UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Generalizing Functions in Sparse Domains

Permalink

https://escholarship.org/uc/item/67f9t5dk

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 41(0)

Authors

Le´on-Villagr´a, Pablo

Lucas, Christopher G.

Publication Date

2019

Peer reviewed

Generalizing Functions in Sparse Domains

Pablo León-Villagrá   Christopher G. Lucas
School of Informatics, University of Edinburgh, United Kingdom

Abstract

We propose that when humans learn sets of relationships they are able to learn the abstract structure or *type* of a family of relationships, and exploit that knowledge to improve their ability to learn and generalize in the future, especially in the face of sparse or ambiguous data. In two experiments we found that participants choose patterns and extrapolate in ways consistent with sets of previously learned relations, as measured by extrapolation judgments and forced-choice tasks. We take these results to suggest that humans can detect shared abstract relations and apply this learned regularity to perform rapid and flexible generalization.

Keywords: generalization, function learning, transfer

Introduction

Many everyday situations require us to generalize from past experience, even if we are faced with a specific problem we have never seen before. For example, in cooking, one regularly has to infer the relationship between ingredients, ratios or quantities, like the amount of sweetener and resulting pleasantness of a dessert, and generalize this relation to new recipes or ingredients. Often, we learn a general relationship that helps us understand related problems. If we learn that as we increase the amount of sugar in a recipe, the sweetness will not change immediately, then increases rapidly and then saturates, one can use this knowledge to reason about similar relationships, as when deciding how much xylitol to add to a cake.

Our example requires two types of generalization. The first, sometimes called transfer or transfer learning, involves transferring information about a relationship between two quantities to help us understand a new and different relationship. The second, extrapolation, involves understanding a single relationship and extrapolating to new instances or data points within, e.g., to new amounts of xylitol. The latter depends on the former – our past experiences shape the inductive biases we bring to a new problem.

Transfer learning expands the task the human learner faces and requires further-reaching and more abstract inferences. Given a set of prediction tasks, how can we capitalize on statistical regularities to aid future prediction? If the tasks exhibit some shared structure, learning a representation capturing this latent structure of the environment (Gershman & Niv, 2010), or learning which aspects of a task change (R. C. Wilson & Niv, 2012) can enable the learner to perform wide-ranging and data-efficient generalization.

The value of transferring knowledge across different tasks is receiving growing attention in machine learning communities. For example, abstract learning and transfer have been successfully applied to challenging control tasks (Hamrick et al., 2017). From a cognitive science perspective, the study of such general learning mechanisms has a long tradition, e.g.,

Harlow (1949). Research in this tradition has highlighted how hierarchical representations can allow for the "blessing of abstraction" (Gershman, 2017), where abstract knowledge is acquired faster than detailed information. In recent years several proposals have been put forward on how hierarchical and structured inductive biases can be acquired through development and how they allow for rapid generalization (Goodman et al., 2008; Tenenbaum et al., 2011).

The second type of generalization has been widely studied in psychology, most commonly in classification tasks in which participants have to learn to predict class labels for unknown objects or entities. Similarly, tasks in which the target to be learned is a continuous quantity have been studied in the domain of function learning research. Research in function learning has emphasized particular human inductive biases. For example, humans learn functions more quickly if the relationship is linear (Brehmer, 1976), and struggle with cyclic functions (Bott & Heit, 2004; Kalish, 2013). More importantly, human extrapolations are strongly biased towards linear relationships, in particular positive linear functions (Brehmer, 1976; DeLosh et al., 1997; Busemeyer et al., 1997; McDaniel & Busemeyer, 2005; Kalish et al., 2004). While this line of research emphasizes simple types of functions, results from experiments with less taxing memory demands have shown that a wide variety of relationships can be learned and inform extrapolation (Lucas et al., 2015; A. G. Wilson et al., 2015; Schulz et al., 2017; León-Villagrá et al., 2018).

In function learning, the hierarchical and abstract representation of the learned relationships has traditionally been reduced to mechanisms that allow generalizing a mapping from criterion to targets. Multiple proposals have been put forward for the nature of these mappings, ranging from rule–like parametric forms (Carroll, 1963; Brehmer, 1976), associative, neuronal network architectures, and hybrids thereof (Busemeyer et al., 1997).

