PROBABILISTIC PROGRAMMING VERSUS META-LEARNING AS MODELS OF COGNITION

Desmond C. Ong [*]
Department of Psychology
University of Texas at Austin
desmond.ong@utexas.edu

Tan Zhi-Xuan & Josh Tenenbaum
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
{xuan,jbt}@mit.edu

Noah D. Goodman
Department of Psychology and Department of Computer Science
Stanford University
ngoodman@stanford.edu

ABSTRACT

We summarize the recent progress made by probabilistic programming as a unifying formalism for the probabilistic, symbolic, and data-driven aspects of human cognition. We highlight differences with meta-learning in flexibility, statistical assumptions and inferences about cognition. We suggest that the meta-learning approach could be further strengthened by considering Connectionist and Bayesian approaches, rather than exclusively one or the other.

Connectionist-*versus*-Bayesian debates have occurred in cognitive science for decades (Griffiths et al., 2010; McClelland et al., 2010), with each side progressing in theory, models, and algorithms, in turn impelling the other side to advance, resulting in a cycle of fruitful engagement. The recent summary of the meta-learning paradigm that Binz and colleagues proposed in the target article bridges the two by proposing how meta-learning in recurrent neural networks can address some of the traditional challenges of Bayesian approaches. But, by failing to recognize and engage with the latest iteration of Bayesian modeling approaches—including probabilistic programming as a unifying paradigm for probabilistic, symbolic, and differentiable computation (Cusumano-Towner et al., 2019)—this article fails to push the meta-learning paradigm as far as it could go.

The authors begin their defense of meta-learning by citing the intractability of exact Bayesian inference. However, this fails to address how and why meta-learning is superior to approximate inference for modeling cognition. As the authors themselves note, Bayesian modelers use a variety of approximate inference methods, including neural-network-powered variational inference (Dasgupta et al., 2020; Kingma & Welling, 2013), Markov chain Monte Carlo (Ullman et al., 2012), and Sequential Monte Carlo methods (Levy et al., 2008; Vul et al., 2009), which have all shown considerable success in modeling how humans perform inference (or fail to) in presumably intractable settings. As such, it is hardly an argument in favor of meta-learning—and against "traditional" Bayesian models—that exact inference is intractable.

This omission is just one way in which the article fails to engage with a modern incarnation of the Bayesian modeler's toolkit—Probabilistic Programming. In the past two decades, we have seen the development of probabilistic programming as unifying formalism for modeling the probabilistic, symbolic, and data-driven aspects of human cognition (Lake et al., 2015), as embodied in probabilistic programming language such as Church (Goodman et al., 2012), webPPL (Goodman & Andreas, 2014), Pyro (Bingham et al., 2019) and Gen (Cusumano-Towner et al., 2019). These languages enable modelers to explore a much wider range of computational architectures than the standard meta-learning setup, which requires modelers to reformulate human cognition as a sequence prediction problem. Probabilistic programming allows modelers to unite the strengths of general-purpose predictors (i.e., neural networks) with theoretically-informed constraints and model-based reason-

---

ing. For instance, Ong and colleagues (2021) showed how reasoning about others' emotions can be modeled by combining the constraints implied by cognitive appraisal theory with bottom-up representations learnt via neural networks from emotional facial expressions. Similarly, several recent papers have shown how the linguistic abilities of large language models (LLMs) can be integrated with rational models of planning, communication, and inverse planning (Wong et al., 2023; Ying et al., 2023), modeling human inferences that LLM-based sequence prediction alone struggle with (Zhi-Xuan et al., 2024).

