Mechanisms of posterior probability re-estimation during gradual evidence acquisition in causal structure learning

Nikolay Dagaev, Elena Gorina, Mario Martinez-Saito
National Research University Higher School of Economics, Moscow, Russia
Presenting author: Elena Gorina

There is a substantial evidence that people do not perform exact Bayesian inference in general, and particularly in causal learning. Indeed, as problem size increases, computational demands quickly become difficult to manage due to cognitive constraints such as working memory or attention capacity. Nevertheless, there is a set of conditions under which a reasonable strategy might be to compute exact posterior probabilities. First, evidence arrives one piece at a time and an access to previous evidence is strongly restricted. Second, learners need to assess posterior probabilities each time a new piece of evidence arrives. Finally, a problem at hand must be simple in terms of computational demands, which implies a fairly small hypothesis space.

The Bayesian re-estimation in this context can be accomplished in two ways: an examples-based strategy and a posterior-based strategy. In the examples-based strategy, previously seen examples (evidence) are retained and, along with the original prior probabilities, used to compute posterior when a new example is received. In the posterior-based strategy, posterior probabilities are retained and used as prior when a new evidence is received. Thus, the examples-based strategy requires more computations as the number of examples seen increases, but for the posterior-based strategy the computational complexity remains constant. Although these two ways to update posterior are mathematically equivalent, they can be differentiated in humans due to memory limitations. In fact, if a subject employs the examples-based strategy, his or her posterior distribution over hypotheses will generally deviate from Bayesian ideal more and more as the number of examples increases since only several recent examples are retained. Under the posterior-based strategy, this posterior distribution will match Bayesian ideal after any number of examples.

We designed a series of experiments to test which of the mechanisms described above is responsible for causal structure learning in humans. In these experiments, a non-deterministic causal system of four binary elements is employed, and participants are sequentially presented with a set of the system's resultant states (examples). As a new resultant state is presented, participants are asked to indicate the most probable scheme of causal connections (hypothesis) in the system given this state and those demonstrated previously. We use two sizes of hypothesis space: two and six possible schemes of connections. To distinguish between the examples-based and posterior-based strategies, we will compare participants' responses to a limited memory model (posterior re-estimation based on several most recent examples) and the ideal Bayesian model.

Under the small hypothesis space size, we expect participants to employ the posterior-based strategy as it is more optimal in terms of memory resources to retain posterior probabilities, not examples. However, by increasing the size of hypothesis space, we expect participants to resort to the examples-based strategy as it becomes difficult to memorize exact posterior probabilities. Additionally, we assume that participants will find it more natural to employ the posterior-based strategy when they are asked not to choose a single most likely option but to provide probability judgments for all response options.