Here we will adopt a general perspective and express the task as Gaussian process regression. A Gaussian process specifies a distribution over functions $f(x) \sim GP(\mu, k)$, where $\mu(x) = E[f(x)]$ and k is the covariance kernel. The kernel specifies a similarity measure over x and allows us to express abstract beliefs about the shape of the function, such as periodicity or smoothness. Gaussian processes have been successful in accounting for both the flexibility in learning, as well as long-range extrapolations (Lucas et al., 2015).

While Gaussian processes allow us to express inductive biases for functions in flexible, non-parametric fashion, only recently more attention has been given to structural and hierarchical aspects of function generalization. This work has emphasized the importance of inductive biases over different

function types (Lucas et al., 2015), the compositional structure of functions (Schulz et al., 2017), or the generalization of functions into dimensions outside the learned space (Lucas et al., 2012).

Here we expand on this line of research and propose that when humans learn relationships they do not maintain sets of data, parametrizations or fixed parametric forms, but that they form flexible and abstract hypothesis spaces. Based on this abstract encoding, we suggest, they are able to capitalize on statistical co-occurrences of abstract information about the *type* of relationship learned. As a result, repeated exposure to similar functions should result in learning about the shared type of relationship, as well as its relevant features. Such exposure should then facilitate extrapolation in sparse contexts and allow far-ranging generalization. We hypothesize that this application of past knowledge does not simply amount to remembering previous data, but extrapolation depends on the induced function type and adapted to the context at hand.

Experiment 1

In this first experiment, we examine if participants prefer functions consistent with the previously learned function type and its shared, defining, features. We train participants on three sets of samples from the same type of function and assess if they subsequently choose extrapolations in concordance with this type and parameters.

Participants

The study was self-certified in accordance with the School of Informatics Ethics Guidelines. We recruited 99 participants ($M_{age} = 32.1$, $SD_{age} = 10.87$, 34 female, 65 male) on Amazon Mechanical Turk. Participants had to have more than 50 approved HITs and an approval rate of 95% or larger. They received \$0.55 for participation and took an average of 7 minutes ($M = 6.46$, $SD = 5.19$) to complete the experiment. Participants were randomly assigned to one of the six conditions ($n_{Cos_1} = n_{Lin_2} = n_{Ou_1} = 17$, $n_{Cos_2} = n_{Lin_1} = n_{Ou_2} = 16$).

Procedure

Participants were instructed that they would learn the relationship between two substances, substance *x*, and substance *y*. They were told that they would be presented with three sets of patterns, each depicting one realization of the same relationship and that they would have to predict the relationship for 10 new points. They also received a visual depiction explaining how they would predict the points. They were instructed that they would see one more pattern from the same relationship, consisting of three points. Then they were instructed to select the pattern from six options that most likely depicted the learned relationship.

Training Phase Each training block took the form of an extrapolation task, where participants saw scatterplots and had to guess the value of the substance on the *y*-axis in an extrapolation range, by selecting the height of the corresponding

value on the plot. Participants were shown the correct value as feedback for one second, and, if their choice deviated by ± 0.025 or more of the true value, had to readjust their selection. Training blocks were presented in randomized order.

Choice Phase After the training blocks, there was a forced-choice task where participants saw the three-point pattern and read that this pattern belonged to the same relationship as the training. Then they saw with six scatterplot patterns, corresponding to one conditional sample for each of the six kernels, in randomized order. Participants had to select the pattern that they deemed the most likely extrapolation for the learned relationship. After the choice tasks, participants completed a short demographic survey.

Materials

The functions in the six conditions corresponded to samples from Gaussian Processes (GPs), with three different types of kernels and mean functions, each with two distinct parametrizations, see Table 1. To allow for characteristic periodic samples, we elected a "pure" cosine kernel, *cos* with $k(r) = \sigma \times cos(r)$, $r(x,x') = \frac{(x-x')^2}{\ell_q^2}$, with an additional intercept. We generated linear samples from a linear kernel *lin* with explicit slope and intercept terms. Finally, we used a Ornstein-Uhlenbeck kernel (*OU*) with an additional intercept, to generate non-smooth samples. The noise variance was fixed to 0.01 for all GPs.