What flexibility does probabilistic programming afford over pure meta-learning? As the article notes, one potential benefit of meta-learning is that it avoids the need for a specific Bayesian model to perform inference over. Crucially, meta-learning achieves this by having access to sufficiently similar data at training and test time, such that the meta-learned algorithm is sufficiently well-adapted to the implied class of data-generating processes. Human cognition is much more adaptive. We do not simply adjust our learning to fit past distributions; we also construct, modify, abstract, and refactor entire theories about how the world works (Rule et al., 2020; Tenenbaum et al., 2011; Ullman & Tenenbaum, 2020), reasoning with such theories on downstream tasks (Tsividis et al., 2021). This capacity is not captured by pure meta-learning, which occurs "offline". By contrast, probabilistic programming allows modeling these patterns of thought: Theory building can be formulated as program induction (Lake et al., 2015; Saad et al., 2019), refactoring as program merging (Hwang et al., 2011), and abstraction-guided reasoning as coarse-to-fine inference (Cusumano-Towner et al., 2018; Stuhlmüller et al., 2015). Inference meta-programs (Cusumano-Towner et al., 2019; Lew et al., 2023) allow us to model how people invoke modeling and inference strategies as needed: one can employ meta-learned inference when one believes a familiar model applies, but also flexibly compute inferences when a model is learned, extended, or abstracted. On this view, meta-learning has an important role to play in modeling human cognition, but not for all of our cognitive capacities.

Another way of understanding the relationship between meta-learning and probabilistic programming is that the former uses implicit statistical assumptions while the latter's assumptions are explicit. Meta-learning assumes that the structure of the world is conveyed in the statistical structure of data across independent instances. With sufficient coverage of the training distribution, flexible deep learning approaches fit this structure and use it to generalize. But they may not do so in a way that may provide any insight into the computational problem being solved by humans. Probabilistic programs, by contrast, explicitly hypothesize the statistical patterns to be found in data, providing constraints that, if satisfied, yield insights for cognition. This implicit-explicit distinction both frames the relative value of the approaches and suggests an alternative relation: a Bayesian model need not subsume or integrate what is learned by a deep learning model, but simply explicate it, at a higher level of analysis. Through this lens, having to specify an inference problem is not a limitation, but a virtue.

The best of both worlds will be to compose and further refine these paradigms, such as using deep amortized inference (like meta-learning for Probabilistic Programming), using Bayesian tools (and other tools for mechanistic interpretation) to understand the results of meta-learning, or constructing neurosymbolic models (e.g., by grounding the outputs of meta-learned models in probabilistic programs, as in Wong et al., 2023). As a very recent example, Zhou, Feinman, and Lake (2024) proposed a neurosymbolic program induction model to capture human visual learning, using both Bayesian program induction and meta-learning, achieving the best of both approaches: interpretability and parsimony, as well as capturing additional variance using flexible function approximators. We believe that the field should move beyond "Connectionist-*versus*-Bayesian" debates to instead explore hybrid "Connectionist-*and*-Bayesian" approaches.

References

Bingham, E., Chen, J. P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletsos, T., Singh, R., Szerlip, P., Horsfall, P., & Goodman, N. D. (2019). Pyro: Deep universal probabilistic programming. *The Journal of Machine Learning Research*, *20*(1), 973–978.

Binz, M., Dasgupta, I., Jagadish, A. K., Botvinick, M., Wang, J. X., & Schulz, E. (forthcoming). Meta-learned models of cognition. *Behavioral and Brain Sciences*.

Cusumano-Towner, M. F., Bichsel, B., Gehr, T., Vechev, M., & Mansinghka, V. K. (2018). Incremental inference for probabilistic programs. *Proceedings of the 39th ACM SIGPLAN Conference on Programming Language Design and Implementation*, 571–585.

Cusumano-Towner, M. F., Saad, F. A., Lew, A. K., & Mansinghka, V. K. (2019). Gen: A general-purpose probabilistic programming system with programmable inference. *Proceedings of the 40th acm sigplan conference on programming language design and implementation*, 221–236.

Dasgupta, I., Schulz, E., Tenenbaum, J. B., & Gershman, S. J. (2020). A theory of learning to infer. *Psychological Review*, *127*(3), 412.

