Turning ON (and OFF)

John Skilling
Maximum Entropy Data Consultants Ltd.
Killaha East, Kenmare, County Kerry, Ireland

Abstract

The frontier of Bayesian methodology lies in algorithm technique. Despite great advances in recent decades, it is the nature of algorithms, rather than any theoretical difficulty of principle, that continues to limit our progress. Perhaps surprisingly, very simple problems continue to be informative here.

One of the common themes of quantitative science is to infer from observations of an object what was present, and how much of it was there. This problem can be formalised in terms of ON/OFF switches for the possible components, allied to quantitation variables in which the observations are often linear.

Exploring the plausible patterns of switching involves digital choices. Conventionally, these choices are described as transition probabilities, but there is an alternative interpretation in terms of a competition between transition engines. This re-interpretation is useful because it leads to a natural way of reducing the rejection steps which tend to plague Bayesian computations. As for quantitation, the most popular technique is Gibbs sampling, but marginalisation is also possible, and offers great advances in efficiency.

By its nature, the digital ON/OFF problem lacks continuity, and the posterior may well break up into distinct islands. This is commonly held to be a cause of severe difficulty. However, the trick of killing off poorly-performing samples samples in favour of better ones allows a program to evade these difficulties to a remarkable degree.

In similar vein, large-scale problems of this type can easily be subject to phase changes which inhibit conventional "annealing" algorithms. After all, there is seldom any guarantee of "nice" behaviour. Indeed, one particular first-order phase change is a fundamental part of the natural Bayesian model of quantitation. The solution to this difficulty is to use the new technique of nested sampling, which explores the state-space more directly, and is incidentally rather easier to implement.

In short, the 21st century has much to say about even a very basic problem.