



## Activation and coherence in memory processes: revisiting the Parallel Distributed Processing approach to retrieval

ANTONINO RAFFONE\* and CEES VAN LEEUWEN†

\*Department of Psychology, University of Sunderland, St Peter's Campus, Sunderland SR6 0DD, UK

email: antonino.raffone@sunderland.ac.uk

†RIKEN Brain Science Institute, 2-1 Hirosawa, Saitama 351-0198, Japan

email: ceesvl@brain.riken.go.jp

*Abstract.* Connectionist models based on activation spreading and attractor dynamics are functionally limited by representational and processing flexibility constraints, the 'feature binding problem' and the need to balance accurately activation and inhibition. We suggest an alternative approach, in which network units are characterized by two variables: *activation* and *phase*. Whereas activation evolves according to a 'classical' connectionist rule, the phase variable is characterized by a chaotic evolution. We present a model of memory retrieval with reference to the paradigmatic McClelland's 1981 'Jets and Sharks' model. The model solves the 'multiple reinstatement problem', i.e. the problem of retrieval of multiple items with overlapping features, implied by its classical predecessor. In our network, multiple pattern reinstatement in terms of activation spreading is disambiguated through selective and differential coherence patterns. The system flexibly represents pattern similarity and feature relationships by means of graded and intermittent synchrony. The domain-general implications of this approach for connectionist 'interactive activation models' and its neurophysiological plausibility are discussed.

*Keywords:* memory retrieval, neural network, connectionism, graded synchronization.

### 1. Introduction

#### 1.1. Connectionist modelling of memory retrieval

Most connectionist or Parallel Distributed Processing (PDP) models assume that memory retrieval occurs through pattern reinstatement and completion in activation spreading processes (e.g. McClelland *et al.* 1986b, Rumelhart *et al.* 1986, McClelland *et al.* 1995). According to these models, when a subset of the features (elementary traces) of a composite representational pattern is reactivated, activation spreads to the remaining constituent features. When an input-cue is insufficient to lead to retrieval of a stored pattern, fragments of the pattern may be reinstated (McClelland *et al.* 1995). For instance, this process may be involved in *source memory failures* occurring when the representational fragments of an episode are retrieved without a recollection of how or when the fragments were acquired (Schacter 1989, Squire 1992, Schacter *et al.* 1998).

Such content-addressable memory models with distributed representations are robust in comparison to memory retrieval systems based on symbolic labels, and exhibit dynamics preventing a representational combinatorial explosion. In attractor-based retrieval networks, pattern reinstatement occurs through the convergence on

to an attractor, representing the information stored in an activation space (Hopfield 1982, Amit 1989, 1995, van Orden *et al.* 1990). The conspicuous virtues of such dynamics are well known (McClelland 1981, McClelland *et al.* 1986a): graceful degradation (even incomplete or partial probes can lead to retrieval); default assignment (inferences based on known similarities between instances); and spontaneous generalization (prototypes may be inferred, even if no actual presentation shares all the typical features). Because of these functional properties, attractor dynamics became prevalent as connectionist models of retrieval.

These models prevailed, despite their conceptual and dynamical limitations which result from the inertness of their dynamics. Attractor memory systems are characterized by limited autonomy, since retrieval processes in terms of convergence on to fixed point attractors are ultimately restricted to a set of stimulus–response pattern associations.

Attractor-based memory systems are also characterized by a limited dynamical flexibility. The working range of these models is critically dependent upon the balance of excitatory and inhibitory interactions. Insufficient inhibition during activation spreading will lead to the activation of all units within the system, irrespective of whether these units code for a target or not; by contrast, an excessive inhibition may cause activation to die out before the target-units are activated. Maintaining the appropriate excitation–inhibition balance is a non-trivial task in the cerebral cortex (see Braitenberg and Shüz 1991). If learning of new items is to be based on the strengthening of excitatory connections, learning will disturb this delicate balance. Compensation by proportionally increasing the overall level of inhibition is implausible, as connections between excitatory (pyramidal) and inhibitory (stellate) cells are generally believed not to be plastic.

Finally, ‘classical’ connectionist models assume that a single variable codes for retrieval in working memory: the activation of features belonging to a given composite representation or multi-feature pattern. In realistic memory retrieval conditions, however, features tend to act as clusters within a cluttered environment, rather than as isolated or well-segregated subsets. For instance, multiple feature clusters could be active simultaneously, working as cues for distinct patterns to be retrieved. We refer to this problem as *multiple reinstantiation problem* in memory retrieval.

For example, let us to consider the classical ‘Jets and Sharks’ retrieval model (McClelland 1981, McClelland *et al.* 1986a). In the original connectionist model (figure 1), the members of the different gangs (the Jets and the Sharks) are represented in terms of their features. When a subset of features is presented, a given individual is recognized as a consequence of activation spreading. In realistic circumstances, however, gang members are usually not encountered alone. In each other’s company, they will provide retrieval cues that are likely to ‘cross over’, leading to a number of spurious activation states and confusion as to which individual each activated feature belongs. Such representational ambiguity in retrieval constitutes a problem for the classical connectionist framework.

### 1.2. *The binding problem at different representational and processing levels*

The multiple reinstantiation problem may be regarded as a version of the well-known binding problem (von der Malsburg 1981, Engel *et al.* 1992, Singer 1994, Gray 1999). In perceptual areas of the cortex, the binding problem occurs when neurons or neural assemblies within the same network code for features of more than one object in terms of activation level (or firing rate).

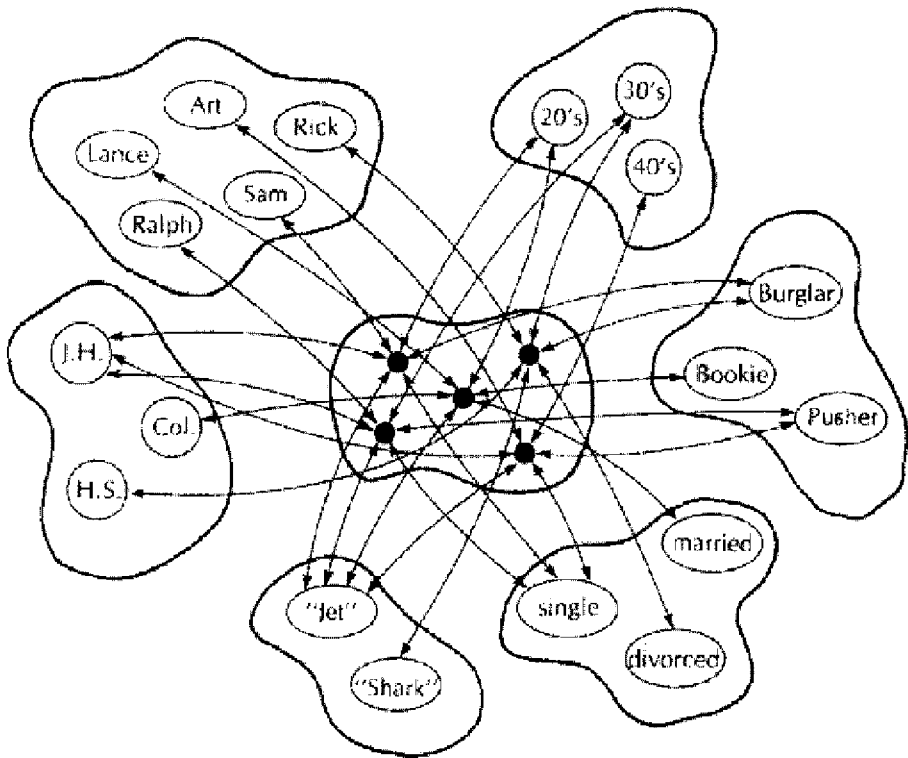


Figure 1. The units and connections for five of the imaginary individuals in McClelland's (1981) Jets and Sharks network. Note the central 'gang member units' and the surrounding 'characteristic units'. (Reproduced from McClelland 1981, with permission.)

In the neurosciences, the classic approach to neural coding is based on the assumption that information is carried by neuronal firing rates over time-scales of hundreds of milliseconds. Firing rates, however, may not adequately disambiguate neural responses when the features of a pattern have to be integrated into a coherent representation (for instance, a visual object), while being segregated from other features (for instance, features belonging to other objects). When a visual cortical network is simultaneously presented with a red square and a green circle, with enhanced firing rates of the neurons coding for shapes and colours, the bindings of *red* and *square*, and of *green* and *circle*, cannot be discriminated in the absence of specific conjunctive representations. A strategy based on an explicit representation of conjunctive features will ultimately lead to a combinatorial explosion (von der Malsburg *et al.* 1981, Engel *et al.* 1992).

It has been suggested that this binding problem may be solved in terms of *temporal coding*, based on the selective synchronization of time-resolved neuronal responses (von der Malsburg 1981, Eckhorn *et al.* 1988, Engel *et al.* 1992, Shastri and Ajjanagadde 1993, Singer 1994, Roelfsema *et al.* 1996). According to this view, the action potentials (spikes) of neurons coding the features of the same object are synchronized, while being uncorrelated to (or actively desynchronized from) the responses of neurons coding the features of other objects. In support of this view, there is evidence that neurons act as coincidence detectors: synchronous synaptic inputs are more effective than asynchronous ones in eliciting spikes of the neurons on

which they converge (Abeles 1991). Temporal neural codes, therefore, may be *read out* in terms of coincidence detection properties.

Modelling studies have demonstrated that temporal synchrony can be used in high-level binding for drawing inferences (Shastri and Ajjanagadde 1993, Shastri and Mani 1997, Shastri 1999). Temporal synchrony may also be involved in handling different instances of the same concept simultaneously, in the so-called multiple instantiation problem (Mani and Shastri 1993, Sougné 1998, Shastri 1999). As these studies indicate, given the appropriate architectures temporal coding may be effective in high-level cognitive processes involving dynamic binding and complex relations between active representations.

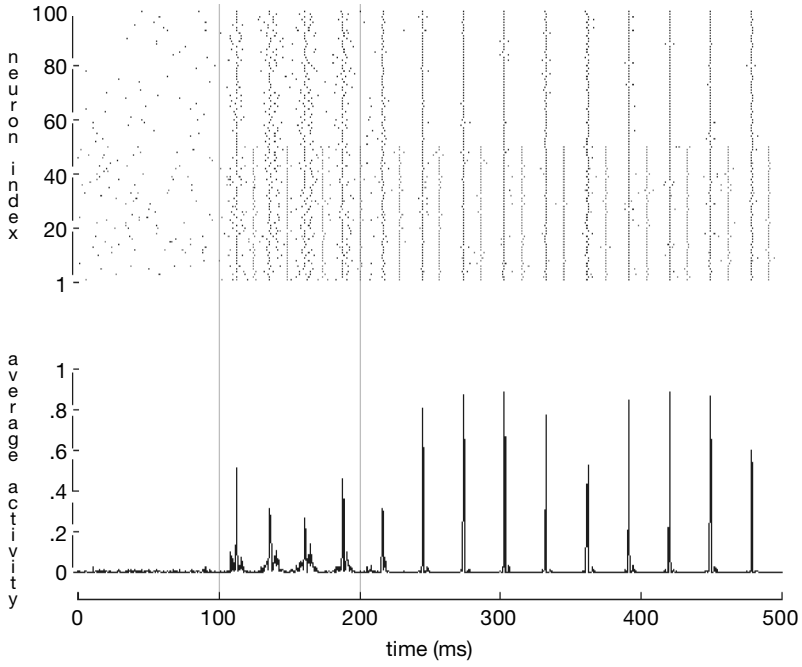
In memory retrieval, the binding problem takes the form of the aforementioned multiple reinstantiation problem. During the retrieval process, cortical neurons or neural assemblies will become activated that code for more than one neural representation of an entity or episode.

Binding processes may also plausibly be involved in memory encoding. Also in encoding, binding has to be achieved while avoiding representational interference problems due to cross-talk (Schacter 1989, Moscovitch 1994, Schacter *et al.* 1998). In encoding, features are linked together according to their episodic, semantic and intrinsic perceptual-structural organization.

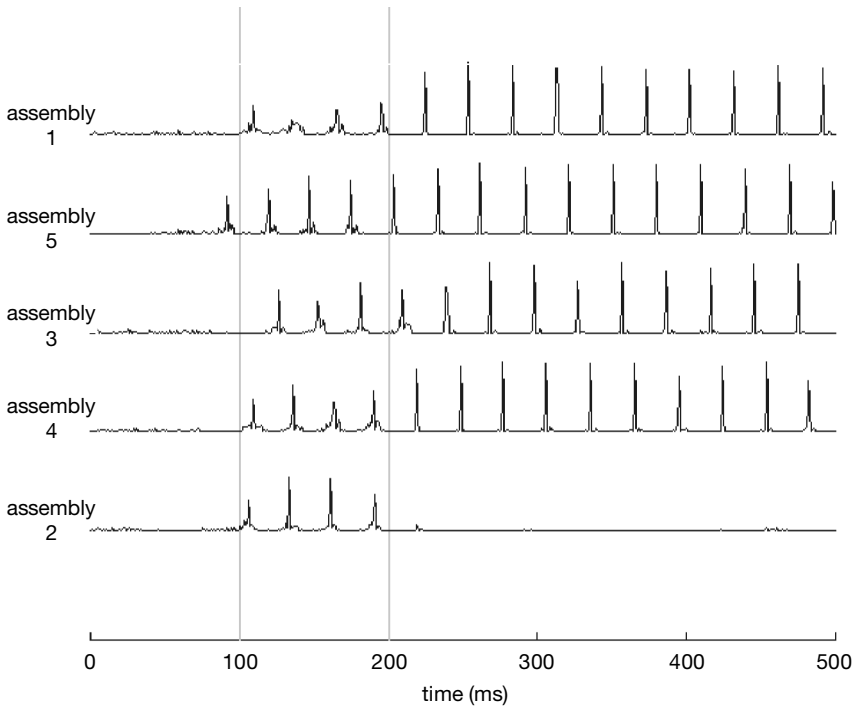
It is commonly agreed that episodic (e.g. Tulving 1985) and semantic organization play important roles in memory encoding (e.g. Miller 1956). The role of structure in memory encoding was studied in a series of articles in the 1970s (Restle 1970, 1976) and 1980s (Geissler and Puffe 1983, Buffart and Geissler 1984, Buffart 1987, van Leeuwen *et al.* 1988, van Leeuwen and Buffart 1989). Geissler and Puffe (1983) and van Leeuwen and Buffart (1989) showed that, respectively, hierarchical and serial patterns are encoded in terms of underlying features of their perceptual structure. There is evidence that the binding processes of memory encoding continue after perception. For instance, in early learning stages visual patterns are encoded in terms of their independent sub-structures, but with increased practise they become stored in terms of more integral features (van Leeuwen *et al.* 1988), and classified accordingly (Goldstone and Medin 1994). The role of visual working memory may be to assemble representational patterns, thus binding increasingly complex sets of features. Shiffrin (e.g. Czerwinski *et al.* 1992) called this process ‘unitization’. In this process, a pattern that is complex at a certain level of representation is coded as a simple pattern in a higher-order feature space.

