# Bayesian learning: Challenges, limitations and pragmatics

Heide, R. de

# Chapter 8

# Discussion and future work

In this chapter I concisely review the previous six chapters of this dissertation, and explore some open challenges and possible directions for future work.

## 8.1 Forward-looking Bayesians

In Chapter 2 we studied the failure of weak truth-merger of Wenmackers and Romeijn's open-minded Bayesians, and we proposed two versions of forward-looking open-minded Bayesians that do weakly merge with the truth when the truth is added at some point in time. In Chapter 2 we only focus on *how* to incorporate new hypotheses. A direction for future research, possibly for me and my co-author on this chapter, is to formalise *when* new hypotheses should be considered, and to investigate how this interacts with the guarantee of truth-merger.

Chapter 2 inspired the following idea for a future project for myself in the area of continuous-armed best-arm identification in machine learning. This protocol can be viewed as similar to the protocol of the forward-looking Bayesians, if we let arms correspond to hypotheses, however, it is still unclear what the relation is between truth-merger and identification. The algorithms proposed in papers on best-arm identification in continuous-armed bandits (Bubeck, Munos and Stoltz, 2009; Carpentier and Valko, 2015; Aziz et al., 2018) employ two phases: First, a finite subset of arms from a continuous reservoir is selected, and subsequently a finite-armed bandit algorithm is run on this subset to identify the best arm. An interesting idea would be to propose an algorithm that decides during the learning process to add (or remove) arms from the finite set under consideration, which might lead to simple regret bounds scaling better in the confidence parameter $\delta$ in the fixed-confidence setting. Another future course would be to propose a Bayesian algorithm for best-arm identification in continuous-armed bandits, which can also be seen as an extension of the algorithms discussed in Chapter 7, see also the upcoming Section 8.4. This is both conceptually interesting because of the link with the forward-looking Bayesians and Bayesian confirmation theory, and also interesting because the Bayesian sampling rules of Chapter 7 do not depend on a confidence parameter or time

horizon. The combination of these two challenges is to propose a Bayesian algorithm for best-arm identification in continuous-armed bandits that adds or removes arms in course of the learning process. This algorithm could also provide some insights for the problem of *when* to add new hypotheses in the framework of the forward-looking Bayesians.

## 8.2 Hypothesis testing

Chapters 3 and 4 deal with the question whether Bayes factor hypothesis testing is robust under *optional stopping*. The bottom line of these chapters is that the answer to this question depends on one's perspective on Bayesianism (see also Section 1.2) and which definition of optional stopping one employs — we give three distinct mathematical definitions in Chapter 4. It is remarkable how resolutely some authors advocate the use of their favourite method for hypothesis testing, and how firm their reproach sometimes is to other authors who nuance or criticise claims about these methods, see for example (Benjamin et al., 2018) and (McShane et al., 2019); and even before being published, Chapter 3 provoked several responses (Rouder, 2019; Wagenmakers, Gronau and Vandekerckhove, 2019; Rouder and Haaf, n.d.) . In light of this fierce defence of some specific methods for hypothesis testing, an interesting project would be to investigate the role of hypothesis testing in the behavioural sciences. In a paper related to this subject, Gigerenzer and Marewski (2014) argue that "determining significance has become a surrogate for good research". The current discussion on optional stopping with Bayes factors that is the subject of Chapter 3 seems to be an example of that shift in focus from the actual goals of science to the surrogate of "mindless mechanical statistics". Goals of science include gaining knowledge about the world around us, and hypothesis testing is one of the means scientists have at their disposal to achieve that. How clear this distinction between goals and means is in current research in the behavioural sciences, and what the role of hypothesis testing in scientific research should be, are subjects to be addressed, possibly by philosophers of science.

In Chapter 5 we proposed a new theory for hypothesis testing based on E-values. From a practical perspective, it is now important to develop software for calculating E-values for common hypothesis tests, so that practitioners can start working with E-value based hypothesis tests. From a theoretical perspective, there are some open questions arising in particular from the combination of Chapter 4 and 5. The former chapter provides results showing that using the right Haar prior in general group invariant cases leads to E-values, however, in Chapter 5 is only shown that these are GROW E-values for the particular (important) case of the $t$-test. An objective for future work is thus to extend this to a general group-invariant setting. Further goals for future work on Safe Testing include the construction of confidence intervals by *inverting* a safe test. When this safe test constitutes a *test martingale*, these confidence intervals are *always valid confidence intervals* in the sense of Howard et al.'s 2018b framework of uniform, nonparametric, non-asymptotic confidence sequences (Darling and Robbins, 1967; Lai, 1984). The intuitions behind the construction of safe tests can lead to other constructions of confidence intervals. Further future objectives are to investigate the connections of safe testing to Shafer and Vovk's 2019 game-theoretic probability framework, and to the framework of always-valid P-values (Robbins, 1970; Robbins and Siegmund, 1972; Robbins and Siegmund, 1974; Johari, Pekelis and Walsh, 2015). The group of prof. Grünwald at CWI is working on these practical and theoretical challenges.