Training Sets We generated the training data by sampling three sets of 35 points each in the range 0.05–0.95 for each of the six conditions. The first 25 points constituted the evidence provided in each training set. Participants had to extrapolate the target value for the last 10 points and received feedback for their choices. To ensure that samples were clearly perceptible and the samples were distinct (within function type and between function types) we generated a set of 20 candidate patterns for the 18 sets. We then selected samples from these candidates for which all points were in the presentation range $[0,1]$, that were ≥ 0.05 of the three transfer points, and rejected uncharacteristic samples[1]. For a full list of kernels and kernel parametrizations, see Table 1, for the training data and the conditional samples, see Figure 1.

Forced Choices In the transfer set three points, $x = \{0.05, 0.1, 0.2\}$, $y = \{0.475, 0.525, 0.5\}$ were. These points were selected to be inconsistent with any of the training materials, in terms of specific point locations. We then generated three samples conditional on the transfer points for each of the six functions. Participants received one of these three samples at random for each of the six kernels in the forced choice task.

---

[1]For example, *OU* samples that did not exhibit any discontinuities and thus looked visually identical to linear relationships, or *cos* samples that had very low amplitudes.
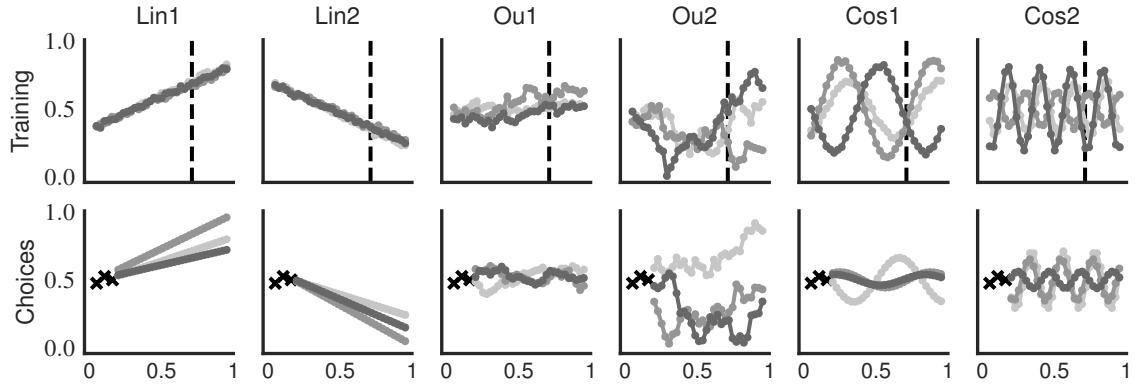
Figure 1: Top row: Training data in the six conditions. For each condition, there were three data sets to be learned. Participants received the first 25 points and had to extrapolate for the 10 remaining points, dashed is the cutoff between presented evidence and training. Bottom row: Samples that constituted the forced-choice options in Experiment 1 & 2. Samples were generated by the particular GP, conditional on the three points in the transfer set.

Table 1: Kernels and kernel parameters generating the training data. For all models we set $\sigma_{noise} = 0.01$.

Kernel	variance	lengthscale	β_0	β_1
Lin_1	0.02	–	0.35	0.47
Lin_2	0.02	–	0.7	-0.47
Cos_1	0.05	0.1	0.5	–
Cos_2	0.05	0.04	0.5	–
OU_1	0.01	1	0.5	–
OU_2	0.08	1	0.5	–

Table 2: MAEs for functions and blocks in Experiment 1

	MAE_{b1}	SD_{b1}	MAE_{b2}	SD_{b2}	MAE_{b3}	SD_{b3}
Lin	.02	.01	.02	> 0.01	.02	> 0.01
OU	.05	.02	.06	.03	.05	.03
Cos	.06	.05	.06	.06	.05	.04

Results

Error Rates

The training functions differed considerably in their mean absolute errors (MAEs)[2], as well in the change of error over blocks, see Table 2.