These findings suggest that binding processes operate in visual working memory processing to serve integration. In an elegant series of studies Luck and Vogel (1997) showed that the storage capacity of visual working memory corresponds to approximately four objects. Significantly, the retention capacity does not depend on the number of features making up the objects. Hence, visual working memory seems to process integrated objects rather than individual features, just as verbal working memory handles higher-order ‘chunks’ instead of individual features or letters.

Raffone and Wolters (2001) have recently developed a neural model of retention, limited capacity and chunking in visual working memory, based on the synchronization and desynchronization of reverberating neural assemblies. The role of oscillatory patterns in verbal short-term retention was put forth by Lisman and Idiart (1995). Raffone and Wolters’ model (figure 2) can parsimoniously account for both the limited capacity of visual working memory and for the temporary binding of multiple assemblies into a single pattern (Raffone and Wolters 2001, Raffone *et al.*

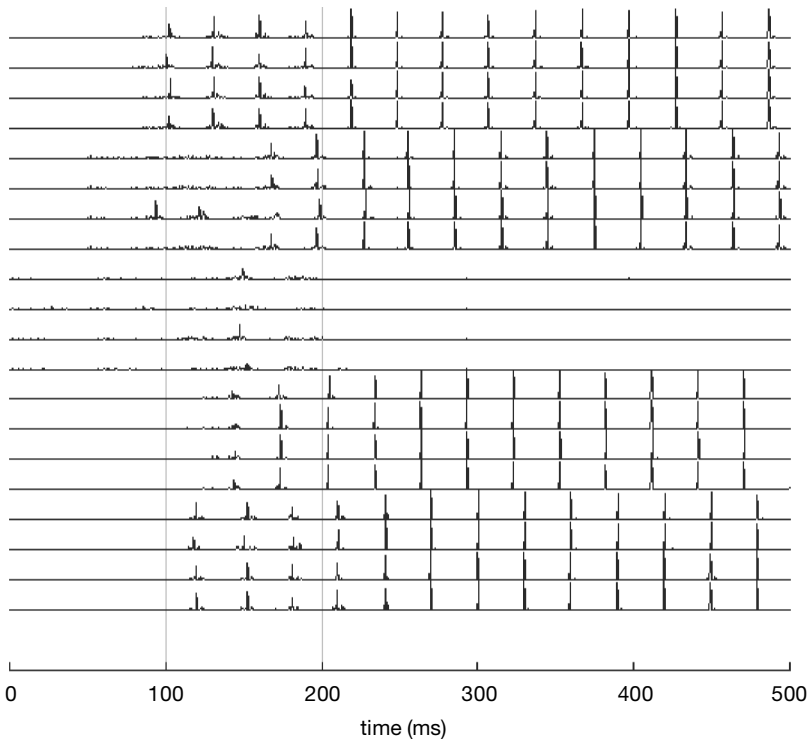


A)



B)

*continued . . .*



C)

Figure 2. Retention, limited capacity and binding in Raffone and Wolters' (2001) model. Individual assembly behaviour with feature input and active feedback. (a) The top part of this panel shows the spike-raster of the 100 neurons of a neural assembly in IT (inferotemporal cortex) and of the 50 neurons of a reciprocally-connected prefrontal assembly. Spikes of the two assemblies are coded by different grey tones. The bottom part of (A) shows the evolution of the average activity of the IT-assembly. Stimulus onset and offset times are marked by the vertical grey lines. Note how oscillatory neuronal firing is maintained after the stimulus-offset, due to feedback from the prefrontal cortex. (B) Phase segregation of IT-assemblies coding for simultaneously presented single-feature objects. Four out of five reverberations remain active. Due to mutual inhibitory activity, the assemblies become spaced in the oscillatory phase, thus allowing a markedly discriminative oscillatory reverberation and retention of the coded objects. Assemblies are shown in an order allowing easy inspection of phase segregation. (C) The combination of within-chunk integration and between-item segregation. Three four-feature objects are retained in terms of three internally synchronized (bound) and mutually desynchronized oscillatory chunks, whereas all the four assemblies coding the features of a fourth object are suppressed. (Reproduced from Raffone and Wolters 2001, with permission.)

2001). A critical capacity of about three to four independent patterns was observed in the simulations, consistent with the results of Luck and Vogel (1997). According to this model, oscillatory activity and selective synchronization play a crucial role in visual working memory representation and processing.

Interestingly, several brain imaging studies (e.g. Buckner 1996, Haxby *et al.* 1996) suggest that the same areas and neural representations that are active during the retention of items in working memory are also active during retrieval from long-term memory. Thus, it may be hypothesized that uniform neurocomputational mechanisms

are operating in perceptual binding, within-chunk integration and between-item segregation in visual working memory (Raffone and Wolters 2001), as well as in multiple reinstatement during memory retrieval. This view is consistent with Phillips and Singer's (1997) proposal of common codes involved in different forms of cortical information processing. Large-scale synchronization processes in cortical networks may play a crucial role in these multifunctional neural computations.

## 2. The graded synchrony hypothesis

Several recording studies from cat and monkey visual cortex have demonstrated that the selective synchronization of oscillatory neuronal discharges is involved in visual grouping and segregation (e.g. Eckhorn *et al.* 1988, Gray *et al.* 1989, Castelo-Branco *et al.* 2000), as well as in binding sensory and motor responses (Murthy and Fetz 1992, Roelfsema *et al.* 1997). Recent EEG studies suggest that oscillatory neural coherence may be related to associative learning and large-scale cognitive integration (Miltner *et al.* 1999, Rodriguez *et al.* 1999), not just to local feature binding. Simulation studies have shown that phase coherence may be translated into behavioural outcomes based on the effectiveness of correlated neural states (Tononi *et al.* 1992).

Although the recording data in many of the electrophysiological studies indicate non-stationary or intermittent coherence episodes, the underlying cortical mechanisms have usually been modelled in terms of networks of phase-locked neural oscillators with a periodic dynamic behaviour (e.g. Sporns *et al.* 1989, König and Schillen 1991, von der Malsburg and Buhmann 1992). Generally, in these models synchronization is mediated by excitatory synchronizing connections, and desynchronization is induced by random noise. However, these models present significant functional limitations: (1) the desynchronization time of the initial phase is relatively long, being proportional to the number of oscillators in a certain assembly and inversely proportional to the squared amplitude of local noise; (2) even (large) weakly coupled assemblies can eventually synchronize their phase (see also Hansel and Sompolinsky 1992); and (3) assuming that a given oscillatory phase codes the set of features belonging to a given object, it is hard to explain phase constancy when the related image size and retinal position are changing (Grossberg and Grunewald 1997).

Several studies have shown that functional limitations of coupled sinusoidal oscillators are not inherent to networks of relaxation oscillators, in the case of long-range co-operation and short-range competition (Somers and Kopell 1993), as well as local co-operation and global competition (Wang and Terman 1997). In both kinds of cortical network models with periodic oscillators, neural synchrony is interpreted as an 'all-or-none' phenomenon, implying a transitive relationship: if a neuron (or assembly) A is synchronized to a neuron B and to a neuron C, then the neurons B and C are mutually synchronized as well (Neven and Aertsen 1992). Thus, the stable synchronization of periodic neural oscillators *cannot* be used for associative coding with multiple active patterns (if their overlap is higher than zero), i.e. neurons cannot code for different objects in the visual scene, and cannot flexibly participate in almost simultaneous but disjoint neural computations. Moreover, this synchronization can also reduce the amount of information that can be encoded, since perfectly synchronous firing in a pool of neurons signals a single event (Ritz and Sejnowski 2000).

Owing to phase-locking into fixed orbits, the stable synchrony of periodic neural

oscillators *cannot* be effective in the multistable and reversible perception of ambiguous patterns, i.e. visual patterns that may be dynamically perceived in multiple ways (e.g. Kubovy and Wagemans 1995, Leopold and Logothesis 1999). By contrast, intermittent patterns of synchronized (non-stationary) activity may enable the integration of local features into global representations, as well as fast switching between alternative global representations (van Leeuwen *et al.* 1997, 2000). Thus, binding dynamics in perceptual neural networks must exhibit dynamic behaviours that enable the intrinsic non-stationarity and multistability of perceptual experiences to be captured.

A flexible temporal coding in neural systems is especially needed in high-level areas of the visual system due to the large receptive fields. Such a coding may be obtained through *graded synchrony* (GS) of non-stationary (e.g. chaotic) signals, i.e. in terms of degrees of inter-neuron synchronization (van Leeuwen *et al.* in press, Raffone and van Leeuwen submitted). Distances in feature spaces, i.e. proximity and similarity relationships, as well as the variable associative strengths between active neural representations, may be effectively coded by relative degrees of synchrony among neurons or neural assemblies.

This type of coding may be at work, for instance, when a network of visual neurons with relatively large receptive fields is presented at the same time with a red square and a red circle (see figure 3 for a related scheme). A transitive synchrony relationship between the periodic signals coding for *red*, *square* and *circle* would lead to the indiscriminate synchronization of all the active neurons. Alternatively, a graded synchrony pattern in which a neuron coding for *red* is synchronized at different times (in an intermittent manner) with the neurons coding for *square* and *circle* would solve the binding problem in the presence of overlapping features. The multiple reinstatement problem in memory retrieval may be solved by graded and intermittent synchronization as well, as we shall show below.

### 3. Chaotic binding processes and intermittency in coupled maps

The creation, maintenance and deconstruction of ordered spatio-temporal patterns has been the subject of intensive investigations in the framework of coupled map (CM) systems (Kaneko 1983, 1990, Gu *et al.* 1984, Waller and Kapral 1984). CM systems are sets of coupled maps exhibiting non-linear dynamic behaviours.

In recent years it has been shown that CM (specifically, coupled map lattices, CML) systems can be used for invariant feature detection (van Leeuwen *et al.* 2000), perceptual segmentation and switching (van Leeuwen *et al.* 1997) and modelling memory retention and forgetting dynamics (van Leeuwen and Raffone 2001). In modelling these processes, there are functional advantages in assuming that the network dynamic patterns are non-stationary. In particular, on the borderline between chaos and ordered behaviour, there are critical regimes characterized by a scaling behaviour similar to that of brain activity. In particular, in systems with many degrees of freedom, such regimes occur spontaneously through a mechanism known as self-organized criticality (Bak *et al.* 1988).

Specifically, regimes that are characterized by *spatio-temporal intermittency* (van Leeuwen *et al.* 1997) present functionally-relevant properties. In the intermittent regime, the system exhibits an ordered (activation) pattern for a certain time interval, escapes from it and, after an interval of irregular behaviour, returns to the ordered pattern, or moves to a different one. The dynamical principle of system intermittency



