Contribution to discussion of paper by Lauritzen and Richardson RSS Ordinary meeting, 12 December 2001

Peter Green (University of Bristol)

The authors' thesis that their data generation processes provide valid interpretations of chain graphs, supporting intervention, is characteristically original and intriguing; central to this is the relationship between intervention by replacement in a data generation process with ordinary conditioning in the equilibrium distribution, and that is what I want to comment on.

This relationship is discussed in Sections 6.3 and 6.4, and directly extends that for undirected graphs in Section 6.2; all I want to say about the extension to chain graphs is that surely convergence to equilibrium could occur concurrently rather than sequentially across the chain components?

Section 6.2 provides helpful examples, rather than a full analysis of the connection between intervention and conditioning; the present discussion is only slightly less incomplete. From these examples, we get the impression that reversibility is the key. The first four processes listed – the systematic and random Gibbs samplers, the time-reversible Markov dynamics and the Langevin diffusions – are all reversible (in the case of the first, at least at the level of individual updates). It is easy to contrive examples of nonreversible generating processes where (14) fails, that is, where intervention by replacement does not lead to ordinary conditioning. However, further study shows that reversibility is not enough.

The vector diffusion (12), $X(t + \mathrm{d}t) = X(t) + CX(t)\mathrm{d}t + \mathrm{d}Z(t)$ with $\mathrm{var}(\mathrm{d}Z(t)) = \Lambda \mathrm{d}t$, is useful in exploring things further. When $\Lambda = I$, this is reversible if and only if C is symmetric, precisely the requirement for (14) to hold, as identified by Proposition 5. But for general Λ, the process is reversible if and only if $C\Lambda = \Lambda C^T$. As long as Λ is diagonal, the impact is straightforward – Proposition 5 can be easily modified, and reversibility continues to imply (14). However, if Λ is not diagonal, this breaks down: you cannot simultaneously have (14) and obtain the correct equilibrium without intervention. Diagonality of Λ is the same as saying that perturbations to the individual variables are (conditionally) independent.

A similar complication arises if you take a broader view of the Gibbs sampler, allowing block updates and directional sampling: again, (14) holds only when the perturbations are independent. When perturbations are independent, intervention by replacement is equally *intervention by conditioning*, so the connection to ordinary conditioning in the equilibrium is perhaps unsurprising.

Curiously, the discrete-time version of (12) behaves differently: suppose $X_{t+1} \sim N(AX_t, \Lambda)$. Reversibility ($A\Lambda = \Lambda A^T$) and independent perturbations (Λ diagonal) are not enough: for (14) you need A diagonal as well, when the system decouples completely. Perhaps the advantage of continuous- over discrete-time processes is of more than 'intuitive appeal'?