60 ideas from “Compositional Inductive Biases in Function Learning” (2017) by Eric Schulz, Joshua Tenenbaum, David Duvenaud, Maarten Speekenbrink, Samuel Gershman

[Link to actual paper]

1. How do people recognize and learn about complex functional structure?
2. Within the framework of Bayesian regression, using a grammar over Gaussian process kernels
3. Participants showed a preference for compositional (over non-compositional) interpolations and extrapolations of functions
4. Experiments designed to elicit priors over functional patterns revealed a compositional inductive bias
5. Compositional functions were perceived as subjectively more predictable than non-compositional functions, and showed other signatures of predictability such as enhanced memorability and reduced numerosity
6. These results support the view that the human intuitive theory of functions is inherently compositional
7. Recognizing functional patterns underlies our perception of time, space, and number
8. Inductive biases are needed to constrain the set of plausible inferences
9. There are two major theoretical accounts of human function learning: similarity-based accounts and rule-based accounts
10. An important characteristic of human learning: strong inferences from small amounts of data
11. The divide and conquer strategy is known to be used in concept learning, language, and visual perception.
12. Using a set of compositional rules, complex structures and representations are built from smaller building blocks.
13. Compositional systems support strong inferences from small amounts of data by imposing structural constraints without sacrificing the capacity for representing an infinite variety of forms
14. The primary claim of this paper is that human function learning is constrained by compositional inductive biases
15. We need a theoretical framework to represent and reason about compositional function spaces
16. In 2015, Lucas, Griffith, Williams, and Kalish presented a normative theory of human function learning using GPs
17. GPs are distributions over functions that encode properties such as smoothness, linearity, periodicity, symmetry, and more
18. We build on the GP approach by studying, both experimentally and theoretically, the compositional nature of inductive biases in function learning.

19. The primary theoretical contribution is to extend the GP formalism to modeling human function learning with a prior that obeys compositionally structured constraints.
20. 10 experiments compare a compositional prior to a flexible, non-compositional one
21. Both models use Bayesian inference, but differ in their inductive biases
22. Mechanistic hypotheses do not directly give insight into inductive biases
23. Different mechanisms may or may not produce the same bias
24. If our goal is to understand human inductive biases, we need a computational level analysis that is agnostic to mechanism
25. A GP is a collection of random variables, any finite subset of which are jointly gaussian distributed
26. From Lucas et. al 2015: In Bayesian linear regression, a hypothesis space of functions is defined…
27. A prior is defined over that space…
28. Predictions are formed by averaging over the posterior probability of y.
29. By Mercer's theorem, any positive definite kernel can be expressed as the outer product of feature vectors…
30. k(x,x') = infinite series with the dth term lamba_d * phi_d(x) * phi_d(x') where lamba_d is the dth eigenvalue and phi_d is the dth eigenfunction of the kernel
31. Priors over functions can be encoded by the kernel
32. Two candidate kernel approximations that express conceptually different inductive biases
33. By Bochner's theorem, any stationary kernel can be expressed as an integral
34. Any stationary distribution can be expressed as a spectral density
35. A spectral density over a kernel space fully defines the kernel
36. Given a set of input-output pairs, the task facing the learner is to identify both the function and the underlying parse tree
37. Pattern recognition as a window into cognitive representations
38. If participants have compositional, structured representations..
39. Then they should prefer pattern completions generated by the compositional kernel.
40. Using a hierarchical Bayesian model to estimate the posterior probability of choosing the compositional completion
41. Discrete wavelet Haar transform
42. Change detection, numerosity perception, short-term memory
43. Perceived numerosity diminishes as structure becomes more discernible
44. Structural regularities distort the units of perception, making them appear less numerous
45. Structured representations facilitate change detection…
46. By storing summary representations of the stimulus in short-term memory
47. Summary representations free up encoding resources for a small number of outliers
48. What are the inductive biases that constrain short-term memory representations?
49. In 2013, Brady and Tenenbaum used Markov random fields to encode information about object features

50. Structural regularities expand memory capacity because they are “compressible.”
51. Chunking is known to be fundamental to exceptional expert memory and story comprehension
52. In a version of the Sternberg task, compositional functions were shown to be more memorable than non-compositional functions
53. Why might an intelligent agent exhibit compositional inductive biases?
54. Testing if participants’ compositional priors over different domains track the structure within those domains
55. If structure exists that a grammar can express, then an agent can save an unbounded number of bits by detecting that structure
56. Compositionality helps memorizing structure by providing naturally occurring chunks
57. These inductive biases may help with the encoding and retrieval of structure in the real world
58. The structure search algorithm that was employed here was recently shown to be statistically efficient
59. Language and object perception have long traditions of emphasizing compositionality.
60. An important direction for future work is to systematically investigate the boundaries of compositional functions.