of explanation, we must first spell out the computational problem which has to be solved. In this case, the problem has two parts. First, what exactly has to be learned? Are an agent's preferences simply a ranking of different choices, or do they have more internal structure? Second, how can these preferences be learned on the basis of observing an agent's choices?

Previous Work

Previous computational models of preference learning have mostly focused on the second question. Central to these models is the *principle of rationality*, according to which people are rational agents whose behavior is directed towards satisfying their preferences. Crucially, these models do not claim that the principle of rationality is descriptively true, but rather that it is assumed by people's intuitive theory of mind. These models treat the problem of preference learning as one of inverting the principle of rationality: in order to infer someone's preferences from their behavior, consider what preferences would have made this behavior rational.

In the inverse-planning model of Baker, Tenenbaum, and Saxe (2007), the task is to determine which of several candidate goals an agent desires on the basis of partial movements towards those goals. The agent is assumed to have a ranking of the goals, from best to worst, and which of the goals is desired is inferred by assuming that the agent is moving efficiently towards the desired goal. Using Bayesian inference, the model recovers a distribution over desired goals by supposing that a given goal makes efficient movement toward that goal highly likely. In this paper, there is a sophisticated link between preference and behavior, but preferences themselves have a simple structure.

Lucas et al. (2009) again use a Bayesian model for inference, but use a more sophisticated model for preferences. Instead of having preferences over outcomes, different possible object features have different utilities (are differentially preferred). Preferences over objects are given by a weighted sum over the objects' features. Agents are assumed to choose the object which gives them the maximal quantity of desirable features.

These models were not intended to capture how preferences might be learned when the goods being chosen between have complex functional relationships. Suppose, for example, that someone is choosing between a bicycle frame and bicycle wheels. It will not generally make sense to say that they prefer the frame to the wheels, or vice-versa. Rather, they most likely want an entire bicycle, meaning that whether they want the wheels will depend on whether they already have a suitable frame. Any computational model of this situation will have to account for this additional structure in the agent's preferences.

We combine the Bayesian modeling of Baker et al. and Lucas et al. with formal models of structured preferences developed by economists. We test our model experimentally by examining adult inferences in tasks that provide information about the interactions between goods. By systematically varying these interactions, we try to determine whether preference learning is sensitive to the structure of the preferences being learned.

[1]The first two authors contributed equally to the paper

Modeling Structured Preferences

Economists are interested in modeling agents who are choosing between goods that can have a range of functional relationships. They are also interested in modeling agents over long spans of time, meaning that they have to account for how novel goods may interact with goods already possessed by the agent. In order to model these interactions and effects of context, economists often will not assign utilities to individual goods, but rather to bundles of goods, where the bundle (A, B) consists of A units of one good and B units of another good. Different interactions between goods in a bundle can be captured by different utility functions on these bundles.

We will be considering three kinds of interactions between goods, each of which is standard in economics (see, e.g., Mas-Colell, 1995). The first class of goods are known as *substitutes*. Two goods will be substitutes if either can be used, independent of the other, to accomplish a single goal. For example, two brands of cherry soda are substitutes for each other: if someone wants to drink cherry soda, they can satisfy that goal by drinking either brand. The second class of goods are *complements*. These goods must both be used in a specific proportion to accomplish a particular goal. Bicycle wheels and a bicycle frame must be put together in a ratio of two to one in order to accomplish the goal of building a functional bicycle. The last class of goods will be described by a *Cobb-Douglas* utility function, which is presented below. These goods perform some goal most effectively in a particular ratio, but can still perform the goal to some extent in other ratios as well. If someone wants to have a fun vacation, it may be best to spend some number of days on the beach and some number of days on a cruise, as an entire vacation on the beach might get boring. However, additional days on either the beach or the cruise will still improve the trip.

These interactions can be captured by three different utility functions on bundles of goods. If two goods are substitutes, then we can set:

$$u(A, B) = \alpha \cdot A + (1 - \alpha) \cdot B \tag{1}$$

where A is the quantity of the first good and B is the quantity of the second good. The weight encodes how well each of the goods serves its function. This equation says that we always get the same amount of utility out of an extra unit of the first good, regardless of how much of the second good we have, and vice-versa. This is essentially the utility function used in Lucas et al., with A and B representing the quantities of two different features.

For complements, we set the utilities of the two goods to be:

$$u(A, B) = min(\alpha \cdot A, (1 - \alpha) \cdot B) \tag{2}$$

Parameter α in this case determines the required proportions between the two goods. Once the goods are in their correct proportions, an additional unit of either good alone will not increase utility. In order to increase utility, extra units of both goods have to be added.