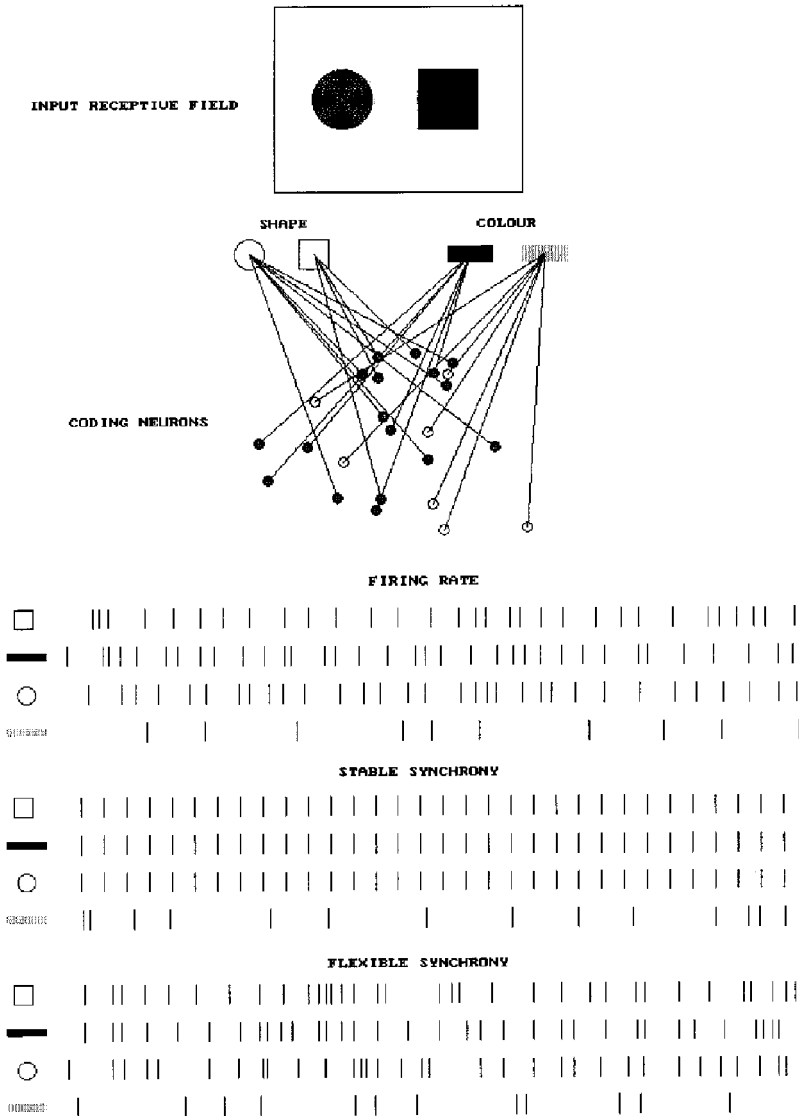


Figure 3. Schematic representation of ‘neural superposition catastrophes in terms of *both* firing rate and coding in terms of a stable synchrony of periodic signals. In this idealized scheme, visual neurons with large receptive fields, coding for specific object shapes or colours, are simultaneously presented with two objects characterized by the same colour. Furthermore, we plausibly assume that lower-level visual areas with smaller receptive fields may not be read-out in high-level perceptual representation, i.e. in visual awareness. In these conditions, both firing rates (see also von der Malsburg 1981) and stable synchrony (phase-locking) cannot discriminate the features belonging to the two separate objects (the active neurons are shown as closed circles). The neurons coding for the colour of the two objects (shared feature) induce the spurious synchronization of the unrelated features of the two objects, since synchronization of phase-locked neural oscillators is *transitive* and may be propagated even through relatively weak connections (Hansel and Sompolinsky 1992). In contrast, a code based on *flexible or graded synchrony* under non-stationary or chaotic spiking conditions enables the neurons coding for the shared colour to be ‘intermittently’ synchronized with the cells coding the features of different objects (see text for more explanations about our graded synchrony hypothesis).

may be related to the assumption that higher-order coherence states in cognitive activity are subject to intrinsic temporal constraints, called *coherence intervals* (Varela 1995), which could play a crucial role in perceptual and memory retrieval processes, as well as in the neural dynamics of awareness.

Like neural network units, the state of coupled maps is updated in discrete time. As a special case, we consider logistic maps, which are described by equation (1). Whereas equation (1) presents only one system variable, CMs are characterized by many system variables ( $x_t$ ), which evolve according to a coupled logistic map system. The logistic map chosen for our model represents a generic family of homeomorphic single-hump maps (Feigenbaum 1979), which all exhibit the same dynamic behaviour. These maps are characterized by a familiar pattern of stability behaviour, called ‘period-doubling route to chaos’. The system may evolve towards fixed points, periodic and chaotic attractors, depending on control parameter  $A$ . This parameter also characterizes oscillation amplitude; small changes in amplitude have generally strong non-linear effects on the behaviour of the system.

$$x^{(t+1)} = A x^{(t)} (1 - x^{(t)}). \quad (1)$$

If  $A < 1$  then  $Ax(1-x) < x$  and the system will approach zero; if  $A > 1$  and  $A < 3$ ,  $x$  will progress to a static point, which only depends on  $A$ . The function will approach a stationary attractor for all initial conditions  $x_0$ . With  $A = 3$ , the stationary point is no longer an attractor. The function will now approach a stable state in which it oscillates between two points with a period equal to 2. This change in the stability of the logistic map as a function of  $A$  is called ‘period-doubling bifurcation’. Still higher values of  $A$  will result in further period-doubling bifurcations, until chaos arises. This phenomenon, known as the period-doubling route to chaos, can be observed in the well-known *bifurcation diagram* (figure 4).

The Lyapunov exponent of the logistic map is given in equation (2). A positive Lyapunov exponent indicates divergence of two signals that start out with infinitesimally small differences. A positive Lyapunov exponent is therefore a necessary condition to call the behaviour of a system ‘chaotic’.

$$\lambda x \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^k \ln \left| A(2x^{(i)} - 1) \right|. \quad (2)$$

In figure 5, the Lyapunov exponent of the logistic map is plotted against the parameter  $A$ . According to this criterion, for values of  $A$  higher than 3.7, the logistic map yields chaotic activity most of the time, except in small bands of periodic activity. These are shown as negative peaks in the Lyapunov function. These negative peaks are known as ‘Arnol’d tongues’. In one of the higher Arnol’d tongues, as is shown in figure 5, a cycle of period 3 appears at  $A = 3.8282$ . Just below this value of  $A$ , the evolution is completely different. The transition is called ‘tangent bifurcation’. *Intermittency* is observed for values of  $A$  just below the tangent bifurcation. Intermittency is characterized by an alternation between ordered activity and episodes of irregular behaviour.

Equations (3)–(6) introduce a simple example of two coupled logistic maps. Their values at iteration  $t$  are  $x(t)$  and  $y(t)$ . In the formulas, *netx* and *nety* are intermediate

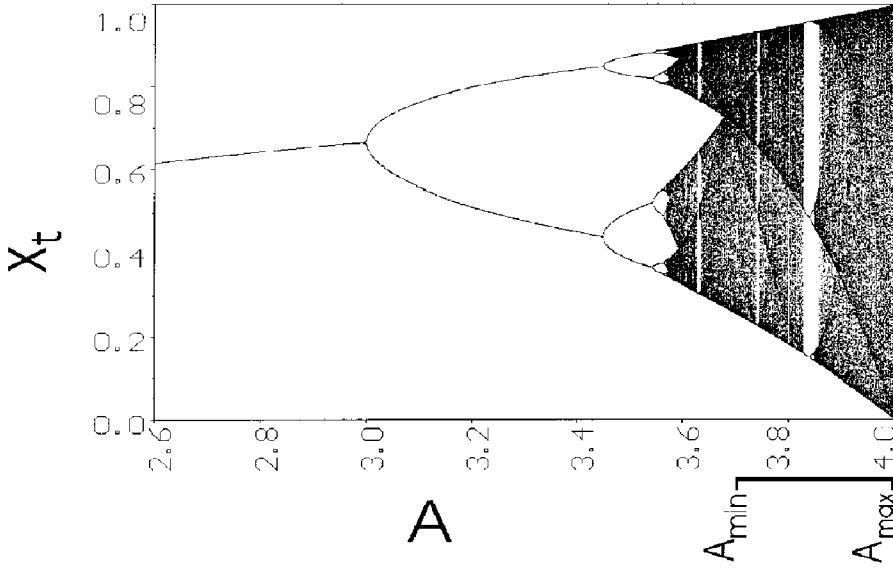


Figure 4. Bifurcation diagram showing the limit behaviour of the logistic map (equation (1)) as a function of chaotic oscillator parameter  $A$ . Chaotic behaviour is observed within the range  $A_{min} - A_{max}$ .

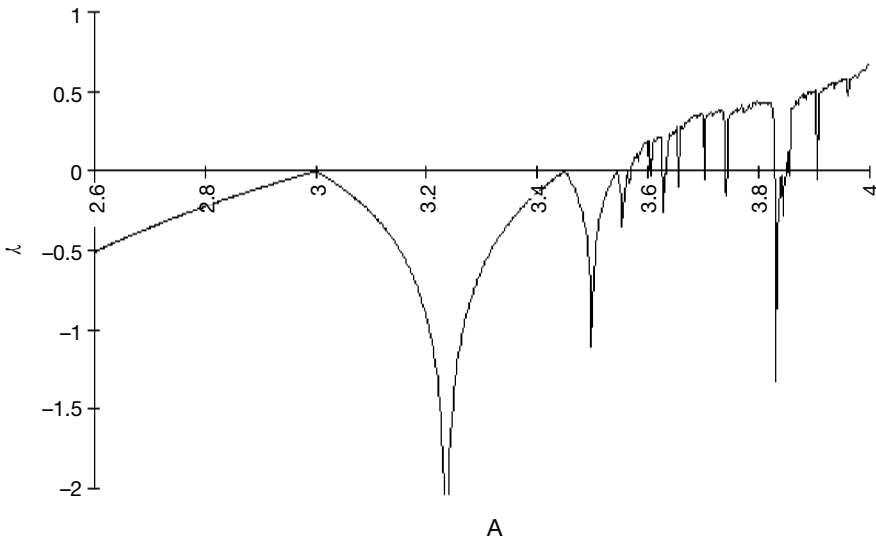


Figure 5. Lyapunov exponent for the logistic map as a function of parameter  $A$ . A positive exponent indicates chaos.

variables representing the net-input to these two units. The parameter  $A$  specifies the variation range of the function. The parameter  $C$  represents the coupling strength. This non-linear coupling scheme was pioneered by Kaneko (1990) who, through numerical studies, explored the stability characteristics of systems with a range of fixed and uniform values of  $A$  and  $C$  for a variety of network structures and sizes. The

parameter  $C$  plays a crucial role in the co-evolution of the states  $x$  and  $y$ . If  $C = 0.5$ , trivially,  $x$  is equal to  $y$ . It can be demonstrated, however, that for values well below  $C = 0.5$ , the system will evolve towards the state  $x = y$ , irrespective of the value of  $A$  (van Leeuwen *et al.* 1997). In other words, even if  $x$  and  $y$  are oscillating chaotically, they will reach a stable synchronization by *exchanging* a certain proportion of their value through their coupling. For certain values of  $C$  that are just below the values leading to a stable synchronization, spatio-temporal intermittency occurs. In this regime, the units of the system will be synchronized during certain time intervals and out of synchrony during certain other intervals. Numerical simulations by van Leeuwen *et al.* (in press) show that spatio-temporal intermittency is a robust phenomenon. Intermittency does not depend on specific ranges of initial values, it is obtained in sufficiently large and compact regions of the  $A$  and  $C$  parameter space, generalizing to larger systems of coupled units. Adaptive CM systems with a sufficient numbers of units will be able to reach these states autonomously, which may explain why self-organized criticality occurs in these systems.

$$netx = C y^{(t)} + (1 - C)x^{(t)} \quad (3)$$

$$nety = C x^{(t)} + (1 - C)y^{(t)} \quad (4)$$

$$x^{(t+1)} = A netx (1 - netx) \quad (5)$$

$$y^{(t+1)} = A nety (1 - nety). \quad (6)$$

#### 4. A model of memory retrieval based on activation and dynamic coherence coding

In order to illustrate the functional advantages of systems based on the combination of activation and graded coherence for modelling memory retrieval, we developed a model closely inspired by the paradigmatic McClelland's (1981, see also McClelland *et al.* 1986a) Jets and the Sharks neural network. The original network, shown in figure 1, is reproduced in most cognitive psychology textbooks to demonstrate the basic connectionist principles in memory retrieval. In this model, the information about persons belonging to different gangs is stored in the connectivity structure, the person nodes being in a central 'cloud', and the attribute nodes in multiple surrounding clouds.

Bi-directional excitatory connections exist in this model between a given instance node and one node of each attribute-set, and inhibitory connections exist between (all) the nodes of the same set. Retrieval takes place as a result of activation spreading from one or more cue nodes. Inhibitory connections mediate activity competition between nodes coding for mutually exclusive instances (gang members in the specific example) or features. This activity competition allows only a restricted number of instance nodes and feature nodes to be activated. Retrieval in this network occurs in terms of activation levels of the nodes after several cycles of activation spreading and mutual inhibition.

This simple, classical network model shares connectionism's conspicuous virtues as well as its vices, as mentioned in the Introduction. On the positive side of the balance, there are graceful degradation, default assignment and spontaneous generalization; on the negative side, superposition catastrophe effects and limited dynamic flexibility. In the following sections, we shall demonstrate that the Jets and the Sharks model and, more generally, connectionist interactive activation models,

may be functionally enriched by flexible dynamics in terms of chaotic or dynamic coherence.

#### 4.1. Model overview and architecture

A neural network with close structural resemblance to the original Jets and Sharks model was studied in our simulations (see figure 6). For the sake of generality, in the following sections we refer to the nodes in the central cloud as ‘instance nodes’ and to the nodes in the surrounding clouds as ‘feature nodes’ belonging to a given ‘attribute or feature-domain set’. The network consists of one set of ‘central’ instance nodes (coding for different entities) and eight sets of feature nodes, with bi-directional excitatory connections between a given instance node and one randomly assigned node of each feature set, and the inhibitory connections between nodes of the same set.

In the simulations we manipulated the number of nodes in the feature sets connected to the instance nodes. The probability of a connection between a certain feature node and an instance node is inversely proportional to the number of other feature nodes in the same set. Similarity between instance nodes may be expressed by the average number of shared features (number of connections to the same feature node). The average similarity between instance nodes is higher with a reduced set of feature nodes (alternatives for each feature dimension). Similarity, therefore, may be specified in the context of the present study as the inverse of the number of feature nodes. Alternatively, similarity may be set uniformly in terms of a (less realistic) hard-wired connectivity schema.

Similarity plays a crucial role in memory encoding and in categorization. For instance, phonological similarity (acoustical confusability) leads to a strong interference in encoding information in short-term memory, as well as in its subsequent retrieval (Conrad 1964). By contrast, semantic similarity gives rise to interference effects in (semantic) long-term memory storage and retrieval (Grossman and Eagle 1970). Very recently, Hommel *et al.* (2001) proposed the Theory of Event Coding (TEC), in which a general representational architecture with feature-overlap between active representations of stimuli and action plans may account for several effects observed with different experimental paradigms.