Only for conditions with linear functions did errors differ significantly between the first and the last block, $t(42.94) = 2.21$, $p = .032$[3]. For OU conditions, errors were lower for the last block, but did not differ significantly, $t(62.21) = 0.95, p = 0.345$. For periodic conditions, error was again lower for the last block, but blocks did not differ significantly $t(57.17) = 0.66$, $p = .509$. The two OU and periodic conditions were highly heterogeneous. While errors for the low variance condition Ou_1 decreased over blocks, errors in Ou_2 remained high. Equally, while errors in Cos_1 decreased, errors for Cos_2 remained high throughout training. For error rates for all conditions, see Figure 2.
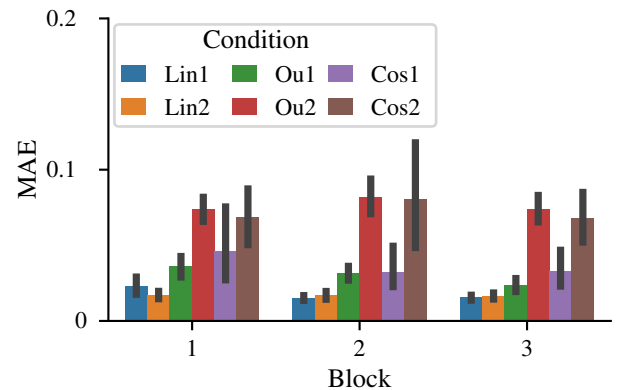

Figure 2: MAEs and 95% confidence intervals for each condition in Experiment 1.

Selecting an Extrapolation Pattern

About 35% of the participants selected the choice corresponding to the correct function type and parametrization. Both for positive and negative linear training conditions, the proportion of chosen true functions was significantly larger than chance (1/6), $Lin_1 = 44\%$, $p = .01$, $Lin_2 = 53\%$, $p < .001$[4]. For periodic functions, Cos_2 was selected significantly above chance, $Cos_2 = 50\%$, $p = .002$, but Cos_1 was not, $Cos_1 = 12\%$, $p = .802$. Instead, participants mostly selected the

---

[2]All MAEs were calculated on extrapolations before the participant had received feedback for that particular value.

[3]All tests are two-sided, unequal variance t-tests. For means and SDs, see Table 2

[4]All tests are one-sided, exact Binomial tests.

other periodic function. The proportion of generally periodic functions over the alternatives for condition Cos_1 was significantly above chance (1/3), $Cos_1 = 59\%$, $p < .027$.

For OU conditions, Ou_1 was not selected significantly above chance, $Ou_1 = 18\%$, $p = .556$, nor did participants prefer OU functions in general, $Ou_1 = 47\%$, $p = .172$.

However, participants trained on Ou_2 selected the true pattern at rates significantly higher than chance $Ou_2 = 38\%$, $p = .05$. For all participant choices, see Figure 5.
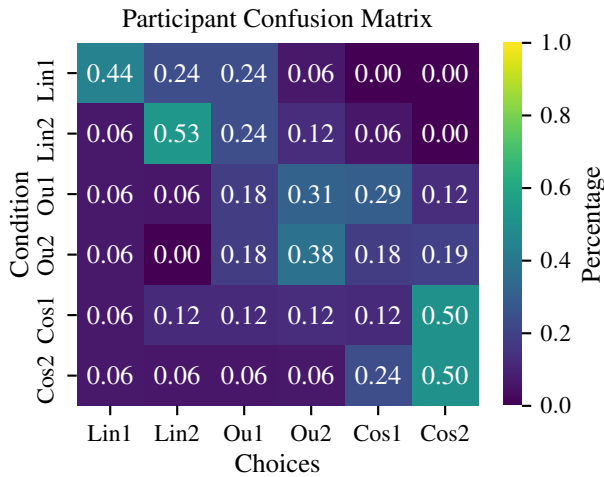

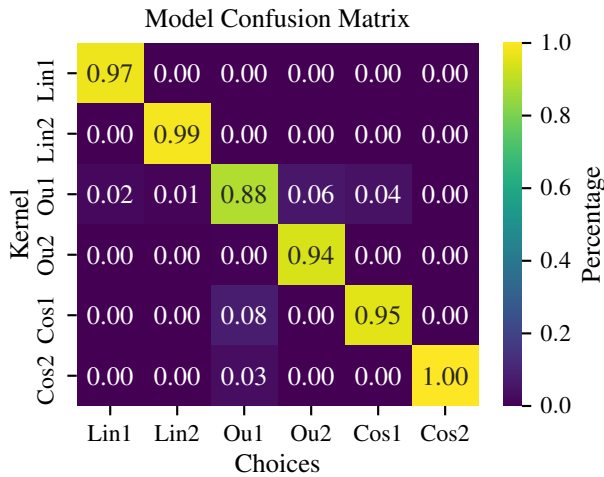Figure 3: Confusion matrix for choices in Experiment 1

Figure 4: Confusion matrix for the generating functions and the choices presented in Experiment 1.