The Cobb-Douglas utility function is defined by:

$$u(A, B) = A^{\alpha} \cdot B^{1-\alpha} \tag{3}$$

As in the case of complements, the parameter α determines the optimal ratio of the two goods. Unlike in the case of complements, additional units of a single good can still increase utility without additional units of the other good. Crucially, there is diminishing marginal utility in each good, holding the other good fixed. In other words, each additional unit of a single good increases utility less and less, unless more of the other good is added.

These utility functions are often the first introduced in microeconomics textbooks, both because of their simplicity and because the structures which they encode are so common. However, the space of possible utility functions and possible interactions between goods is enormous (Mas-Colell, 1995). Though we will only be considering these three cases, our model of preference learning, which we describe next, is compatible with any parameterized utility function.

Computational Modeling

Economists have supposed that individual preferences can be represented by the utility functions above and that agents choose bundles with highest utility. By assuming that preferences have this form, predictions can be made about choice in unobserved situations, by computing the relevant utilities.

On our model of human intuitive preference inference, people's intuitive inferences use the same representation of preferences as economists—people are intuitive economists. They can represent different utility function types, corresponding to the way goods interact in a particular situation. For a given utility function type, they can represent different relations between the goods involved—e.g. that a particular ratio of one good to another is optimal. Again following economics, we propose that people employ the principle of rationality, connecting preference to choice.

In order to model reasoning about agents' choices in situations where the objects of choice interact with each other, we integrate the structured utility functions discussed above with Bayesian inference and the principle of rationality. We assume that the observed agent's preferences over a given set of objects is well-described by one of the utility functions discussed above; which of the utility functions is appropriate in a particular situation will depend on both the causal relations between the objects and the agent's goals. Whatever the utility function of the agent, we assume that she is approximately rational and softly maximizes her utility. The probability of a choice given her utility function is therefore given by the Luce-Choice rule:

$$P(c|u) \propto exp(\beta \cdot u(c)) \tag{4}$$

where u is the agent's utility function, c is a choice, and β is a noise parameter that determines how close the agent is to rational. We set β equal to 15 throughout.

We do not attempt to model how an observer might determine, before seeing any of an agent's choices, which utility function the agent is using in a particular situation. Rather, we assume that the observer has a prior distribution P(u) over the possible utility functions which apply in the situation. For example, we will not attempt to model how one might learn that bicycle wheels and bicycle frames are complements. In general, P(u) will depend on the observer's prior knowledge about the relations between the goods in question. The observer is also given a prior P(α) for each utility function u over the parameter alpha of the utility function. Throughout, we assume P(α) follows a Beta(2,2) distribution.

These assumptions allow us to use Bayesian inference to infer an agent's utility function from her choices. Given observed choices C, the posterior over utility functions is given by:

$$P(u|C) \propto P(C|u) \cdot P(u) \quad (5)$$

where u is the utility function with parameter α. This formula inverts the generative model of the agent's behavior captured by her utility function and the Luce-Choice rule (Equation 4). From observed choices C, we can infer the probability of a novel choice c by averaging over the agent's utility functions:

$$P(c|C) = \int P(c|u)P(u|C)du \quad (6)$$

Using the Luce-Choice rule to give the likelihood $P(c|u)$, we recover the mixed multinomial logit model which is used in Lucas et al. as well as many papers in econometrics (McFadden & Train, 2000).

Experiment

In the following experiment we test the quantitative predictions of the structured preference learning model. We designed scenarios in which an agent chose between bundles of goods, with each bundle containing *A* amount of one good and *B* amount of another good. In order to distinguish our model from previous utility-based models (Baker et al., 2007; Lucas et al., 2009) we varied the functional relationship between the goods described in the scenarios in order to match the three utility function types described above. Thus we constructed three separate groups of scenarios. In the Substitutes Group, the two goods were substitutes for each other. Hence, economists would model the preferences of an agent over these goods using a substitutes utility function (Equation 1). In the Complements Group of scenarios, the goods were complements for each other and so preferences over them would be modeled by a complements utility function (Equation 2). In the Cobb-Douglas Group, the goods were related in the way described by Cobb-Douglas utility function (Equation 3).

Our experiment aims to question whether people are intuitive economists. That is, whether they recognize the functional relationship between goods in a particular scenario and model an agent's preferences with a utility function over bundles of the goods that is appropriate to that functional relationship. A key prediction of the theory from economics is that different functional relationships will lead to different patterns of preference across bundles of the same size. For example, once you have two wheels for your bike, additional wheels are worth very little, while after two days at the beach, further days might still be very desirable. To test this, we designed scenarios for which the numerical properties of the bundles could be held fixed while the functional relationship between the goods in the bundles was varied.

