p7q6
JGD

Model Answer – Neural Computing, Question 1. 2000

(A)
A Hopfield content-addressable, associative memory allows patterns to be stored in such a way that each one can be retrieved by using just a portion, or a corrupted version, of it as its "address." The memory functions dynamically by detecting an association between an input pattern and a stored memory, which it then retrieves in its entirety.

The network consists of N "neurones" that are fully connected to each other. Their connections have weights of adjustable strengths, defining an NxN connectivity matrix. This matrix of weights is adjusted during a learning phase that stores the patterns in memory. Each of the N neurones has a binary state, and the N-dimensional space of their states is called the configuration space. A stored "memory" is just a particular point in this N-dimensional configuration space. The training phase consists in setting all the NxN weights such that a particular memory is a stable state consistent with all the other stable states, so that once the network reaches such a state, no neurones will change their state further. Then, whenever a new input pattern first appears (implemented by re-setting the states of the N neurones), the collective state of the network will evolve by each neurone following a threshold rule: The $N-1$ inputs that it receives from all the other neurones (each of whose state is either a $+1$ or a -1), each multiplied by the corresponding weight for that connection, are all added together and compared against a threshold. If this inner product (sum) exceeds the threshold, then the neurone's state is set to $+1$; else it is set to -1. This process reiterates for all the N neurones until a stable state is reached, when no neurone changes its state anymore. This collective state is called an stable attractor. Each stable attractor represents one memory. The range of different input states that will all converge eventually to a particular such stable attractor is called the basin of attraction for that memory. Because the states within this basin of attraction represents similar, but corrupted (or partial) versions of the stable state, a Hopfield memory is capable of "completing" a memory from just parts of it serving as a trigger, rather like happens in human memory. Hence it is deemed content-addressable and associative. It is also capable of overcoming noise and corruption of an input pattern, provided that the input remains within the basis of attraction of the attractor. The network capacity of a Hopfield network is about $(0.15)N$, where N is the number of neurones. This capacity would be greater if only orthogonal patterns were stored.

[10 marks]

....Continued....

(B)

1. The purpose of vision is not to reproduce faithfully the 2D retinal image, but rather to construct an internal 3D model of the surrounding world and of the 3D objects that populate it. In this regard, vision is a kind of "inverse graphics." (In graphics, a 3D world model is projected into a 2D screen image; in vision, just the reverse must be accomplished.) Evidence that what we see is our own "graphics:"

- Perceptual size invariance: all the faces in a lecture theatre appear to be roughly the same size. Yet in terms of the 2D retinal image, some faces may actually be 20 or 30 times larger than others, depending on their distance.
- Motion compensation: every eye movement causes the retinal image to shift ballistically, in tremendous sweeps and jerks. Yet the world appears to be stable.
- Homogeneity of spatial resolution: high visual resolution exists only within the fovea (the central 1 or 2 degrees). Yet the world appears to have uniform resolution everywhere.
- Homogeneity of colour signals: the colour-sensitive photoreceptors ("cones") exist only near the fovea; outside the fovea, only black and white information is transduced. Yet the world appears uniformly full of colour.
- Invisibility of the silhouette of the retinal blood vessels: the retinal image is corrupted by a dense tree of opaque blood vessels, lying in front of the actual photoreceptors. Yet we discount their shadows, and "fill-in" the gaps with graphics. Similarly, the blind spot is filled in; we are not aware of this large hole.
- Colour constancy: a banana continues to look yellow, whether reddish light or bluish light is shown on it. We see the spectral reflectance properties of the pigmented surface, rather than the actual wavelength mix received on the retina.
- Interpreted scene properties: we perceptually interpret events like occlusion, or motion and rotation in depth, in terms of solid 3D objects and configurations.

One implication of these observations for machine vision is that we should cast the problems of vision as "inverse problems," for which the goal is to invert the processes of graphics (projections, ray-tracing, radiosity, etc) to try to arrive at the unique 3D world situation that could produce such a 2D signal as the one received on the retina.

[8 marks]

2. The geometrical illusions (distortions of form, angle, length, shape) may be inevitable consequences of trying to achieve such goals. They may reveal, for example, the existence of near-range competition and far-range cooperation in the processing of orientation information. Mere fidelity to the retinal image should *not* be a primary goal of vision, as it might be in a sound reproduction system. The goal is to *understand* the world; who cares if this *changes* its appearance (inverting Karl Marx...)

[2 marks]