Goodman, N. D., & Andreas, S. (2014). The design and implementation of probabilistic programming languages. *R etrieved from http://dippl. org*.

Goodman, N. D., Mansinghka, V., Roy, D. M., Bonawitz, K., & Tenenbaum, J. B. (2012). Church: A language for generative models. *arXiv preprint arXiv:1206.3255*.

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*(8), 357–364.

Hwang, I., Stuhlmüller, A., & Goodman, N. D. (2011). Inducing probabilistic programs by bayesian program merging. *arXiv preprint arXiv:1110.5667*.

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, *350*(6266), 1332–1338.

Levy, R., Reali, F., & Griffiths, T. (2008). Modeling the effects of memory on human online sentence processing with particle filters. *Advances in Neural Information Processing Systems*, *21*.

Lew, A. K., Matheos, G., Zhi-Xuan, T., Ghavamizadeh, M., Gothoskar, N., Russell, S., & Mansinghka, V. K. (2023). SMCP3: Sequential monte carlo with probabilistic program proposals. *International Conference on Artificial Intelligence and Statistics*, 7061–7088.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences*, *14*(8), 348–356.

Ong, D. C., Soh, H., Zaki, J., & Goodman, N. D. (2021). Applying probabilistic programming to affective computing. *IEEE Transactions on Affective Computing*, *12*(2), 306–317.

Rule, J. S., Tenenbaum, J. B., & Piantadosi, S. T. (2020). The child as hacker. *Trends in Cognitive Sciences*, *24*(11), 900–915.

Saad, F. A., Cusumano-Towner, M. F., Schaechtle, U., Rinard, M. C., & Mansinghka, V. K. (2019). Bayesian synthesis of probabilistic programs for automatic data modeling. *Proceedings of the ACM on Programming Languages*, *3*(POPL), 1–32.

Stuhlmüller, A., Hawkins, R. X., Siddharth, N., & Goodman, N. D. (2015). Coarse-to-fine sequential monte carlo for probabilistic programs. *arXiv preprint arXiv:1509.02962*.

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.

Tsividis, P. A., Loula, J., Burga, J., Foss, N., Campero, A., Pouncy, T., Gershman, S. J., & Tenenbaum, J. B. (2021). Human-level reinforcement learning through theory-based modeling, exploration, and planning. *arXiv preprint arXiv:2107.12544*.

Ullman, T. D., Goodman, N. D., & Tenenbaum, J. B. (2012). Theory learning as stochastic search in the language of thought. *Cognitive Development*, *27*(4), 455–480.

Ullman, T. D., & Tenenbaum, J. B. (2020). Bayesian models of conceptual development: Learning as building models of the world. *Annual Review of Developmental Psychology*, *2*, 533–558.

Vul, E., Alvarez, G., Tenenbaum, J., & Black, M. (2009). Explaining human multiple object tracking as resource-constrained approximate inference in a dynamic probabilistic model. *Advances in Neural Information Processing Systems*, *22*.

Wong, L., Grand, G., Lew, A. K., Goodman, N. D., Mansinghka, V. K., Andreas, J., & Tenenbaum, J. B. (2023). From word models to world models: Translating from natural language to the probabilistic language of thought. *arXiv preprint arXiv:2306.12672*.

Ying, L., Zhi-Xuan, T., Mansinghka, V., & Tenenbaum, J. B. (2023). Inferring the goals of communicating agents from actions and instructions. *Proceedings of the AAAI Symposium Series*, *2*(1), 26–33.

Zhi-Xuan, T., Ying, L., Mansinghka, V., & Tenenbaum, J. B. (2024). Pragmatic instruction following and goal assistance via cooperative language guided inverse plan search. *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*.

Zhou, Y., Feinman, R., & Lake, B. M. (2024). Compositional diversity in visual concept learning. *Cognition*, *244*, 105711.