Methods

Participants Participants were 480 individuals on Amazon's Mechanical Turk who received a small compensation for their time.

Materials and Procedure As noted above, we designed three groups of scenarios: Substitutes, Complements and Cobb-Douglas. Scenarios differed across the groups in how the goods being acquired related functionally, and hence in which utility function models preferences for the goods in the scenario.

An example stimulus for the Cobb-Douglas group is the following:

> Last year, John and his wife took a one-week vacation in the Caribbean. When John was booking the trip he had the choice between two package deals, which both cost the same amount:
>
> (A) 4 days on the beach and 2 days on a cruise
>
> (B) 3 days on the beach and 4 days on a cruise.
>
> John chose package (B) and thought that was the best choice given his options. This year John and his wife are planning another trip to the Caribbean. John needs to book the trip in advance. He has to decide between two package deals. Both deals cost the same amount:
>
> (C) 5 days on the beach and 5 days on a cruise
>
> (D) 8 days on the beach and 2 days on a cruise.
>
> Which option should John take?

As this example illustrates, the scenarios consisted of three parts: (1) setup of the first choice situation, (2) the agent chooses one bundle of goods over another, and (3) the agent faces a new choice situation. Each scenario was shownin 16 numerical conditions: these were the quantities of goods in each bundle in the two choice situations. In the example above, the bundle (3,4) was chosen over the bundle (4,2), and participants were asked to choose between the bundles (5,5) and (8,2). In different numerical conditions, the quantities in these bundles were varied. In pre-testing, we found that participants were making inferences about the cost of each bundle on the basis of bundle size; as a result, we subsequently stated in each scenario that the bundles being chosen from were the same price. Because this was incompatible with one of the scenarios, it was removed from subsequent testing. This left a total of 11 scenarios: three in the Cobb-Douglas Group, four in the Substitutes Group, and four in the Complements Group. Pre-testing also suggested possible order effects in some of the numerical conditions, which were controlled in subsequent experiments. The scenarios were crossed with the numerical conditions, for a total of 11x16 inference questions.

Participants each received five or six distinct scenarios; numerical conditions were randomized across participants.

If people are sensitive to the structure of the agents' utility functions, then identical numerical conditions should give rise to different patterns of inference, depending on the scenario group. For example, if the goods are complements, then the observer should be uncertain about the optimal proportion of one good to the other. They will take the agent's initial choice as evidence about the optimal proportion, and will subsequently choose the bundle which fits this inferred proportion as closely as possible. On the other hand, if the goods are substitutes, then they will take the initial choice as evidence about how well each good serves its function, and will subsequently choose the bundle which maximizes the weighted sum of the two goods. If both bundles have the same size, then they will choose the one with the most of the better good.

Model Predictions

In order to make model predictions, we used the multinomial logit model (Equation 6) to find the probability of taking option (C) or (D) conditional on the agent's first choice. This probability was approximated through MCMC simulations, using the Metropolis-Hastings method.

We computed these probabilities under a variety of settings for the noise parameter β in the soft-max equation (Equation 4) and for the prior distribution Beta(α, α) over the utility functions' weight parameter. The noise parameter was set to 15 and α was set to 2 based on the fit of the model to the data over all conditions. Figure 1 shows the sensitivity of the model's predictions to the value of β. The mean squared error of the model varied between 0.04 and 0.1 as β varied between 0.1 and 20. We similarly analyzed the model's sensitivity to the value of α. We found that the mean squared error of the model varied between 0.04 and 0.1 as α varied between 1 and 10.

Our computational model predicts agent preferences by matching the appropriate utility function to the scenario. Thus, e.g. scenarios in the Complements group will be modeled with very strong prior probability on the Complements utility function and likewise for the other two groups.

In order to assess the distinctive predictions of our model, we implemented three heuristic models of preference inference. These heuristics implement rules for making preference inferences that are not based on calculations of utility. The heuristics worked as follows. The Sophisticated Ratio Heuristic was a model of ratio-matching, in which goods are selected based on how close their ratio is to an optimal ratio. The heuristic considers a range of possible optimal ratios and uses Bayesian inference to integrate over these ratios in making predictions. This differs from the complements function in being insensitive to absolute quantities of the goods. The Crude Ratio Heuristic was similar to the Sophisticated Ratio Heuristic, except that instead of considering a range of possible ideal ratios, it treats the ratio of goods in the observed choice as ideal. The Max Heuristic assumes that one of the two goods is preferred, and that agents will always choose the bundle that contains the greatest quantity of the preferred good. This differs from the substitutes utility function in being insensitive to the quantity of the non-preferred good. These three heuristics do not assign utilities to bundles but instead simply select one outcome as the best.