We also calculated the confusion matrix resulting from the likelihood assigned to each of the presented samples for each of the functions generating the materials. We then converted these likelihoods into proportions via the softmax function[5].

_____
[5]We also evaluated these proportions for a model with an addi-

While the model consistently favors the true generating function and does not produce preferences resembling the participants' choices, some interesting parallels are apparent. First, both kernels resulting in the strongest preference for the true function (Cos_2 and Lin_2) also correspond to conditions with fairly peaked human preferences. In contrast, kernels resulting in more dispersed likelihoods, and as a result, lower preference for the true function (Ou_1, Ou_2) resemble the systematic preferences for alternative functions by human participants. For a confusion matrix displaying the asymmetric choices of participants, see Figure 3, for model confusions, see Figure 4.

Experiment 2

In Experiment 1, participants were able to select from a set of candidates realizations corresponding to the learned type of function and, in many cases, the specific features of the set of training examples. In this control experiment, we contrasted participants' choices in Experiment 1 with a condition in which no training was provided.

Participants

We recruited 50 participants ($M_{age} = 34.7$, $SD_{age} = 10.53$, 25 female, 24 male, 1 other) on Amazon Mechanical Turk. Participants received $0.2 for participation and took an average of 1.5 minutes, ($M = 1.46$, $SD = 6.05$) to complete the experiment.

Procedure

Participants were instructed that they would be presented with a relationship between two substances, consisting of three pairs of values. Then they were instructed that they would have to select a pattern from six options that most likely depicted the relationship. The choices were the same as in Experiment 1.

Results

In the absence of training data participants preferred periodic functions over OU and linear, $Lin_1 = 10\%$, $Lin_2 = 0\%$, $Ou_1 = 18\%$, $Ou_2 = 14\%$, $Cos_1 = 28\%$, $Cos_2 = 30\%$, see also Figure 1. Given the low rates of choices of Lin_1 and the high proportion of chosen periodic functions, these results suggest that participants interpreted the three points presented as generated from a deterministic, non-monotonic relationship, rather than a low-noise linear or low-variance OU relationship.

These results suggest that the strong preference for periodic samples in Experiment 1 did not solely result from the training but were also reflective of a higher preference to ascribe periodicity to the test points.

Experiment 3

We have shown that participants can use the knowledge acquired in the training sets, to inform their choices about which

_____
tional temperature parameter T that we fitted to the human choices. Unsurprisingly, this temperature parameter was estimated to be low, $T \approx 10$, and produce less peaked distributions.
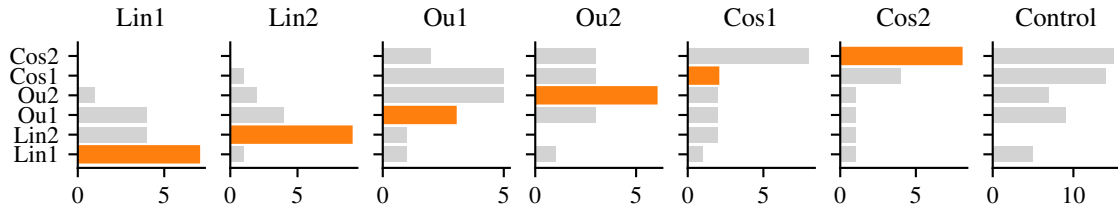
Figure 5: The selected choices (correct choice in orange) for Experiment 1 and Experiment 2 (Control). Participants generally selected the function type and parametrization consistent with their training. When instead they preferred other options they mostly selected samples from the same type of function. Without training, participants favored periodic and OU.

particular type of function generated the data. However, it is possible that these choices do not reflect participant's true belief about the underlying functions, but are merely best guesses given a set of unsatisfactory options. In this last experiment, we will analyze if these choices correspond to actual extrapolation behavior.

