Limitations of learning with distributed representations

Horace Barlow[1] and Tony Gardner-Medwin[2]

Workshop on Neuroinformatics, Potsdam, June 24-27, 1998

[1]Physiological Laboratory, Cambridge CB2 3EG
[2]Departmentof Physiology, University College London, London WC1E 6BT

Many years ago the "redundancy reduction hypothesis" suggested that the statistical structure of the natural world is exploited in sensory coding, but further progress has been hampered by inability to specify the cost, or negative survival value, of different possible ways of communicating and representing sensory information in the brain. In man-made systems the increased cost of inappropriate coding results mainly from wasted channel capacity, but we think the major biological cost is that it increases the amount of data required to discover the associations present in the environment - ie it hinders learning.

Finding out whether two events are associated requires counting their joint and separate occurrences, but there is a problem in doing this for the distributed representations thought to occur in the brain. This is because a given representational element is active for different environmental events, which introduces both random and systematic errors. We show that to minimise these errors the representations should a) be sparse, with elements active at a small fraction of their maximum rate; b) represent rare events by states with a higher activity ratio than common events; c) represent biologically important events with a higher activity ratio than unimportant events, and d) represent distinct events with uncorrelated sets of cells or sets that are correlated to reflect similarities in their associations.

These rules seem to be obeyed (at least approximately) in higher sensory representations, so the original hypothesis needs modifying to say that external redundancy from the environment is converted (rather than simply reduced) to create a sparse perceptual representation that helps rather than hinders efficient learning.