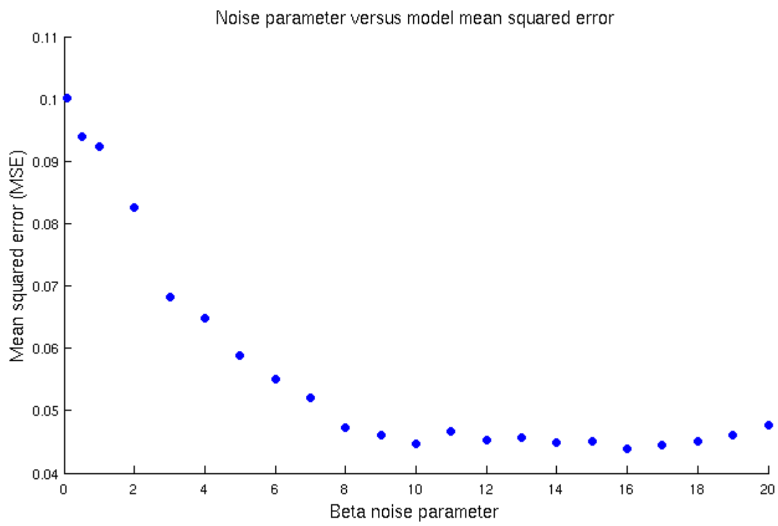
Table 1: Correlations between the three different utility models and subject judgments. Scenarios 1-3 were intended to be best fit by a Cobb-Douglas utility function; scenarios 4-7 by a Substitutes utility function; and scenarios 8-11 by a Complements utility function.

Scenario	Cobb	Substitutes	Complements
1	**0.79**	0.83	0.50
2	**0.62**	0.72	0.30
3	**0.80**	0.70	0.80
4	0.85	**0.95**	0.45
5	0.69	**0.83**	0.32
6	0.87	**0.91**	0.49
7	0.81	**0.93**	0.36
8	0.64	0.34	**0.87**
9	0.63	0.41	**0.86**
10	0.78	0.74	**0.69**
11	0.46	0.23	**0.87**

Table 2: Correlations between the best heuristic models and participant judgments.

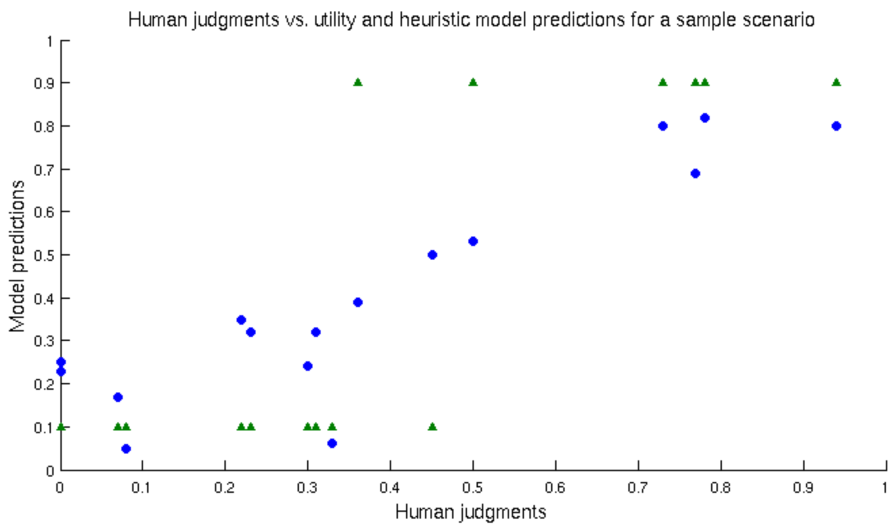
Scenario	Crude Ratio	Max
1	0.55	0.89
2	0.57	0.73
3	0.91	0.76
4	0.48	0.95
5	0.37	0.76
6	0.61	0.93
7	0.55	0.96
8	0.82	0.44
9	0.94	0.49
10	0.70	0.80
11	0.87	0.26

Results. We first consider the model fit when all of the weight of the prior probability distribution is placed on a particular utility function type. As noted above, scenarios were grouped based on which utility function we believed would best fit the goods in the scenario. If participants are sensitive to the functional relationship between the goods in each scenario, then their judgments should be best fit when a high prior probability is placed on the group's utility function



M

Figure 1: Sensitivity of model performance to the value of the noise parameter β.



