Human-like Function Learning and Transfer

Pablo León Villagrá, Christopher G. Lucas

School of Informatics,
University of Edinburgh

Function learning (or regression) problems are ubiquitous in human experience and machine learning. Humans can generalise in diverse ways that respect the abstract structure of a problem and can use knowledge in one context to inform decisions in another. Knowledge transfer is common in applied statistics, as when a practitioner recognises that kinds of regression problems involve certain parametric relationships. It is also at the heart of scientific progress, e.g., when analogies lead to new hypotheses and discoveries [5]. In some situations, data are plentiful and transfer of knowledge is relatively unimportant, but when data are sparse, having appropriate prior knowledge is essential.

In this work, we explore human-like generalisation in regression problems, using psychological experiments and probabilistic models. Specifically:

- We present evidence that humans can learn and generalise from relationships in ways that reflect the compositional structure of those relationships.
- These learned relationships are re-usable: they shape subsequent inferences and lead to structured extrapolations in the face of extremely sparse data.
- We describe a model that explains qualitative features of human judgements in cases where previous models fail, and re-uses compositional representations to extrapolate from sparse data.

1 Learning a Language of Shared Functional Expressions

Under a probabilistic perspective, extrapolation in function learning problems[1] involves inferring the distribution of the target variable $y^* \in \mathbb{R}$ given predictors $x^* \in \mathbb{R}^d$ and data $\mathbf{x}_n = x_{1,\ldots,n}$ $\mathbf{y}_n = y_{1,\ldots,n}$. If the relationship follows some unknown function $f(\cdot)$, then $p(y^*|x^*, \mathbf{x}_n, \mathbf{y}_n) = \int_{\mathcal{F}} p(y^*|x^*, f)p(f|\mathbf{x}_n, \mathbf{y}_n)df$, where $p(f|\mathbf{x}_n, \mathbf{y}_n)$ is determined by the data likelihood and prior beliefs about f. In this work we posit that our mental representations of these functions are, in all but the simplest cases, constructed from a vocabulary of simpler functional components. This idea follows in the spirit of other computational models of human learning and concept formation, e.g., [1, 6], but our fully probabilistic approach and its application to function learning are both novel.

Our model builds in part on Lucas et al. [4], using Gaussian processes as a unifying representation for both parametric and and arbitrary smooth relationships. In addition, we expand on recent unsupervised structure learning models [2] and formalise the construction and re-use of these compositional elements

[1] That is, regression problems presented to human learners.

using Pitman-Yor adaptor grammars [3]. Where [2] showed the value in composing Gaussian process kernels to discover structure, ours allows the grammar itself to be expanded as new problems are encountered, enabling extrapolation according to complex patterns even in sparse domains. Using representational elements that originate from an infinite distribution over base types, an agent can introduce both compositions of existing elements and new, real-valued latent variables that make it possible to learn both structure and a set of parametrized primitives.

Under this representation, the task becomes one of finding a vocabulary of concepts, latent classes of relationships, and parameters that explain the totality of the agent's past experience, across diverse and heterogeneous problems (or approximate distributions over the same).

2 Inference for Functional Compositions

Given our use of adaptor grammars, which are traditionally used to discover structure in natural language, we can think of each relationship as being analogous to a word, having a distribution over possible morphological "parses". We draw samples from the distribution over structural parses for individual relationships using blocked Metropolis-Hastings (MH) sampling, with priors based on a locally finite PCFG approximation of an adaptor grammar, as described in Johnson, Griffiths, and Goldwater [3]. This step of inference is conditional on parametrised primitive kernels. Interleaved with this structure-resampling step, we update the primitive kernels, also using blocked MH. Our base grammar includes linear, periodic, squared-exponential, and white-noise kernels, as well as product and sum operations (see [2] for a description of these kernels and the kinds of relationships that can be represented by composing them).

