Gerhard Schurz (Düsseldorf): Meta-inductive Probability Aggregation

The choice of a suitable priori probability distribution is a deep question of Bayesian epistemology. Not all priors can be outwashed by evidential updating in the long run. A counterexample is the uniform prior over possible world which constantly predicts averages; yet there are environments in which this prior leads to maximal predictive success. Even if one restricts oneself to priors which are continuous over the space of possible frequencies (or frequency limits) and thus are able to learn inductively, different priors of this sort can make a big different for the predictive success in the short run. Moreover there is the question whether the observed random sequence is an IID sequence or a Markov chain, and whether conditionalization on external events can increase predictive success.

In the talk it is suggested that meta-inductive probability aggregation can offer a new solution to this problem. A probabilistic prediction game is a pair ((e),PR), consisting of a sequence (e) of events $e_1,e_2,\ldots$ and a finite set PR = $\{P_1,\ldots,P_m\}$ of probabilistic predictors or methods which are identified with their probability functions. At each time point n the task of every predictor P_i consists in delivering a (conditionalized) probabilistic distribution for the event of next round prediction of the next round. The predictions are scored by a proper loss function; this guarantees that predicting the correct probabilities will maximize the score.

In different environments, different methods will be successful. The strategy of meta-induction can be utilized to define an optimal aggregation of the different probabilistic methods, in the form of a weighted average $P_{MI,n+1} = \Sigma_{1\leq i\leq m} w_{i,n}\cdot P_{i,n+1}$, with weights $w_{i,n}$ that depend on the achieved success rates of the methods. P_{MI}'s predictive success is guaranteed to be optimal in the long run among the pool of methods $\{P_i{:}1\leq i\leq m\}$, with tight upper bounds for possible losses that quickly vanish with increasing number of rounds. From the aggregated conditional distribution of MI the prior distribution of MI can be calculated 'post-facto'.