# COMMENTARY ON GRENANDER & MILLER:
## Representations of Knowledge in Complex Systems

*David Mumford*
*Department of Mathematics, Harvard University*

Tonight's paper introduces an intriguing algorithm which combines the small stochastic steps of simulated annealing algorithms with large 'jumps' to produce a Markov chain sampling a complex Gibbs field of the type encountered in vision problems. This paper is remarkable in coming to grips with the necessity of making large jumps which change the topology of the sought-for pattern, while also making small incremental improvements with a fixed topology. Shah and I proposed something of this kind [Mumford & Shah, 1985], but had been defeated by its complexity.

One noteworthy antecedent is the work of Brandt, Ron & Amit [1986] on a multi-grid algorithm for sampling the Ising model with external field. In fact, their random variables, the spins $x_\alpha = \pm 1$, are equivalent to those of the authors when there is only one type of organelle and only cyclic graphs. Brandt *et al* employ a hierarchy of moves, in which larger and larger square blocks of pixels are flipped. In the authors situation, this could introduce a new organelle or radically reshape the partition, chopping a big block out of an organelle, etc. Brandt's method, however, *doesn't accept or reject a move immediately*, but works back down the hierarchy, improving the result with smaller flips, before deciding.

This raises an issue in the present paper which is not stressed: it seems to be computationally intractable to actually calculate the transition probabilities $Q(x, dy)$ because of the need to sum over *all* possible new curves $y(1)$ that might be introduced. These new curves can only be sampled, and the illustrations suggest that even more drastically, only small circles were considered. It is exactly the need to sample well the more probable new curves that forced Brandt *et al* to a complex algorithm which delayed acceptance/rejection.

A logical route is to mimic *genetic algorithms*, and rather than entertain one or a series of global moves on *one* sample, consider a small population of samples simultaneously. The moves are now the stochastic evolution of the individual members of the population, and the splicing of parts of one sample with parts of another. In the authors paper, this is especially simple:

we may combine two patterns $x(m_1), x(m_2)$ by taking $k_1$ of the objects $c(j)$ in $x(m_1)$ together with $k_2$ of the objects in $x(m_2)$ and forming a new pattern $x(k_1 + k_2)$ out of their union. I believe this will often be faster and more effective.

Brandt A, Ron D & Amit D, 1986, Multi-level approaches to discrete-state and stochastic problems, Lecture Notes in Math, 1228, pp. 65-98, Springer-Verlag: New York.

Mumford D & Shah J, 1985, Boundary Detection by Minimizing Functionals I, in *1985 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Science Press: Los Alamitos CA.