CWI-Inria
INTERNATIONAL LAB

CWI

Inria
inventeurs du monde numérique

Workshop 2020
September 24, 2020, online

Rianne de Heide **(CWI)**
6PAC Associate team

Title: « Bayesian Best-Arm Identification »

Abstract:

In multi-armed bandits, a learner repeatedly chooses an arm to play, and receives a reward from the associated unknown probability distribution. We study the task of best-arm identification (BAI), where the learner is not only asked to sample an arm at each stage, but is also asked to output a recommendation (i.e., a guess for the arm with the largest mean reward) after a certain period. Unlike in another well-studied bandit setting, the learner is not interested in maximising the sum of rewards gathered during the exploration (or minimising regret), but only cares about the quality of her recommendation. We investigate a Bayesian-flavoured sampling rule called Top-Two Thompson sampling (TTTS). In particular, we justify its use for fixed-confidence BAI. We further propose a variant of TTTS called Top-Two Transportation Cost (T3C), which disposes of the computational burden of TTTS. As our main contribution, we provide the first sample complexity analysis of TTTS and T3C when coupled with a very natural Bayesian stopping rule, for bandits with Gaussian rewards, solving one of the open questions raised by Russo (2016). We also provide new posterior convergence results for TTTS under two models that are commonly used in practice: bandits with Gaussian and Bernoulli rewards and conjugate priors.