Rather than modelling a specific set of experimental data on memory retrieval, which would need domain-specific network architectures, in this study we are aiming at generality and simplicity. A pattern (representing an arbitrary entity or event) is coded by a set of nodes (units). One of these nodes is an instance node labelling the pattern; the other nodes coding for the pattern are feature nodes, with only one node being randomly sampled from a given feature domain. We assume that a mutual exclusion and competition between alternative feature nodes in the same domain (or ‘cloud’) takes place. For instance, the shape of a simple visual object cannot be circle-like and square-like at the same time.

The performance of the network is assessed in terms of *ambiguity resolution*, i.e. the degree to which during a stimulus-cued activation spreading process, the activated feature-coding units are preferentially synchronized with the units coding for one of the patterns. For ambiguity to be resolved, dynamic coherence between the eight feature nodes and an instance node belonging to the same pattern must be highest amongst the activated and reciprocally-connected nodes. This coherence will be selective and non-transitive, due to its intermittent (chaotic) character. These

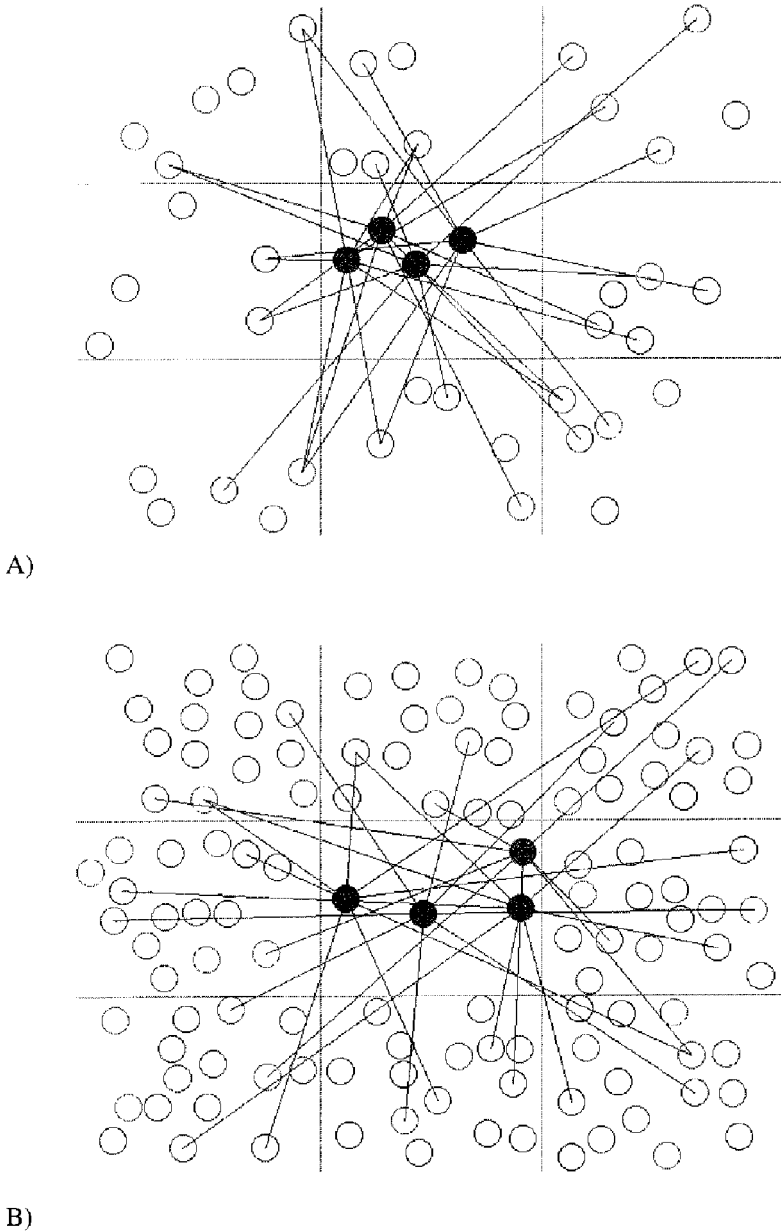


Figure 6. Scheme of the general architecture of our memory retrieval network. A set of nodes labelling generic instances (e.g. gang members, stimuli or multi-feature patterns), i.e. instance nodes, shown in the central box, are (reciprocally) connected to one of the feature units for each of the eight attribute boxes. Thus, a specific pattern is coded by a dedicated instance node and a set of eight feature nodes, which may be involved in multiple distributed representations (A) In a high-similarity condition with five units in each feature set, the probability of 'shared' features is higher than in a low-similarity condition with 15 units in each feature box (B). Note the higher number of connection converging links in (A), which facilitates a rapid activation spreading in the network.

properties will prevent superposition catastrophe effects due to the parallel and distributed architecture.

#### 4.2. A computational scheme with interacting activation and chaotic phase variables

The computational logic of the model we are introducing is based on updating two state variables for each unit at a given iteration. Each unit corresponds to a (micro)-assembly or group of densely connected neurons (see also Tononi *et al.* 1992 for a similar assumption). The first variable  $x_i$  is a chaotic phase, which is assumed to code the average onset of chaotically dispersed bursts within a local neural assembly, i.e. it is a variable related to spike emission timing over a time-resolved scale, defined in reference to an internal clock cycle. The other variable  $a_i$  is an activation variable, and it is assumed to be related to the average firing rate of the neurons of a given assembly. Hence, we assume an implicit separation of two dynamically related time-scales: a fast time-scale of evolution of phase interactions and a slower time-scale of evolution of activation level states. An iteration corresponds to a time period of several tenths of milliseconds, with usually only one spike-burst being emitted by each neuron within the micro-assembly. We assume that bursts are chaotically dispersed over the fast time-scale periods (iterations) due to non-linear interactions between membrane variables (see also Hansel and Sompolinsky 1992). Moreover, due to strong coupling within micro-assemblies, it may be safe to assume a low dispersion of spike onsets within assemblies, i.e. the average (unit) phase is representative of the spike timing of individual neurons within the assembly. The chaotic phase of unit  $i$ ,  $x_i$ , is computed in equation (7), as in equation (1).

$$x_i^{(t+1)} = A_i \text{Net}_i^{(t)} (1 - \text{Net}_i^{(t)}). \quad (7)$$

In the simulations, the modulation term  $A_i$  is assigned within the chaotic oscillation range [3.7, 4]. In these simulations, we assume that  $A_i$  models intrinsic brain *ground states*, which are dependent on thalamic and reticular modulatory inputs. Initial phase values are randomly assigned within a broad interval [0.25–0.75]. In previous models (van Leeuwen *et al.* 1997, van Leeuwen and Raffone 2001), based on chaotic phase dynamics only, stimulus input modulation was simulated through the  $A_i$  term.  $\text{Net}_i^{(t)}$  depends on the dynamic scaling of the  $x_i^{(t)}$  value at the last iteration and the weighted local field (*WLF*) of chaotic phases, the latter being computed in a previous algorithmic step of the same time iteration

$$\text{Net}_i^{(t)} = x_i^{(t)} (1 - C_i) + \text{WLF}_i^{(t)} C_i. \quad (8)$$

The weighted local field of a given unit  $i$  depends on both weights and activation of the ‘efferent’ units. An efferent unit  $j$  is more effective in contributing to the local field of a target unit  $i$  when its activation value  $a_j$  is high. Thus, in this computation different degrees of interaction between activations and phases may be considered. This interaction is neurobiologically plausible, since the number of synchronous spikes is proportional to the firing rates of neurons, and in turn the synchrony of spikes affects the firing rates of neurons. Several studies (Neven and Aertsen 1992, Roelfsema *et al.*

1996, Riehle *et al.* 1997) suggest that spike rates and temporal coding are complementary and in mutual interaction. For instance, attentional modulation of binding by synchrony may be computed with our present equations.

$$WLF_i^{(t)} = \frac{\sum_{j=1(i \neq j)}^N (x_j^{(t)} a_j^{(t)} w_{ij})}{\sum_{j=1(i \neq j)}^N (a_j^{(t)} w_{ij})}. \quad (9)$$

The activation value decays with a decay term  $\lambda$  and asymptotically approaches the variation range extremes 0 or 1. The input contributions from other connected units are specified by two terms, *Act* and *phAct*, the respective contribution of which is scaled by a proportion term  $p$ .

$$\text{if } Act^{(t)} p + phAct^{(t)} (1-p) \geq 0 \text{ then} \\ a_i^{(t+1)} = a_i^{(t)} \gamma + [Act^{(t)} p + phAct^{(t)} (1-p)] (1 - a_i^{(t)} \gamma) \quad (10a)$$

$$\text{else} \\ a_i^{(t+1)} = a_i^{(t)} \gamma + [Act^{(t)} p + phAct^{(t)} (1-p)] (a_i^{(t)} \gamma). \quad (10b)$$

The first term *Act* is computed in terms of a ‘classic’ connectionist scheme, i.e. it is given by the weighted (algebraic) sum of the activation values of the connected units

$$Act^{(t)} = \sum_{j=1}^N (a_j^{(t)} w_{ij}). \quad (11)$$

Since neurons act as coincidence (correlation) detectors (Abeles 1982 1991), i.e. synchronous synaptic inputs are more effective in triggering action potentials of target neurons, we assume that the time-averaged firing rate (the unit variable) is affected by time-resolved correlations of afferent signals (phase values of the  $j$  units). Thus, a phase-dependent activation term *phAct* is computed (equation (12)).

$$phAct^{(t)} = \frac{\sum_{j=1(i \neq j)}^N \exp\left(\frac{-|x_j^{(t)} - WLF_i^{(t)}|}{\tau}\right)}{N-1}. \quad (12)$$

In this equation  $\tau$  (ranging from 0 to 1) corresponds to a neuronal membrane time constant. When it is close to zero the input synchrony advantage is higher, and vice versa. Thus, the behaviour of the simulated neural units as coincidence detectors or as integrators may be specified by this parameter (see also Abeles 1991). Note that if the  $x_i$  phases of the incoming input are only slightly different, the sum of the modules of their differences from the *WLF* is low, and possibly equal to zero. But as such difference modules are the arguments (multiplied by  $-1$ ) of an exponential function, their sum is higher when they are close to zero, and decays depending on the  $\tau$  parameter. In the present simulations we do not use this phase-dependent contribution to



the units, thus setting the  $p$  term equal to one. However, we assume that coherence between the units is *read-out* in terms of neural computations reflecting coincidence detection (see equation (6)).

In previously developed perceptual-memory CM models (van Leeuwen *et al.* 1997, van Leeuwen and Raffone 2001), the coupling weights are plastic on different time-scales according to Hebbian dynamics. Weights are then computed through time-averaging (low-pass filtering) of phase coherence values. Depending on the time constant, different time-scales of weight plasticity may be realized in the present model.

### 4.3. Data analysis

To evaluate the selectivity of node-couplings in the network, the following measures of coherence and coactivation between two units  $i$  and  $j$  were used:

(i) A *phase coherence (PC)* measure averaged over a period  $T$

$$PC_{ij} = \frac{\sum_{t=t_0}^{t_0+T} \left( \exp \left( \frac{-|x_i - x_j|}{\tau} \right) \right)}{T} \quad (13)$$

in which the phase difference between two units is the argument of an exponential function and  $\tau$  (set equal to 0.1) plays the role of a neuronal time constant (see equation (12)). We assume that read-out assemblies (nodes), acting as coincidence detectors, are sensitive to the phase coherence of the network nodes.

(ii) A *coactivation (CA)* measure, averaged over a period  $T$

$$CA_{ij} = \frac{\sum_{t=t_0}^{t_0+T} \left( \sqrt{a_i a_j} \right)}{T} \quad (14)$$

which measures to what extent two units are both active during a given period.

(iii) An *effective phase coherence (EPC)* measure, given by the product of phase coherence and coactivation at a given iteration averaged over a period  $T$ , which may be related to the number of synchronous spikes emitted by two assemblies (nodes) over a given period

$$EPC_{ij} = \frac{\sum_{t=t_0}^{t_0+T} \left( \exp \left( \frac{-|x_i - x_j|}{\tau} \right) \sqrt{a_i a_j} \right)}{T}. \quad (15)$$

#### 4.4. Model behaviour

The retrieval dynamic behaviour of the network was tested in different simulation conditions, in terms of activation and phase coherence. In the simulations, we manipulated the following variables:

- The similarity of the 15 instance nodes, by alternatively considering attribute sets of five features (high similarity condition), of 10 features (intermediate similarity condition) and of 15 features (low similarity condition).
- The inhibitory strength coefficient  $W_{inh}$ , scaled to the excitatory strength  $W_{exc}$  coefficients and the number of feature nodes  $N_{feat}$ , according to the following rule

$$W_{inh} = \frac{\beta W_{exc}}{N_{feat} - 1}, \quad (16)$$

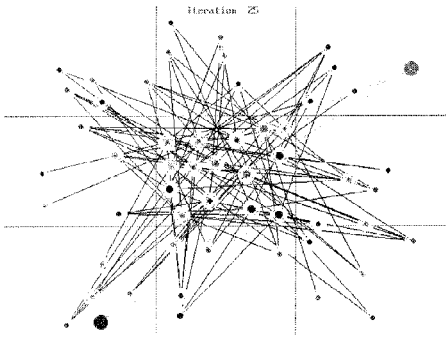
where  $\beta$  is usually set equal to 4.0 in low inhibition conditions and to 8.0 in high inhibition conditions.  $W_{exc}$  was set equal to 0.02.

- The chaos level, by setting the phase modulation term  $A_i$  equal to 3.7 in weak chaos conditions and equal to 4 in strong chaos conditions.