Participants

We recruited 91 participants ($M_{age} = 30.53$, $SD_{age} = 6.941$, 34 female, 57 male) on Amazon Mechanical Turk. Participants received \$0.65 for participation and took an average of 10 minutes (M=10.29, SD=10.56) to complete the experiment. Participants were randomly assigned to one of the 6 conditions ($n_{Cos1} = n_{Cos2} = n_{Lin1} = 15$, $n_{Lin2} = n_{OU1} = 16$, $n_{OU2} = 14$).

Procedure & Materials

Instructions and training were identical to Experiment 1. However, instead of the forced-choice task participants performed an extrapolation task. In the extrapolation task, participants received the same three points that generated the conditional samples in experiment 1 and had to extrapolate for 30 values of *x*, without feedback, following the same procedure as in the training sets. The 30 extrapolation criteria were the same as the ones used to generate the forced-choice patterns in Experiment 1.

Results

Error Rates

As in Experiment 1, conditions differed considerably in their MAEs, as well in the decrease in error, depending on the particular function, see Table 3, for errors, see Figure 6. In contrast to Experiment 1, errors in conditions with linear functions did not differ significantly between the first and the last block $t(56.04) = 0.81$, $p = .423$. Neither did errors differ significantly in OU conditions, $t(50.13) = 0.37$, $p = .716$. However, errors for periodic conditions differed significantly between the first and the last block, $t(42.62) = 2.38$, $p = .022$. As in Experiment 1 most conditions exhibited very low errors. In contrast Cos_2 and Ou_2 were characterized by large MAEs. For error rates for all conditions, see Figure 6.

Table 3: MAEs in Experiment 3.

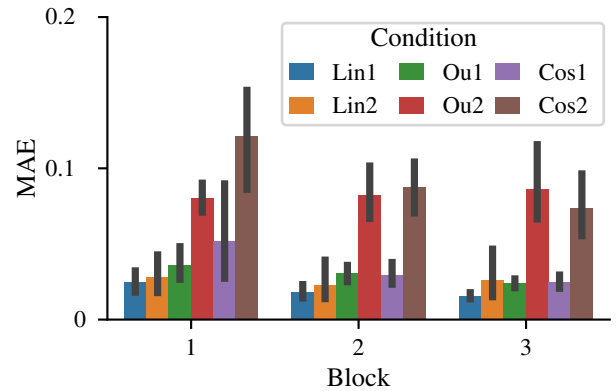| Function	MAE_{b1}	SD_{b1}	MAE_{b2}	SD_{b2}	MAE_{b3}	SD_{b3}
Lin	0.03	0.02	0.02	0.02	0.02	0.02
OU	0.06	0.03	0.05	0.04	0.05	0.05
Cos	0.09	0.08	0.06	0.04	0.05	0.04

Figure 6: MAEs and 95% confidence intervals for each condition in Experiment 3.

Extrapolating

Visual inspection of the extrapolation strongly suggested that variances between OU-, frequencies for periodic- and slopes for linear conditions reflected training functions, see Figure 1. To evaluate if these patterns were also well aligned with the generating models, and if samples reflected the differences in function parametrization, we performed maximum-likelihood estimation (MLE) for each individual participant and each generating GP. We then used the type of the generating GP with the highest likelihood to predict which training samples the participant had been assigned to. This approach allowed us to evaluate if the experimental manipulation resulted in extrapolation patterns consistent with the generating GPs. Our classification procedure classified 22 out of 30 participants in the OU conditions correctly, a proportion that was significantly larger than expected by chance (1/3),
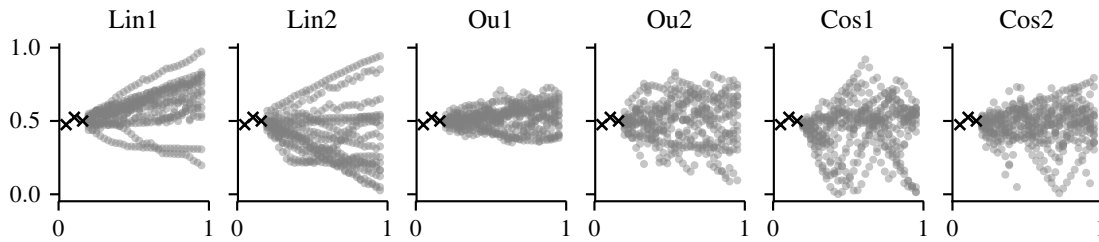
Figure 7: Participant extrapolations in Experiment 3. Extrapolations closely matched the learned type of function and its detailed parametrization, see Figure 1