M

Figure 2: Human judgments vs. utility model predictions (blue dots) and heuristic predictions (green triangles) for a complements scenario. Other scenarios showed a similar relationship between human judgments and the predictions of the two models.

type. In all of the Substitutes scenarios and three of the four Complements scenarios, the human data was best fit by the scenario's corresponding utility function type. Participants' judgments in two of the three Cobb-Douglas scenarios are best fit by the substitutes utility function. These results are shown in Table 1.

We next tested the fit of our heuristic models. The free parameters of the heuristics were chosen to maximize overall correlation with the human data. The Crude Ratio and Max heuristics had the highest correlations, so we restrict our attention to them. The Max Heuristic had high (above 0.7) correlations with human judgments on all of the Cobb-Douglas and Substitutes scenarios. The Crude Ratio Heuristic had similarly high correlations on all of the Complements scenarios. These results are shown in Table 2.

Discussion

Our experiment aimed to investigate two questions. The first was whether people are sensitive to the functional relationships between goods when they make inferences about people's preferences. The second was, if people are sensitive to these functional relationships, what do they learn about others' preferences from their observations? Our results provide evidence that participants were sensitive to the functional relationships between goods. In particular, participants appeared to distinguish between the Complements scenarios and the other scenarios, although we did not find evidence that they distinguish between the Cobb-Douglas and Substitutes scenarios. This is indicated by the fit of both the utility and heuristic models. A single utility model (with all of its prior weight on the substitutes utility function) and a single heuristic (the Max Heuristic) correlate well with human judgments in the Cobb-Douglas and Substitutes scenarios. On the other hand, these two models fare poorly on most of the Complements scenarios. The complements utility model and the Crude Ratio Heuristic provide better fits for these scenarios.

The performance of the Max Heuristic on a particular scenario is very well predicted by the performance of the substitutes utility function on that scenario; the same is true of the Crude Ratio Heuristic and the complements utility function. The Max Heuristic naively tracks a relationship which is important to the substitutes utility function, namely which good is favored over the other. The Crude Ratio Heuristic gives a noisy estimate of the optimal ratio between goods, which is similarly important to the complements utility function. This provides evidence that participants were sensitive to the optimal ratio of goods in the Complements scenarios, and to which was the preferred good in the other scenarios.

It is an interesting question whether the utility-based or heuristic models provides a more promising approach to modeling preference learning. The best heuristic models were able to capture the qualitative shape of participants' judgments. The heuristics were able to correctly classify bundles as more or less likely to be chosen. A representative scenario is shown in Figure 2. On the other hand, human judgments showed a gradedness that was better captured by the utility models. As predicted by the utility models, participants varied across numerical conditions in how strongly they thought one bundle should be chosen over another. This is also shown in Figure 2.

Conclusion

We have presented a computational model of structured preference learning. Our model incorporates Bayesian inference with economic models of internally complex preferences. Using a standard tool from econometrics, it can be used to model how individuals infer what someone prefers on the basis of past observations. We experimentally tested the model, and found that participants were sensitive to the functional relationships between goods in some of the ways that the model predicts. Future experiments could help to differentiate the utility-based and heuristic models. These models will make distinct predictions, for example, when the sizes of the bundles being chosen between are different. Bigger bundles will, all things being equal, be preferred by the utility-based models, while this is not always true for the heuristics.

Further work needs to be done in systematically varying the structure of the preferences being learned. This may be done by explicitly varying the causal structures of the choice scenarios so that subjects' prior beliefs about these relations do not come into play. This suggests a further line of research: studying whether individuals can make use of more complex functional relationships than the ones that were considered here. If individuals can effectively learn others' preferences in a range of situations, then it is unlikely that these preferences will be captured by the simple functional forms studied here.

References

Baker, C. L., Tenenbaum, J. B., Saxe, R. R. (2007). Goal inference as inverse planning. In *Proceedings of the twenty ninth annual conference of the Cognitive Science Society.*

Kushnir, T., Xu, F., Wellman, H. (2008). Preschoolers use sampling information to infer the preferences of others. In *Proceedings of the thirtieth annual conference of the Cognitive Science Society.*

Lucas, C., Griffiths, T. L., Xu, F., Fawcett, C. (2009). A rational model of preference learning and choice prediction by children.In *Advances in Neural Information Processing Systems* 21.

Mas-Colell, A., Whinston, M. D., Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press.

McFadden, D. Train, K. E. (2000). Mixed MNL models of discrete response. *Journal of Applied Econometrics*, 15, 447-470.

Repacholi, B. M., Gopnik, A. (1997) Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, 33(1), 12-21.