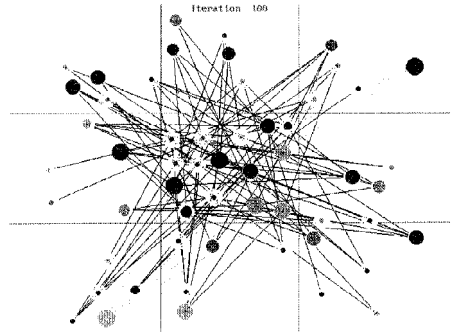
In the general simulation procedure, multiple trials for each condition were run by activating one or two cue nodes in different attribute sets and then observing the network evolution for 400 iterations in each trial. Different random connectivity schemes, initial values and cue nodes were used in each run. The initial activation values of the cue nodes were set equal to 0.75, and the activation values of the other units equal to zero. No autonomous activation decay was considered in the simulations ( $\lambda = 1.0$ ). The  $C$  parameter (equation (8)) was set equal to 0.375 for all the units. The main aim of the simulations was to demonstrate the representational effectiveness of phase coherence and chaotic correlations when the activation spreading in the network gives rise to superposition catastrophe effects and to show the relationships between activation and coherent states.

As shown in figure 7, a few iterations after the cue nodes' activation, the activation spreads to the related instance nodes, and then to the other feature nodes which are jointly activated by the instance nodes with a relatively high activation level. The cue node effectiveness, in terms of the strong activation of a limited number of instance and feature nodes, is maximal when cue nodes act in a conjunctive manner (McClelland 1981, McClelland *et al.* 1986a). The speed of the retrieval spreading, as well as the activation level in the entire network at the equilibrium state, depend upon the strength of the excitatory connections, the relative strength of the inhibitory connections, the number of retrieval cues and the instance node similarity, i.e. on their 'conjunctions' in terms of feature nodes. In the *high similarity* condition ( $N_{feat} = 5$ ), the instance nodes spread their activation by converging on to a limited number of feature nodes.

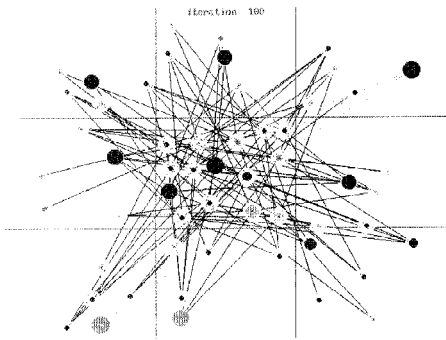
The simulations evidenced a trade-off between retrieval activation spreading and disambiguation, i.e. which features belong to which instance (pattern), which features are reciprocally related and which instances are mutually related. Disambiguation is resolved effectively in terms of dynamic coherence, without a rigid reference to the cued-retrieval associative path. Thus conceived, the model may be regarded as



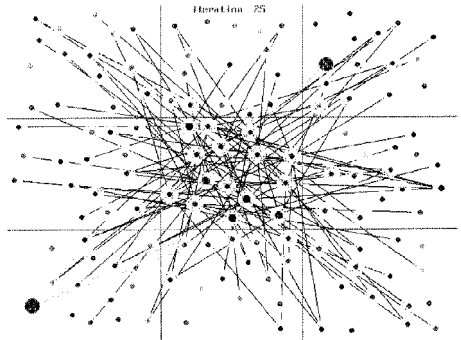
A)



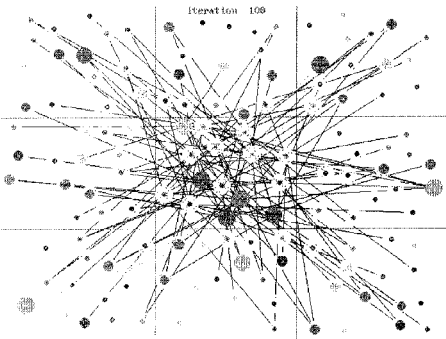
B)



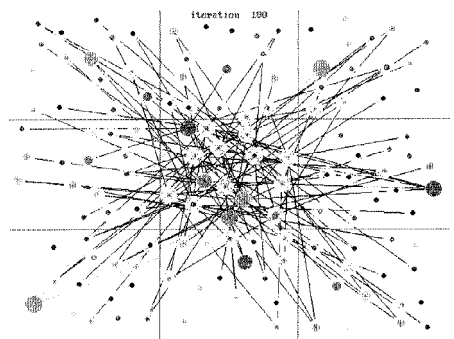
C)



D)

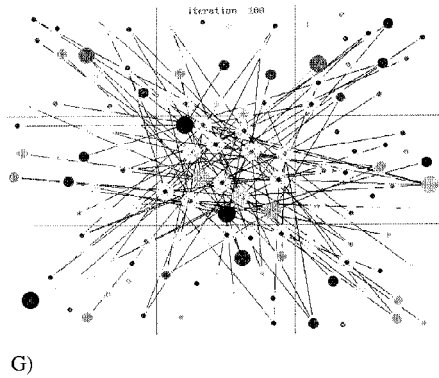


E)



F)

*continued . . .*



G)

Figure 7. Activation spreading and chaotic phase differentiation in our content-addressable memory model. Different clusters of chaotically-evolving phases enable the discrimination of active sets of units coding for more than one pattern. Instance nodes are shown in the central box; feature nodes are in the surrounding boxes. The activation level is coded by the radius of the nodes, i.e. more active nodes are depicted with a larger radius. The chaotic phase at the shown iteration, ranging between 0 and 1, is depicted by a grey level. (A) Simulation with high similarity among the instance nodes (implemented as the probability of a connection between the instance node and a given feature node). In the high similarity condition, five feature nodes are in each box. This condition is characterized by a low inhibition ( $\beta = 4$ ) between nodes in the same box, and a weakly chaotic phase evolution ( $A_i = 3.7$ ). The network state is shown after 25 iterations: activation spreading originates from two feature nodes that act as memory retrieval cues. These two nodes are shown as relatively large circles; their connections to instance nodes are grey. (B) The same simulation condition shown in (A), after 100 iterations. Note the increased activation spreading due to the recurrent connections. (C) In these conditions the connectivity diagram of the simulation is the same as in (B), but with a high inhibition ( $\beta = 8$ ) between nodes; note the lower number of active nodes due to a higher activity competition. The network evolution with the same patterns (recurrent connectivity condition) is shown with low similarity (15 nodes in each feature box), low inhibition and weak chaos, after 25 iterations (D) and after 100 iterations (E). (F) The same condition as in (E), but with high inhibition. (G) The same condition as in (E), but with strong chaos ( $A_i = 4.0$ ); note the higher phase variance in terms of differentiated grey-tunes of the units.

a first step towards the realization of autonomous, or self-addressable, memory functions.

At the same time as activation spreading, *phase coherence spreading* occurs in the network. This coherence spreading is related to the activation pattern, as the more active units are more effective in determining the phase value of coupled units (see the computational scheme). As shown in figures 7–13, even with a high level of network activation, phase coherence is selective for a given connectivity pattern. Chaotic phase coherence may effectively code for instance–feature relationships (which feature belongs to which pattern) (figures 8–11), the similarity between multiple instance nodes (figure 12) and the relationships between feature nodes (figure 13). The same instance unit may be simultaneously or subsequently engaged in transient synchronous states with multiple coupled feature units (figure 9).

According to the graded synchrony hypothesis, read-out of coherence states in the network will take place over a relatively long observation period, as in the classic firing rate read-out. However, short-lived coherence episodes, cumulated over a longer period, are temporally intermittent and spatially non-transitive, thus involving

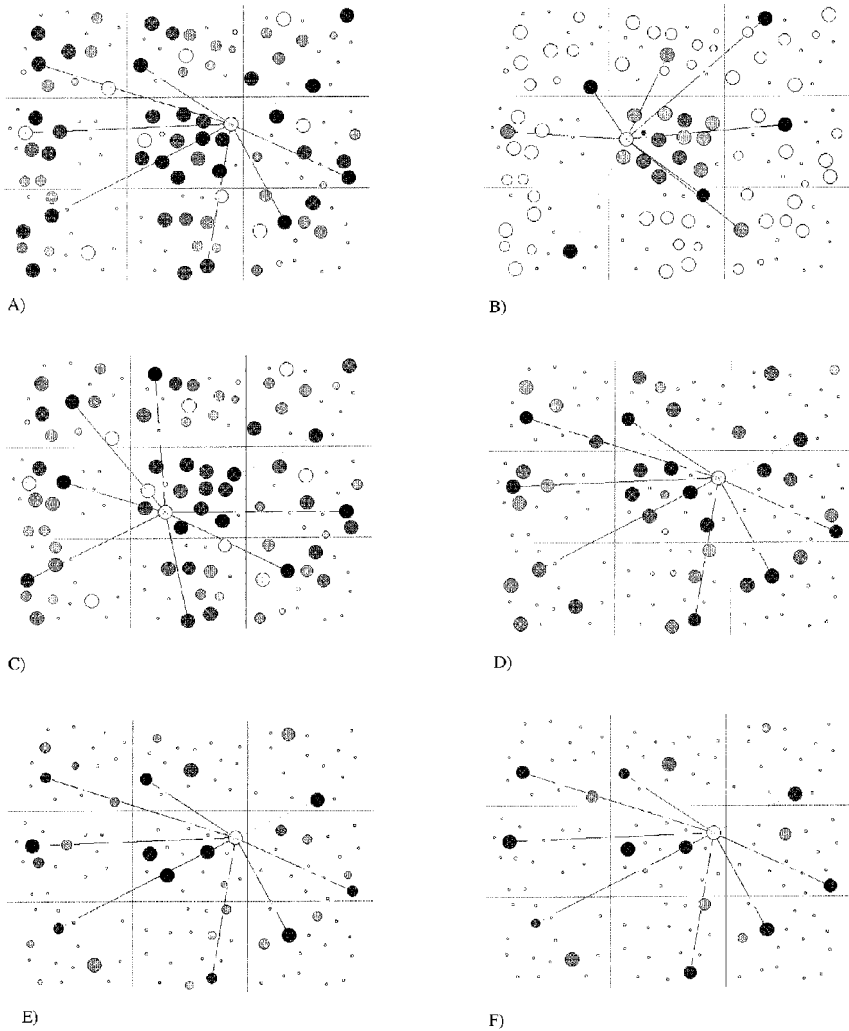
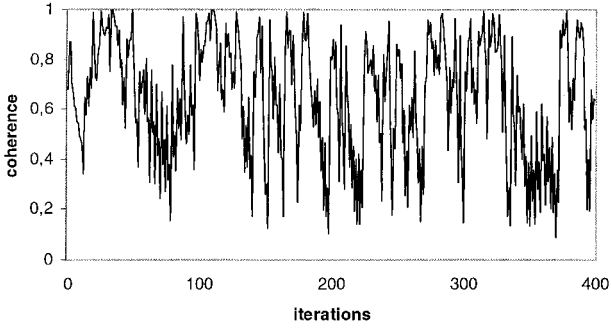
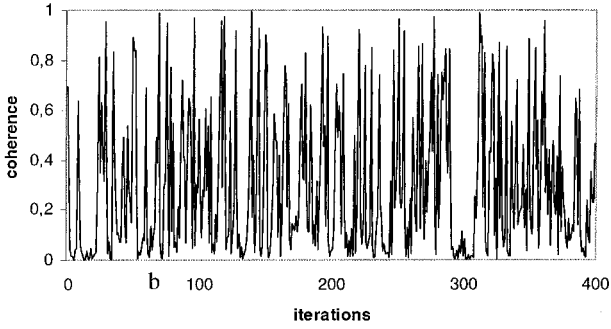


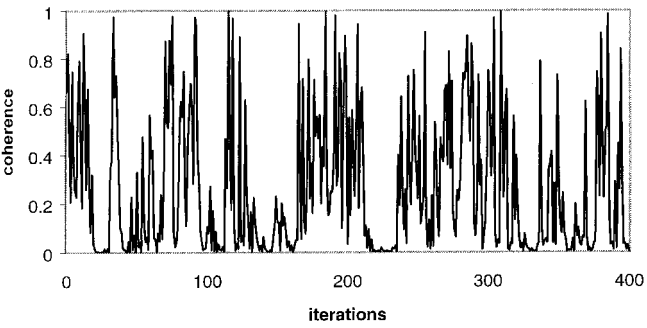
Figure 8. Coherence between a selected (reference) ‘instance node’ and the other instance and feature nodes in the network, averaged over a period from 200 to 400 iterations. The effective phase coherence (see the data analysis section), coded by the grey level, is normalized to a maximum value for each reference node (the grey-level ‘white’ codes for a coherence value equal to zero and ‘black’ codes the maximum coherence value observed for a given reference node in the central box). Connections of sampled reference nodes (white nodes in the central box from which the connections diverge) to the feature nodes of a given pattern are shown. (A) The instance node with index three in the instance node vector is the reference node, with weak chaos ( $A_i = 3.7$ ) and very low inhibition ( $\beta = 2$ ). Despite the high number of coactive nodes, coherence between connected nodes in the same pattern is higher than coherence between uncoupled nodes belonging to different patterns. (B) In the same simulation and observation period, the coherence of the ‘instance node 13’ is now shown; note the different coherence configuration related to the same network activation pattern of the ‘instance node 3’ shown in (A). In (C), the ‘instance node 9’ is shown, in the same conditions described in the two previous cases. The ‘instance node 3’ is shown with strong chaos ( $A_i = 4.0$ ) and low inhibition ( $\beta = 4$ ) in (D), with high inhibition ( $\beta = 8$ ) in (E) and very high inhibition ( $\beta = 12$ ) in (F). Thus, selective coherence may be an effective code to disambiguate pattern features with different levels of activation spreading.