$p_{OU} < .001$[6]. In the periodic conditions, 17 out of 30 participants were classified correctly, $p_{Cos} = .007$. However, for linear samples, only 10 out of 31 participants were classified correctly, $p_{Lin} = .617$. For the full confusion matrix, see Figure 8.
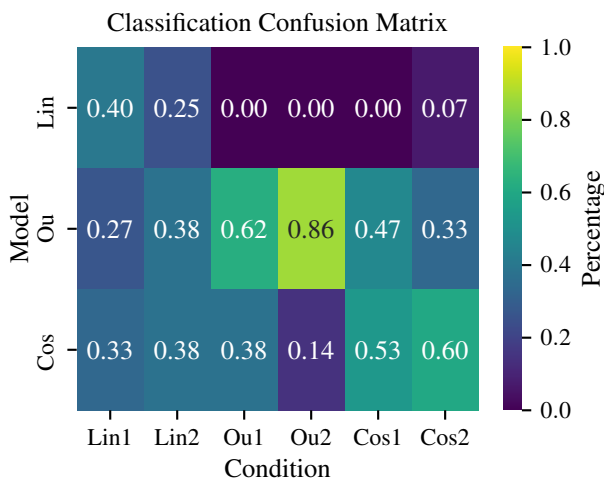

Figure 8: Confusion matrix for the proportion of participants assigned to each model. Our method was able to accurately recover which type of function participants were trained on for OU and periodic, but not linear conditions.

To compare the inferred parametrizations across training conditions of one function type, we contrasted the parameters obtained via MLE for the true model. For linear functions, the MLE parameters for slopes differed significantly between conditions, $M_{Lin1} = 0.2$, $SD_{Lin1} = 0.25$, $M_{Lin2} = -0.1$, $SD_{Lin2} = 0.34$, $t(27.49) = 2.82$, $p = .009$[7], with the signs of the inferred slopes matching the training. Neither intercept, variance or noise estimates differed significantly between conditions (all $p > .1$). The inferred parameters for variance in the OU conditions did not differ significantly, but were reflective of differences in training, $M_{OU1} = 0.002$, $SD = 0.002$, $M_{OU2} = 0.007$, $SD_{OU2} = 0.008$, $t(15) = -1.95$, $p = .071$.

---

[6]All tests are one-sided, exact Binomial tests.

[7]All tests in this section are unequal variance, two-sided t-tests.

The inferred length scale did not differ significantly between conditions, but was slightly higher for $OU1$, $M_{OU1} = 0.38$, $SD_{OU1} = 0.39$, $M_{OU2} = 0.21$, $SD_{OU2} = 0.28$, $t(26.31) = 1.46$, $p = .157$. Both intercept and noise estimates did not differ significantly between conditions (all $p > .5$). The inferred parameters for periodic conditions did not differ significantly for length scale, $M_{Cos1} = 0.08$, $SD_{Cos1} = 0.06$, $M_{Cos2} = 0.08$, $SD_{Cos2} = 0.1$, $t(22.47) = 0.27$, $p = .79$. Instead, conditions differed significantly for variance $M_{Cos1} = 0.02$, $SD_{Cos1} = 0.02$, $M_{Cos2} = 0.01$, $SD_{Cos2} = 0.01$, $t(20.1) = 2.25$, $p = .036$. Estimates for intercepts and noise were not significantly different between conditions (all $p > 0.1$).

Discussion

We found evidence that participants choose patterns and extrapolate in ways consistent with the learned function type. Furthermore, contrasting the extrapolations in the transfer set within function conditions, suggested that these patterns differed in ways consistent with our experimental manipulation.