2.1 Human Biases in Compositional Transfer

We ran an experiment to test the hypotheses that (1) people can discover structure in relationships that are complex by the standards of past function learning tasks, but have a simple compositional structure; and (2) people can discover and re-use that structure to make sense of sparse data in subsequent tasks. In our experiment, participants had to sequentially extrapolate from two functions that were presented as scatterplots. When participants were told that a third functional relationship was similar to the previous data, their extrapolations tended to reflect the structure of previous functions, where past models and experiments suggest that participants should extrapolate linearly. We have also performed a preliminary model evaluation, examining the ability to transfer functional components to new domains. In our first evaluations the model extrapolated based on small-scale periodic characteristics in the sparse dataset, an effect that was not apparent when we did not assume the first two datasets to be related. See Figure 1 for examples of human and model extrapolations.

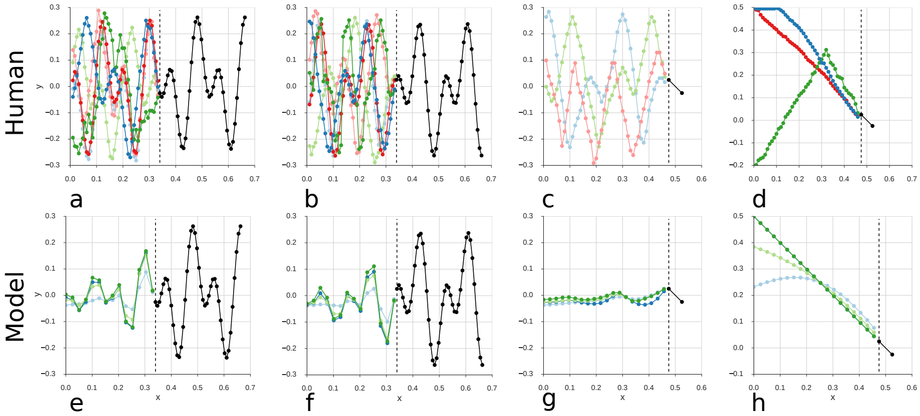

Fig. 1. Black points denote training data; coloured points denote model and human extrapolations. **Top Row**: Extrapolations of 6 participants in our experiment. After seeing two blocks of data and performing extrapolations (a, b), participants extrapolate given only two points (c, d). The participants in (c) were told that the relationship was the same as in (a,b). Participants in (d) were told that the relationship was different. **Bottom Row**: Extrapolations for 3 samples from our model. For (e-g), a model was trained on the full dataset (all black points in e-g). In (h), extrapolations reflect only the two single-problem points. The alphabet of base kernels that could be combined consisted of two different linear, two periodic, two squared-exponential, and two white-noise kernels, with diffuse priors on their parameters.

References

[1] Eyal Dechter et al. "Bootstrap learning via modular concept discovery". In: *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*. AAAI Press, 2013, pp. 1302–1309.

[2] David Duvenaud et al. "Structure Discovery in Nonparametric Regression through Compositional Kernel Search". In: *Proceedings of the 30th International Conference on Machine Learning*. Vol. 28. 2013, pp. 1166–1174.

[3] Mark Johnson, Thomas L Griffiths, and S Goldwater. "Adaptor grammars: A framework for specifying compositional nonparametric Bayesian models". In: *Advances in Neural Information Processing Systems* 19 (2007), p. 641. ISSN: 10495258.

[4] Christopher Lucas et al. "A rational model of function learning." In: *Psychonomic bulletin & review* 22.5 (2015), pp. 1193–215. ISSN: 1531-5320.

[5] Robert Oppenheimer. "Analogy in science." In: *American Psychologist* 11.3 (1956), p. 127.

[6] Steven T Piantadosi, Joshua B Tenenbaum, and Noah D Goodman. "Bootstrapping in a language of thought: A formal model of numerical concept learning". In: *Cognition* 123.2 (2012), pp. 199–217.