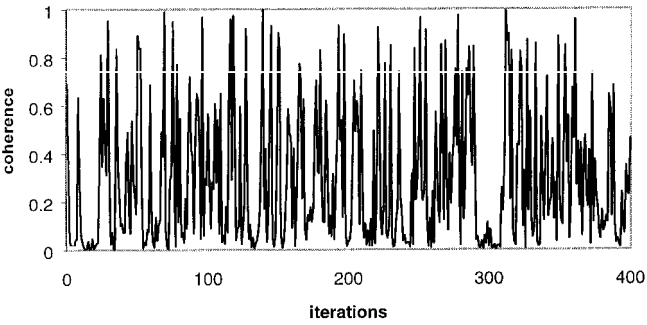
A)



B)



C)



D)

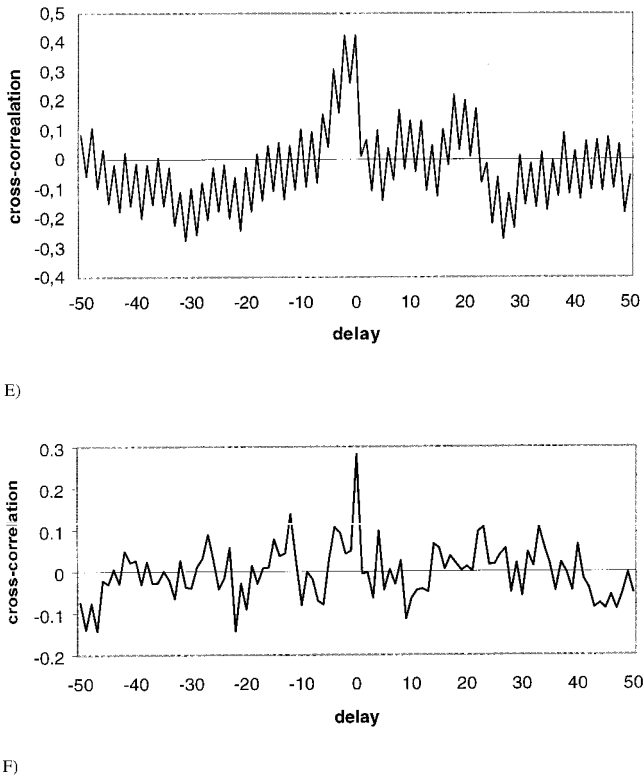


Figure 9. Temporal evolution of the effective phase coherence (see data analysis section) between a given instance node and two feature nodes in the pattern labelled by this instance node. Coherence between a reference instance node and the connected feature nodes of box (feature domain) 1 (A) and box 2 (B), with *weak chaos*. Coherence between a reference instance node and the connected feature node of box 1 (E) and box 2 (D), with *strong chaos*. (E) Cross-correlation of the time series shown in (A) and (B), and (F) of the time series shown in (C) and (D). Note the tendency to synchronization, and the fluctuations related to intermittent coupling and decoupling between different nodes.

different unit pairs at different times. This intermittency resolves an ambiguous read-out of active neural representations with overlapping representations.

Weak chaos (with  $A_i = 3.7$ ) induces a relatively stable phase coherence between coupled nodes, whereas with strong chaos (with  $A_i = 4.0$ ) short-lived coherence intervals are typically observed, with a given node being intermittently bound to other nodes. Thus, chaos may be useful in solving the connectionist superposition problem even with multiple overlapping patterns, as the involved network units may switch from one pattern to another in intermittent coherence dynamics.

Interestingly, due to chaotic phase spreading, time-resolved correlations may arise even between non-directly connected nodes. Such dynamic couplings may code relationships between feature nodes that are not directly connected, particularly with a high level of chaos (figure 13). In such a dynamic scheme, units coding for features that are repeatedly ‘conjoint’ in the instance-set, tend to be frequently synchronized. The resulting correlation over a given sampling or read-out period would provide a robust code for further processing dynamics.

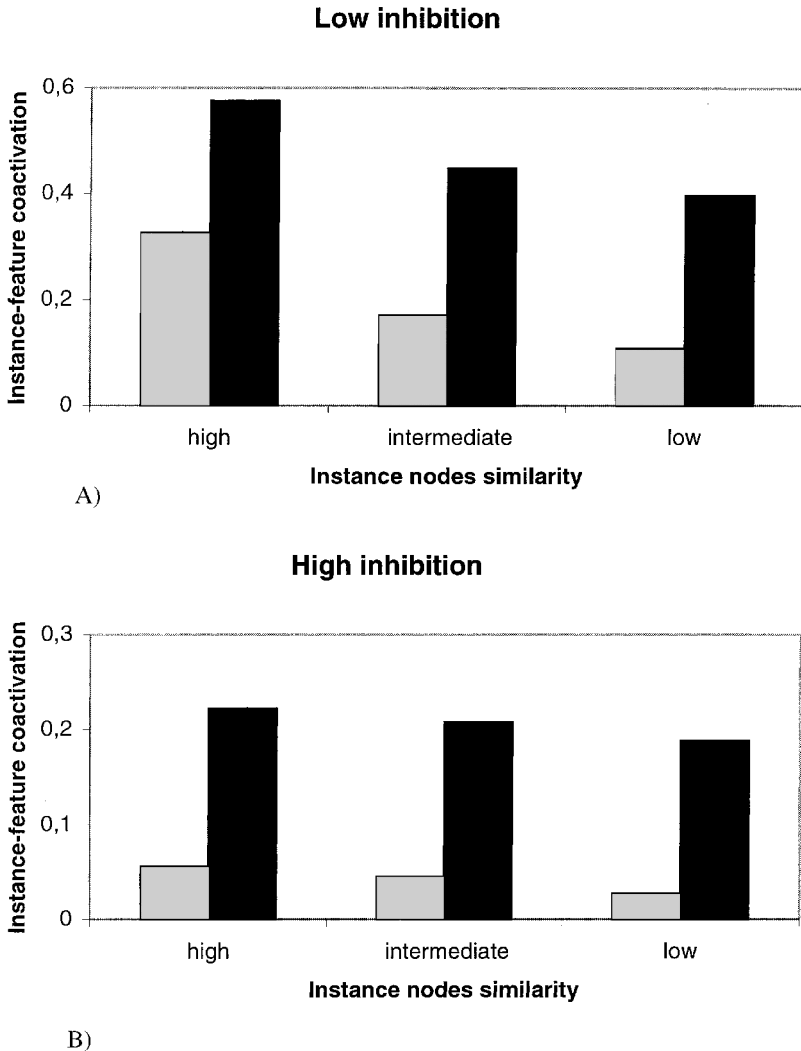
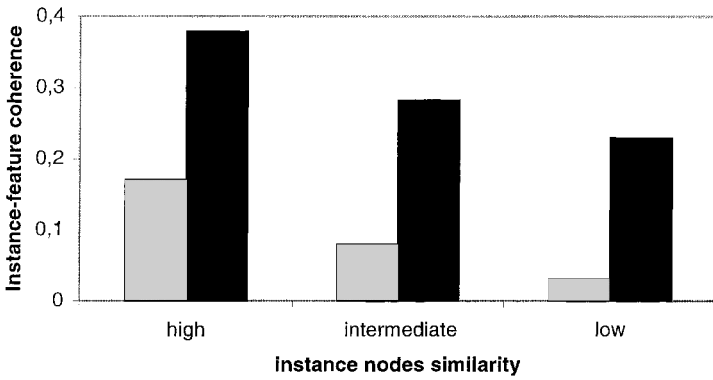


Figure 10. Coactivation (see text for more detailed explanations) of instance nodes and the coupled (in the same pattern) feature nodes (black bars), and of the same instance nodes and the feature nodes belonging to other patterns (grey bars), averaged over 10 trials with different connectivity (patterns). For a certain condition, a diagnostic measure of the discriminative network behaviour is given by the ratio of coupled over uncoupled average coactivation measures, i.e. by the ratio of the matched black and grey bar heights. (A) With low inhibition, the discriminative network behaviour in terms of coactivation is poor, particularly with high similarity of the instance (pattern-labelling) nodes. (B) With high inhibition, since only a subset of coupled network nodes tend to be activated, the network behaves in a more discriminative manner. In the high similarity condition, the high convergence of activation spreading involves a high number of unrelated nodes.

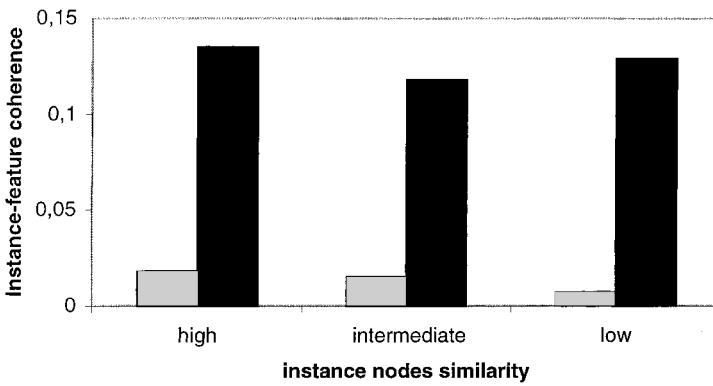


**Weak chaos - Low inhibition**



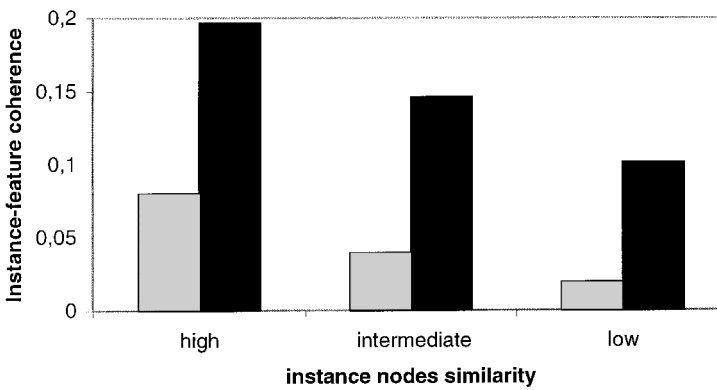
A)

**Weak chaos - High inhibition**



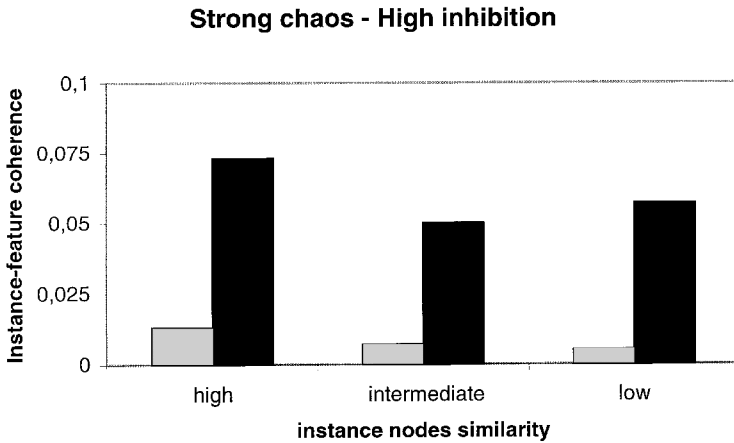
B)

**Strong chaos - Low inhibition**



C)

*continued . . .*



D)

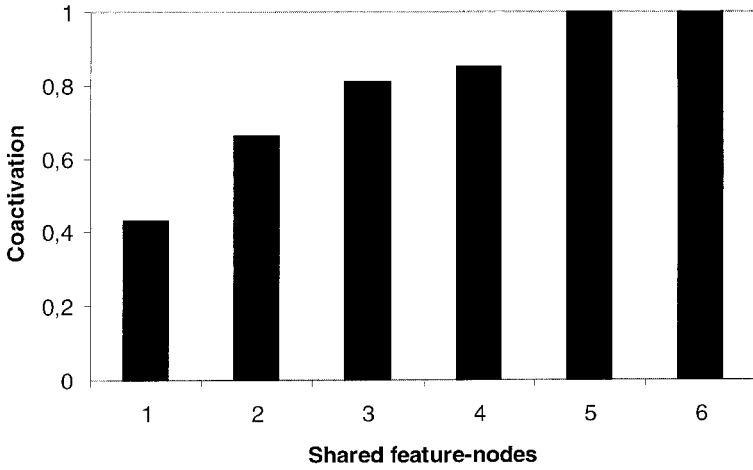
Figure 11. Effective phase coherence (see text for explanations) of coupled (black bars) and uncoupled (grey bars) instance–feature node pairs, averaged over 10 trials with different connectivity patterns. All network instance nodes and feature nodes were considered in averaging. (A) With weak chaos ( $A_i = 3.7$ ) and low inhibition ( $\beta = 4$ ); (B) with weak chaos and high inhibition ( $\beta = 8$ ); (C) with strong chaos ( $A_i = 4.0$ ) and low inhibition; (D) with strong chaos and high inhibition. When one compares these plots with the coactivation plots in figure 10, the network’s discriminative behaviour (nodes in the same pattern versus nodes in different patterns) is higher in terms of chaotic phase coherence. Note the relatively high ratio between the heights of black and grey bars, even in the high similarity condition.

## 5. Discussion

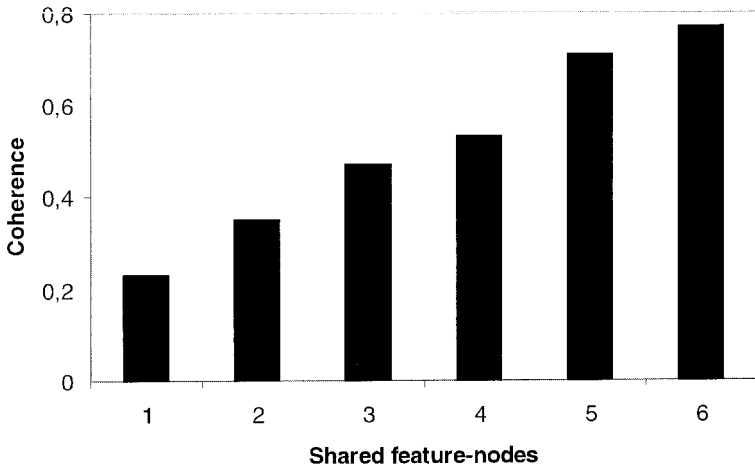
Our simulations have shown that unambiguous neural representational states are enabled by graded and intermittent coherence unit activity in a classic connectionist network, even during the retrieval of multiple patterns with overlapping features. Representational interference problems in pattern retrieval (leading to the superposition catastrophe) were minimized by temporal coding mediated by chaotic coherence states.