While participants' judgments generally reflected the functions they learned during training, our results also highlight characteristic human biases. In the Cos_1 condition, participants preferred high-frequency periodic samples over the true low-frequency samples. Similarly, participants in the Ou_1 conditions, preferred the higher variance samples, or even periodic samples over the trained low-variance samples. One explanation for this biases could be that people have a strong preference for particular functions because these parametrizations are well adapted to environmental regularities. As a result, these functions would be robust and applicable to a wide range of task in the environment. This explanation would be consistent with recent results in human exploration, where participants exhibited a tendency to undergeneralize spatial correlations, but this undergeneralization resulted in comparable or even better performance than a ground-truth matching model (Wu et al., 2018).

To better describe these characteristic human biases and explore their potential rational grounding, future research should more closely examine which statistical patterns can be generalized and under which circumstances these generalizations are performed. For example, while our experiment imposed that all patterns followed the same relationship, in

reality this information is rarely available. Thus, future research should examine under which circumstances task regularities are inferred to be similar, and what kinds of notions of similarity can guide these generalizations. One exciting prospect is to link notions of hierarchical- and compositional representations and function generalization. We are currently exploring how such compositional regularities can aid transfer and generalization in sparse domains.

Acknowledgements

References

Bott, L., & Heit, E. (2004). Nonmonotonic extrapolation in function learning. *Journal of Experimental Psychology*, *30*(1), 38.

Brehmer, B. (1976). Learning complex rules in probabilistic inference tasks. *Scandinavian Journal of Psychology*, *17*(1), 309–312.

Busemeyer, J. R., Byun, E., Delosh, E. L., & McDaniel, M. A. (1997). Learning functional relations based on experience with input-output pairs by humans and artificial neural networks. In K. Lamberts & D. R. Shanks (Eds.), *Studies in cognition. knowledge, concepts and categories* (p. 408-437). The MIT Press.

Carroll, J. D. (1963). Functional learning: The learning of continuous functional mappings relating stimulus and response continua. *ETS Research Bulletin Series*, *1963*(2), i–144.

DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology*, *23*(4), 968.

Gershman, S. J. (2017). On the blessing of abstraction. *The Quarterly Journal of Experimental Psychology*, *70*(3), 361–365.

Gershman, S. J., & Niv, Y. (2010). Learning latent structure: carving nature at its joints. *Current opinion in neurobiology*, *20*(2), 251–256.

Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive science*, *32*(1), 108–154.

Hamrick, J. B., Ballard, A. J., Pascanu, R., Vinyals, O., Heess, N., & Battaglia, P. W. (2017). Metacontrol for adaptive imagination-based optimization. *arXiv preprint arXiv:1705.02670*.

Harlow, H. F. (1949). The formation of learning sets. *Psychological review*, *56*(1), 51.

Kalish, M. L. (2013). Learning and extrapolating a periodic function. *Memory & Cognition*, *41*(6), 886–896.

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: knowledge partitioning and function learning. *Psychological Review*, *111*(4), 1072.

León-Villagrá, P., Preda, I., & Lucas, C. G. (2018). Data availability and function extrapolation. In *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Lucas, C. G., Griffiths, T. L., Williams, J. J., & Kalish, M. L. (2015). A rational model of function learning. *Psychonomic Bulletin & Review*, *22*(5), 1193–1215.

Lucas, C. G., Sterling, D., & Kemp, C. (2012). Superspace extrapolation reveals inductive biases in function learning. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 34).

McDaniel, M. A., & Busemeyer, J. R. (2005). The conceptual basis of function learning and extrapolation: Comparison of rule-based and associative-based models. *Psychonomic Bulletin & Review*, *12*(1), 24–42.

Schulz, E., Tenenbaum, J. B., Duvenaud, D., Speekenbrink, M., & Gershman, S. J. (2017). Compositional inductive biases in function learning. *Cognitive Psychology*, *99*, 44–79.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*(6022), 1279–1285.

Wilson, A. G., Dann, C., Lucas, C. G., & Xing, E. P. (2015). The human kernel. In *Advances in Neural Information Processing Systems* (pp. 2854–2862).

Wilson, R. C., & Niv, Y. (2012). Inferring relevance in a changing world. *Frontiers in human Neuroscience*, *5*, 189.

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*(12), 915.