## 8.3 Safe-Bayesian generalised linear regression

Chapter 6 provides theoretical evidence that $\eta$-generalised Bayes can outperform standard Bayes for generalised linear models, and provides empirical evidence for Bayesian lasso and logistic regression. We also provided MCMC samplers for the generalised Bayesian lasso and logistic regression. The Gibbs sampler for the latter is based on a Pólya-Gamma latent variable scheme, in which the Pólya-Gamma random variable is approximated by a truncated sum of weighted Gamma random variables. Our current implementation is slow and unable to deal with high-dimensional data, presumably because of the approximation via the truncated sum. There exist another implementation of Bayesian logistic regression, in the programming language STAN (Carpenter et al., 2017), using No-U-Turn-Sampling (Hoffman and Gelman, 2014), which is an extension of Hamiltonian Monte Carlo (HMC) (Duane et al., 1987). An interesting direction for future work, possibly for a master's or PhD student, would be to develop HMC algorithms for $\eta$-generalised Bayesian methods. This could also lead to a better and possibly faster implementation of $\eta$-generalised Bayesian logistic regression.

An issue with generalised Bayesian methods is the dependency on the learning rate parameter $\eta$. Grünwald's 2012 Safe-Bayesian algorithm provably finds the appropriate $\eta$ for bounded excess loss functions and likelihood ratio's, and experiments of Grünwald and Van Ommen (2017) and Chapter 6 indicate that SafeBayes performs excellently in the unbounded case as well, but theoretical guarantees still need to be established. Furthermore, a drawback of the Safe-Bayesian algorithm is that it is computationally very slow. Another future objective is to propose a faster algorithm for learning $\eta$, possibly based on cross-validation, naturally together with theoretical guarantees, e.g. that the data distribution satisfies the central condition at the learning rate $\eta$ output by the algorithm.

Objectives for future work thus are:

- providing a better MCMC sampler for $\eta$-generalised logistic regression, possibly via Hamiltonian Monte Carlo,
- providing MCMC samplers for other $\eta$-generalised GLMs,
- providing guarantees on the Safe-Bayesian algorithm for the unbounded case,
- proposing a faster algorithm than SafeBayes for learning the appropriate learning rate $\eta$, together with
- providing theoretical guarantees for this algorithm.

## 8.4 Pure exploration

In Chapter 7 we studied two Bayesian sampling rules, `TTTS` and `T3C`, for best-arm identification (BAI) in the fixed confidence setting. We introduced the notion of asymptotic $\beta$-optimality and proved that `TTTS` and `T3C` are asymptotically $\beta$-optimal. This optimality notion has two drawbacks. First, in order to be optimal, we would need the unknown true optimal $\beta^\star = \arg\max_{\beta\in[0,1]} \Gamma^\star_\beta$. Secondly, the guarantees are asymptotic, whereas finite-time sample complexity bounds would be more practicable.

Evident objectives for my future work are:

- fixed-confidence guarantees with online tuning of $\beta$ for TTTS and T3C,
- finite-time sample complexity bounds,
- an extension to continuous-armed bandit models (see Section 8.1 above), and
- fixed-budget guarantees.

Furthermore, Chapter 7 provides a piece of the puzzle of the following two bigger pictures.

**Any-time sampling rules** BAI has been studied in different frameworks: the fixed-budget setting, the fixed-confidence setting, which has been studied in Chapter 7, and the *any-time* BAI setting, introduced by Jun and Nowak, 2016. In the any-time setting, the sampling rule does not depend on the risk parameter or the budget. The first sampling rule for BAI that does not depend on the risk parameter is the *tracking rule* proposed by Garivier and Kaufmann (2016). The sampling rules studied in Chapter 7, TTTS and T3C, are also examples of any-time sampling rules. This sparks the question: does there exist a sampling rule that is, albeit with modifications depending on the setting and objective, optimal in all settings? Thompson sampling (TS) could be a possible candidate for this: vanilla TS for regret minimization, TTTS for fixed-confidence best-arm identification, and (see below), Murphy sampling for the minimum of means problem.

**Pure-exploration objectives** Pure exploration problems can have other objectives than finding the best arm. Naturally, different objectives require different sampling rules. However, an interesting avenue for future work is to investigate how the lower bounds and sampling rules for the different objectives and frameworks relate. Here are two pure-exploration problems with objectives different from BAI.

Kaufmann, Koolen and Garivier (2018) study a problem related to BAI: They consider the task of adaptively learning how the minimum mean of a finite set of arms compares to a given threshold. They provide a lower bound on the sample complexity in the fixed-confidence setting, and propose an algorithm inspired by TTTS, called *Murphy Sampling*. Murphy Sampling is, just as TTTS and T3C, an any-time sampling rule. An open problem is to find a fixed-budget lower bound and algorithm for this problem.

Antos, Grover and Szepesvári (2010) and Carpentier et al. (2011) study the problem of estimating the means of a finite number of arms in the fixed-budget setting uniformly well. The objective is to minimise the worst expected squared error loss of the arms, and the performance of the algorithm is measured by comparing its loss to that of the optimal allocation algorithm, that is, *regret*. This notion of regret is however not cumulative, and this problem is therefore more related to the pure-exploration setting than to the standard MAB framework. This is also reflected in the property that good strategies for this problem should play all arms linearly in the number of draws, whereas in the standard stochastic bandit setting suboptimal arms should be played logarithmically in the number of draws. The problem can be extended to learning the transition probabilities of Markov Chains (Talebi and Maillard, 2019). An open problem is to find problem-dependent lower bounds for this problem. Furthermore, the algorithms proposed in both papers depend on the budget and/or the confidence level. An interesting avenue for future work is to find a problem-dependent lower bound and to propose an any-time, possibly Thompson Sampling related sampling rule.