In the present model, pattern similarity and feature relationships are represented in terms of *graded synchrony*. Unlike other models of feature binding in terms of phase-locking of periodic oscillators (e.g. Sporns *et al.* 1989, von der Malsburg and Buhmann 1992), our model reflects a neurocomputational logic based on activation-weighted chaotic correlations, and thus enables flexible coding of proximity in feature spaces and *population coding* of complex feature constellations. By contrast, neural binding dynamics in terms of phase-locking only allow for the formation of internally synchronized and mutually desynchronized representational assemblies.

An interesting visual feature binding model, based on both activation and phase variables, was developed by Tononi *et al.* (1992). In their computational scheme, the phase variable is stochastically computed according to an activation-based weighting procedure. However, Tononi *et al.* implemented an arbitrary quantization of the phase variable (in terms of phase bins) and the activation variable was not characterized by an autonomous evolution, as in the present model. Moreover, intrinsic non-stationarity of the phase evolution is optimal for allowing the network units to switch from one active representational pattern to another.

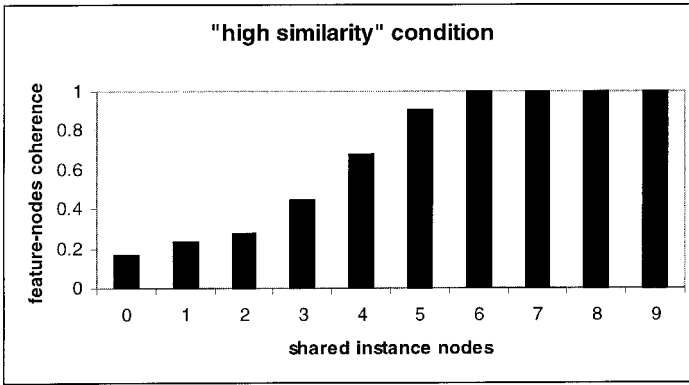


A)

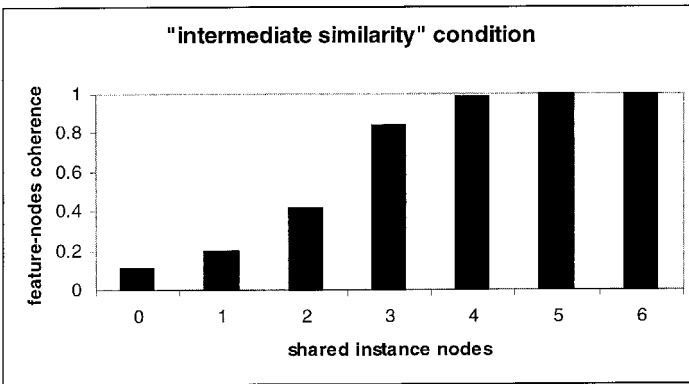


B)

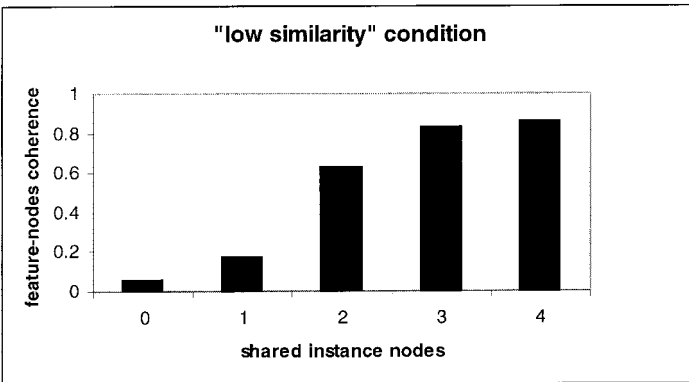
Figure 12. Coactivation and coherence between instance nodes labelling different patterns, with low inhibition. Average coactivation (A) and effective phase coherence (B) of instance nodes as a function of the number of shared feature nodes, in a low similarity condition with strong chaos. Note how coherence is more effective than the simple coactivation in discriminating the instance node relationships, i.e in flexibly coding the similarity of the instance nodes or patterns labelled by the instance nodes.



A)



B)



C)

Figure 13. Coherence between feature nodes, with low inhibition. The effective phase coherence between feature nodes is a function of the number of shared instance nodes, i.e. the number of patterns in which these features are associated, in (A) high, (B) intermediate and (C) low similarity conditions with strong chaos. As in the case of the couplings between the instance nodes of figure 12, coherence is effective in coding the relationships between the different features, i.e. the number of occurrences when they jointly pertain to the same instance.

The present model endorses a *dual coding* view, in which activation (firing rate) and synchronization-based coding play complementary roles in cortical information processing. As observed in the primary motor cortex of monkeys in a recent study of Riehle *et al.* (1997), synchronization and firing rate play different roles in neural coding. In their study, spike synchronization without firing rate modulation was related to internal cognitive events, while spike synchronization and rate modulation were shown to co-occur in processing external, behaviourally-relevant events. Thus, the brain may use different coding strategies, depending on functional requirements (Riehle *et al.* 1997). Our recent simulations with networks of coupled chaotic model neurons have shown that synchronization and firing rates may be differentially affected by structural and dynamic coupling conditions (Raffone and van Leeuwen submitted).

Spike rate coding and time-resolved coding may not be mutually exclusive, but rather complementary in cortical information processing, thus interacting to various degrees (Neven and Aertsen 1992, Roelfsema *et al.* 1996, Raffone and Wolters 2001). A coherence based on time-averaged (in the order of 100 ms or more) spike rates may reflect the recruitment of new neurons into an already active assembly, thus giving rise to new or enlarged assemblies, even when the sub-assemblies are located in spatially distant cortical sites. Complementarily, a coherence based on correlated individual spikes (over a time-scale of several milliseconds) may plausibly organize the simultaneously active groups into internally coherent (intra-assembly integration) and mutually incoherent assemblies (inter-assembly segregation) (see also Neven and Aertsen 1992). Attentive top-down signals could bias the activation and integration of rate-coherent assemblies, while suppressing the firing rate of other neural assemblies (Usher and Niebur 1995). Consequently, the binding/unbinding operations in terms of time-resolved coherence are less likely to give rise to ambiguous or illusory responses, due to a lower number of active neurons or lower-order neural assemblies.

In our present retrieval network, each node may participate in several correlational schemes, thus coding for multiple disjoint patterns. This dynamic behaviour is neurobiologically plausible, since it has been shown that during a time interval of tenths or hundreds of milliseconds, a given neuron may be dynamically recruited in different correlational assemblies, thus participating in multiple disjoint computations (Vaadia *et al.* 1989, 1995, Diesmann *et al.* 1999). In associative cortices, robust combinatorial codes given by spatio-temporal spiking patterns may be implemented in terms of multiple overlapping 'synfire chains', as demonstrated by Diesmann *et al.* (1999). Particularly robust and flexible associative codes in the cerebral cortex may be provided by combining synfire neurocomputational architectures with chaotically spiking (or bursting) neurons. More generally, it has been suggested that (non-stationary) spatio-temporal spike patterns appear to be more likely candidates for perceptual and working memory cortical coding than 'static' attractors or phase-locked neural oscillators (Fujii *et al.* 1996).

Using coupled logistic maps to generate chaotic states, as in our model, may not be considered to be neurophysiologically realistic. Nevertheless, the model may capture some relevant aspects of neuro-cognitive representation and processing. In simulations with Hindmarsh–Rose chaotic spiking neurons (Hindmarsh and Rose 1984), we have recently observed that feature overlap between stimulus patterns may be flexibly coded by neurons exhibiting chaotic behaviour, but not by phase-locking of periodically oscillating neurons (Raffone and van Leeuwen submitted). Thus, our

coupled map coherent behaviour may approximate the dynamic coupling of chaotic spiking neurons.

Recently, Tsuda (2001) has discussed the relevance of a neural coding scheme based on *chaotic itinerancy* in high-dimensional dynamical systems, as well as the related neurophysiological evidence. According to Tsuda, chaotic itinerancy arises when an intermediate state between order and disorder appears, and the dynamics of such a state may be regarded as being those of an itinerant process, indicating a correlated transition among states. A peculiar characteristic of this state is that the 'trace' of an original attractor remains, in spite of the generation of unstable directions in the neighbourhood of the attractor. Such an itinerant process often becomes chaotic.

Given that a similar dynamics is observed in our system, we endorse Tsuda's notion of chaotic itinerancy. Our interpretation of its neuro-cognitive relevance, however, differs. Given the strong evidence of functional specialization in the cerebral cortex, we emphasized the role of time-resolved coherence and its interplay with firing rates in neural binding (Raffone and van Leeuwen 2001). Tsuda (2001) favours a holistic dynamic view over selective binding, mediated by correlational neural assemblies. In our view, the functional differentiation evident in the organization of the cerebral cortex cannot be neglected.

The more general implications of our present approach for connectionist modelling may be summarized by the following points: (1) dynamic coupling among representational units may take place on several time-scales; (2) distributed neural representations are intrinsically non-stationary, and the same unit may participate in several representational and processing schemes during a 'psychological time-scale' period; (3) autonomous representational and processing behaviour may be modelled in the absence of inert attractors, but in terms of non-stationary trajectories in state evolutions; (4) in this functional logic, representation (conceived as states in the attractor paradigm) and processing (trajectories of state evolution) are not separable, as both are aspects of a system's dynamic behaviour; (5) thus, at the microstructural level of cognition, the so-called *state* and *process* descriptions (Simon 1962) should not be separated at all; and (6) the classic neural network activation and learning rules may coexist with CM dynamic coupling rules, and their relative effect on system dynamics is a matter of degree, depending on the computational (functional) requirements.

The possible relevance of chaotic coherent states for memory storage and retention has been suggested by van Leeuwen and Raffone (2001). The current model extends the synchronization-based binding approach to memory retrieval. These studies predict that the coherence of neural signals in distributed cortical areas increases during retrieval. For instance, correlated oscillatory signals in areas of the visual cortex may be generated during reactivation of visual memory traces. We expect that such coherence spreads from higher- to lower-level visual areas by backprojections, as has been shown for activation levels (Roland and Friberg 1985, Kosslyn 1994).

Another hypothesis that follows from the model is the non-stationarity of coherence. The lifetime of inter-neuron coherence is called 'coherence interval'. The distribution of coherence intervals may be  $1/f$ . Its modal duration is expected to increase from lower (e.g. V1) to higher (e.g. IT) visual cortical areas.

This model extends earlier studies (van Leeuwen *et al.* 1997, 2000). In these studies, CMs were used to model chaotic activity in the visual system. The units of the system represented activity of local assemblies, but their status remained unclear. By introducing two variables for representing assemblies, activation and phase, a step has

been made towards a more realistic application of CM in computational models of the visual system, even though more work in this direction needs to be done. These developments are expected to result from new computational applications. In our simulations we made the assumption that the phase variables depend on activation levels of other coupled units. This assumption was made in order to minimize interference from weakly active units coding for irrelevant features. Future simulations will systematically investigate the effect of phase coherence on unit activation levels. Since computer simulations with coupled maps are computationally convenient, the current phase activation algorithm could be applied to the segmentation of complex (e.g. medical) images (Wang and Terman 1997) and other pattern recognition problems. Further applications of more flexible interactive activation models of cognitive processes could be developed. For instance, in word recognition (McClelland and Rumelhart 1981) the simultaneous presentation of multiple words with shared features or letters (as in reading) may give rise to representational interference problems which could be resolved in terms of graded and intermittent coherence.

### Acknowledgements

We would like to thank three anonymous reviewers for their valuable comments, which significantly improved the manuscript. We would also like to thank Angela Tagini for her kind help in preparing the manuscript.

### References

- Abeles, M., 1982, *Local Cortical Circuits* (Berlin: Springer).
- Abeles, M., 1991, *Corticonics: Neural Circuits of the Cerebral Cortex* (Cambridge: Cambridge University Press).
- Amit, D. J., 1989, *Modeling Brain Function* (Cambridge: Cambridge University Press).
- Amit, D. J., 1995, The Hebbian paradigm reintegrated: local reverberations as internal representations. *Behavioral and Brain Sciences*, **18**: 617–657.
- Bak, P., Tang, E., and Wiesenfeld, K., 1988, Self-organized Criticality. *Physical Review A*, **38**: 364–374.
- Braitenberg, V., and Schüz, A., 1991, *Anatomy of the Cortex* (Berlin: Springer).
- Buckner, R., 1996, Beyond HERA: contributions of specific prefrontal areas to long-term memory retrieval. *Psychonomic Bulletin Review*, **3**: 149–158.
- Buffart, H., 1987, Seven minus two and structural operations. In: E. E. Roskam and R. Suck (eds) *Progress in Mathematical Psychology*, Vol. 1 (Amsterdam: North Holland), pp. 117–149.
- Buffart, H., and Geissler, H.-G., 1984, Task-dependent representation of categories and memory-guided inference during classification. In: E. Degreef and J. van Bruggenhaut (eds) *Trends in Mathematical Psychology* (Amsterdam: North Holland).
- Castelo-Branco, M., Goebel, R., Neuenschwander, S., and Singer, W., 2000, Neural synchrony correlates with surface segregation. *Nature*, **405**: 685–689.
- Conrad, R., 1964, Acoustic confusions in immediate memory. *British Journal of Psychology*, **55**: 75–84.
- Czerwinski, M., Lightfoot, N., and Shiffrin, R. M., 1992, Automatization and training in visual search. Special issue: views and varieties of automaticity. *American Journal of Psychology*, **105**: 271–315.
- Diesmann, M., Gewaltig, M. O., and Aertsen, A., 1999, Stable propagation of synchronous spiking in cortical neural networks. *Nature*, **402**: 529–533.
- Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., and Reitboeck, H. J., 1988, Coherent oscillations: A mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biological Cybernetics*, **60**: 121–130.
- Engel, A. K., König, P., Kreiter, A. K., Schillen, T. B., and Singer, W., 1992, Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends in Neuroscience*, **15**: 218–226.
- Feigenbaum, M. J., 1979, The universal metric properties of nonlinear transformations. *Journal of Statistical Physics*, **22**: 186–223.
- Fujii, H., Ito, H., Aihara, K., Ichinose, N., and Tsukada, M., 1996, Dynamical cell assembly hypothesis – theoretical possibility of spatio-temporal coding in the cortex. *Neural Networks*, **9**: 1303–1350.
- Geissler, H. G., and Puffe, M., 1983, The inferential basis of classification: from perceptual to memory representation systems. Part 1: Experiment on discrete feature processing. In: H.-G. Geissler, H. F.

