A simpler explanation for how binding neurons develop

Consider a linked set of three neurons at different stages of the ventral visual pathway:

(i) Neuron 1 (in a lower visual layer) represents a low level visual feature.
(ii) Neuron 2 (in a higher visual layer) represents a high level visual feature.
(iii) Neuron 3 is a hidden neuron within a local layer, say the same layer as either neuron 1 or 2, which may learn to become a binding neuron.

Assume that there are the following three synaptic connections between these three neurons:

(i) A connection from neuron 1 to neuron 2. (This is the traditional bottom-up type of connection that has always been incorporated in VisNet.)
(ii) A connection from neuron 1 to neuron 3. (This could be either a lateral or bottom-up connection depending on which layer neuron 3 is in.)
(iii) A connection from neuron 2 to neuron 3. (This could be either a lateral or top-down connection depending on which layer neuron 3 is in.)

Denote the axonal delay from neuron j to neuron i as dt(i,j).

If and only if neuron 1 is participating in driving neuron 2, then each spike emitted by neuron 2 will usually occur dt(2,1) after a spike emitted by neuron 1.

If we have a set of three axonal delays such that

$dt(3,1) = dt(2,1) + dt(3,2)$ Eqn (1)

then the spikes from neurons 1 and 2 will converge on neuron 3 simultaneously if and only if neuron 1 is participating in driving neuron 2.

Assume the hidden neuron 3 operates as a 'coincidence detector' (with short time constants), and fires only when the volley of spikes from neurons 1 and 2 arrive simultaneously. In this case, neuron 3 will behave as a binding neuron. That is, neuron 3 will fire if and only if neuron 1 is participating in driving neuron 2.

Spike time dependent plasticity (STDP) will further strengthen the connections from neurons 1 and 2 onto the binding neuron 3.

But we have begun by fortuitously assuming the required relationship in the axonal delays given by Eqn (1). However, in fact, what we assume is that there are hundreds of thousands of nearby hidden neurons k with random axonal delays dt(k,1) and dt(k,2). A small subset of these hidden neurons will fulfill the condition of Eqn (1). This subset of hidden neurons will become binding neurons that fire if and only if neuron 1 is participating in driving neuron 2.

It is also possible that neurons 1 and 2 each send several connections to the hidden neuron 3 with random different axonal delays. In this case, STDP will strengthen the subset of these connections that result in the spikes from neurons 1 and 2 arriving simultaneously at neuron 3 (if and only if neuron 1 is driving neuron 2). This is another mechanism by which the binding neurons may develop their selectivity.

These mechanisms by which binding neurons may develop are very simple. Although the mechanisms utilize an SNN model with distributions of randomized axonal delays, it does not require the full self-sustaining attractor dynamics of polychronization because the system will be driven by external visual input.

It is hypothesised that a large spiking neural network model with a distribution of randomized axonal delays will develop many such binding neurons if the model incorporates the following elements:

(i) Bottom-up, top-down, and lateral connections;
(ii) A randomised distribution of axonal delays;
(iii) Spike time dependent plasticity;
(iv) Diluted synaptic connectivity, leaving many hidden neurons to be driven primarily by lateral and top-down connections;
(v) Inhibitory interneurons mediating competition within layers, and therefore competitive learning needed for binding neurons to refine their selectivity by STDP.

The binding neurons that develop would carry measurable information about which low level features are driving (and hence part of) which high level features. Binding neurons would develop at every layer of the feature hierarchy, and for every spatial scale within a natural visual image.

This argument seems to suggest that binding neurons will develop automatically in a biologically realistic neural network. They may be one important reason why the visual system has extensive lateral and top-down connections in addition to the bottom-up connections that are traditionally modeled.

These binding neurons would only develop within a biological neural network architecture; engineering networks such as backpropagation of error would not develop such binding neurons and so cannot solve binding.

Binding information is essential to the rich semantic analysis and interpretation of visual images performed by the human brain.