- J. M. Buffart, E. L. J. Leeuwenberg and V. Sarris (eds) *Modern Issues in Perception* (Amsterdam: North Holland).
- Goldstone, R. L., and Medin, D. L., 1994, Time course of comparison. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **20**: 29–50.
- Gray, C. M., 1999, The temporal correlation hypothesis of visual feature integration: still alive and well. *Neuron*, **24**: 31–47.
- Gray, C. M., König, P., Engel, A. K., and Singer, W., 1989, Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, **338**: 334–337.
- Grossberg, S., and Grunewald, A., 1997, Cortical synchronization and perceptual framing. *Journal of Cognitive Neuroscience*, **9**: 117–132.
- Grossman, L., and Eagle, M., 1970, Synonymity, antonymity, and association in false recognition responses. *Journal of Experimental Psychology*, **83**: 244–248.
- Gu, Y., Tung, M., Yuan, J. M., Feng, D. H., and Narducci, L. M., 1984, Crises and hysteresis in coupled logistic maps. *Physical Review Letters*, **52**: 701–704.
- Hansel, D., and Sompolinsky, H., 1992, Synchronization and computation in a chaotic neural network. *Physical Review Letters*, **68**: 718–721.
- Haxby, J. V., Ungerleider, L. G., Horwitz, B., Maisog, J. M., Rapoport S. I., and Grady, C. L., 1996, Face encoding and recognition in the human brain. *Proceedings of the National Academy of Sciences USA*, **93**: 922–927.
- Hindmarsh, J. L., and Rose, R. M., 1984, A model of neuronal bursting using three coupled first order differential equations. *Proceedings of the Royal Society of London B*, **221**: 87–102.
- Hommel, B., Müseler, J., Ascherslaben, G., and Prinz, W., 2001, The Theory of Event Coding (TEC): a framework for perception and action planning. *Behavioral and Brain Sciences*, **24** (4) (in press).
- Hopfield, J. J., 1982, Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA*, **79**: 2554–2558.
- Kaneko, K., 1983, Period-doubling of kink-antikink patterns, quasiperiodicity in antiferro-like structures and spatial intermittency in coupled logistic lattices. *Progress of Theoretical Physics*, **72**: 480–486.
- Kaneko, K., 1990, Clustering, coding, switching, hierarchical ordering and control in a network of chaotic elements. *Physica D*, **41**: 137–172.
- König, P., and Schillen, T. B., 1991, Stimulus-dependent formation assembly formation of oscillatory responses: I. Synchronization. *Neural Computation*, **3**: 155–166.
- Kosslyn, S. M., 1994, *Image and Brain: The Resolution of the Imagery Debate* (Cambridge MA: MIT Press).
- Kubovy, M., and Wagemans, J., 1995, Grouping by proximity and multistability in dot lattices. *Psychological Science*, **6**: 225–234.
- Leopold, D. A., and Logothetis, N. K., 1999, Multistable phenomena: changing views in perception. *Trends in Cognitive Sciences*, **3**: 254–264.
- Lisman, J. E., and Idiart, M. A. P., 1995, Storage of 7 [+]<sup>2</sup> short-term memories in oscillatory subcycles. *Science*, **267**: 1512–1515.
- Luck, S. J., and Vogel, E. K., 1997, The capacity of visual working memory for features and conjunctions. *Nature*, **390**: 279–281.
- Mani, D. R., and Shastri, L., 1993, Reflexive reasoning with multiple instantiation in a connectionist reasoning system with a typed hierarchy. *Connection Science*, **5**: 205–242.
- McClelland, J. L., 1981, Retrieving general and specific information from stored knowledge of specifics. In *Proceedings of the Third Annual Conference of the Cognitive Science Society*, Berkeley, CA, pp. 170–172.
- McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C., 1995, Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, **102**: 419–457.
- McClelland, J. L., and Rumelhart, D. E., 1981, An interactive activation model of context-effects in letter perception: Part I. An account of basic findings. *Psychological Review*, **88**: 483–524.
- McClelland, J. L., Rumelhart, D. E., and Hinton, G. E., 1986a, The appeal of parallel distributed processing. In J. L. McClelland, D. E. Rumelhart and the PDP Research Group (eds) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 2: Psychological and Biological Models (Cambridge MA: MIT Press/Bradford Books), pp. 3–44.
- McClelland, J. L., Rumelhart, D. E., and the PDP Research Group, 1986b, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 2: Psychological and Biological Models. (Cambridge MA: MIT Press/Bradford Books).
- Miller, G. A., 1956, The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, **63**: 81–97.
- Miltner, W. H. R., Braun, C., Arnold, M., Witte, H., and Taub, E., 1999, Coherence of gamma-band EEG activity as a basis for associative learning. *Nature*, **397**: 434–436.
- Moscovitch, M., 1994, Interference at retrieval from long-term memory: the influences of frontal and temporal lobes. *Neuropsychology*, **4**: 525–534.



- Murthy V. N., and Fetz, E. E., 1992, Coherent 25 to 35 Hz oscillations in the sensorimotor cortex of awake behaving monkeys. *Proceedings of the National Academy of Sciences USA*, **89**: 5670–5674.
- Neven, H., and Aertsen, A., 1992, Rate coherence and event coherence in the visual system: a neuronal model of object recognition. *Biological Cybernetics*, **67**: 309–322.
- Phillips, W. A., and Singer, W., 1997, In search for common foundations for cortical computation. *Behavioral and Brain Sciences*, **20**: 657–722.
- Raffone, A., and van Leeuwen, C., 2001, Chaos and neural coding: Is the binding problem a pseudo-problem? *Behavioral and Brain Sciences*, **24** (4) (in press).
- Raffone, A., and van Leeuwen, C., 2001, Graded synchrony in flexible neural coding of visual scenes (submitted).
- Raffone, A., and Wolters, G., 2001, A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience*, **13**: 766–785.
- Raffone, A., Wolters, G., and Murre, J. M. J., 2001, A neurophysiological account of working memory limited capacity: within-chunk integration and between-item segregation. *Behavioral and Brain Sciences*, **24**: 139–141.
- Restle, F., 1970, Theory of serial pattern learning: structural trees. *Psychological Review*, **77**: 481–495.
- Restle, F., 1976, Structural ambiguity in serial pattern learning. *Cognitive Psychology*, **8**: 357–381.
- Riehle, A., Grün, S., Diesmann, M., and Aertsen, A., 1997, Spike synchronization and rate modulation differentially involved in motor cortical function. *Science*, **278**: 1950–1953.
- Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., and Varela, F. J., 1999, Perception's shadow: long-distance synchronization of human brain activity. *Nature*, **397**: 430–433.
- Roelfsema, P. R., Engel, A. K., König, P., and Singer, W., 1996, The role of neuronal synchronization in response selection: a biologically plausible theory of structured representations in the visual cortex. *Journal of Cognitive Neuroscience*, **8**: 603–625.
- Roelfsema, P. R., Engel, A. K., König, P., and Singer, W., 1997, Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature*, **385**: 157–161.
- Ritz, R., and Sejnowski, T. J., 2000, Synchronous oscillatory activity in sensory systems: new vistas on mechanisms. *Current Opinion in Neurobiology*, **7**: 536–546.
- Roland, P. E., and Friberg, L., 1985, Localization of cortical areas activated by thinking. *Journal of Neurophysiology*, **53**: 1219–1243.
- Rumelhart, D. E., McClelland, J. L., and the PDP Research Group, 1986, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 2: Foundations (Cambridge MA: MIT Press/Bradford Books).
- Schacter, D. L., 1989, Memory. In M. Posner (ed.) *Foundations of Cognitive Science* (Cambridge MA: Bradford Books), pp. 683–726.
- Schacter, D. L., Norman, K. A., and Koutstaal, W., 1998, The cognitive neuroscience of constructive memory. *Annual Review of Psychology*, **49**: 289–318.
- Shastri, L., 1999, Types and quantifiers in SHRUTI – a connectionist model of rapid reasoning and relational processing. In S. Wermter and R. Sun (eds) *Hybrid Neural Symbolic Integration*, Lecture Notes in Computer Science, Lecture Notes in Artificial Intelligence (Berlin: Springer) (in press).
- Shastri, L., 1999, Advances in SHRUTI – a neurally motivated model of relational knowledge representation and rapid inference using temporal synchrony. *Applied Intelligence*, **11**: 79–108.
- Shastri, L., and Ajjanagadde, V., 1993, From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings. *Behavioral and Brain Sciences*, **16**: 417–494.
- Shastri, L., and Mani, D. R., 1997, Massively parallel knowledge representation and reasoning: taking a cue from the brain. In J. Geller, H. Kitano and C. Suttner (eds) *Parallel Processing for Artificial Intelligence 3* (Amsterdam: Elsevier Science).
- Simon, H. A., 1962, The architecture of complexity. *Proceedings of the American Philosophical Society*, **106**: 467–482.
- Singer, W., 1994, Putative functions of temporal correlation in neocortical processing. In C. Koch and J. L. Davis (eds) *Large-scale Neuronal Theories of the Brain* (Cambridge MA: MIT Press).
- Somers, D. C., and Kopell, N., 1993, Rapid synchronization through fast threshold modulation. *Biological Cybernetics*, **68**: 393–407.
- Sougné, J., 1998, Connectionism and the problem of multiple instantiation. *Trends in Cognitive Sciences*, **2**: 183–189.
- Sporns, O., Gally, J. A., Reeke, G. N., Jr, and Edelman, G. M., 1989, Reentrant signaling among simulated neuronal groups leads to coherency in their oscillatory activity. *Proceedings of the National Academy of Sciences USA*, **86**: 7265–7269.
- Squire, L. R., 1992, Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, **4**: 232–243.
- Tononi, G., Sporns, O., and Edelman, G. M., 1992, Reentry and the problem of integration of multiple cortical areas: simulation of dynamic integration in the visual system. *Cerebral Cortex*, **2**: 310–335.

- Tsuda, I., 2001, Towards an interpretation of dynamic neural activity in terms of chaotic dynamical systems. *Behavioral and Brain Sciences*, **24** (4) (in press).
- Tulving, E., 1985, How many memory systems are there? *American Psychologist*, **40**: 385–398.
- Usher, M., and Niebur, E., 1995, Modeling the temporal dynamics of IT neurons in visual search: a mechanism for top-down selective attention. *Journal of Cognitive Neuroscience*, **8**: 311–327.
- Vaadia, E., Bergman, H., and Abeles, M., 1989, Neuronal activities related to higher brain functions – theoretical and experimental implications. *IEEE Transactions Biomedical Engineering*, **36**: 25–35.
- Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, U., Slovin, H., and Aertsen, A. M. H. J., 1995, Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature*, **373**: 515–518.
- van Leeuwen, C., and Buffart, H., 1989, Facilitation of retrieval by perceptual structure. *Psychological Research*, **50**: 202–210.
- van Leeuwen, C., Buffart, H., and van der Vegt, J., 1988, Sequence influence on the organization of meaningless serial stimuli: economy after all. *Journal of Experimental Psychology: Human Perception and Performance*, **14**: 481–502.
- van Leeuwen, C., and Raffone, A., 2001, Coupled nonlinear maps as models of perceptual pattern and memory trace dynamics. *Cognitive Processing*, **2**: 67–111.
- van Leeuwen, C., Raffone, A., and Simionescu, I., 2001, Flexible synchronization in cognitive processing. *Proceedings of the International Conference of Neural Information Processing (ICONIP 2001)* (in press).
- van Leeuwen, C., Steyvers, M., and Nooter M., 1997, Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. *Journal of Mathematical Psychology*, **41**: 1–26.
- van Leeuwen, C., Verver, S., and Brinkers, M., 2000, Visual illusions, solid/outline-invariance, and non-stationary activity patterns. *Connection Science*, **12**: 279–297.
- van Orden, G. C., Pennington, B. F., and Stone, G. O., 1990, Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, **97**: 488–522.
- Varela, F., 1995, Resonant cell assemblies: a new approach to cognitive functions and neuronal synchrony. *Biological Research*, **28**: 81–95.
- von der Malsburg, C., 1981, *The Correlation Theory of Brain Function*. Internal Report 81–2, Max Planck Institute for Biophysical Chemistry, Göttingen.
- von der Malsburg, C., and Buhmann, J., 1992, Sensory segmentation with coupled neural oscillators. *Biological Cybernetics*, **54**, 29–40.
- Waller, I., and Kapral, R., 1984, Spatial and temporal structure in systems of coupled nonlinear oscillators. *Physical Review A*, **30**: 2049–2055.
- Wang, D. L., and Terman, D., 1997, Image segmentation based on oscillatory correlation. *Neural Computation*, **9**: 805–836.