5.1.93/NOB

SYSTEMATICITY IN THE VISION TO LANGUAGE CHAIN

Niels Ole Bernsen, CCI, Roskilde University and Risø National Laboratory

Summary: The paper addresses the challenge to distributed connectionism from adherents of the physical symbol systems paradigm for cognitive science. The challenge is presented semi-formally and is then addressed both theoretically and in two distributed connectionist simulations of increasing sophistication processing 'real' semantic material supposed to have come from a machine vision front-end. It is claimed that these simulations and their theoretical underpinnings successfully meet the systematicity challenge and that it is up to adherents of the classicalist position if the systematicity debate is to continue. The prospects for this to happen appear slim as the classicalist position is not geared to assume the particular burden of proof involved.

Keywords: Systematicity in distributed connectionism, spatial cognition.

1. Introduction

Connectionism seems likely to have come to stay as a second computational paradigm for cognitive science in addition to the paradigm of classical AI (for the latter, see, e.g., Pylyshyn 1984). Connectionism with distributed representations throughout has been proposed as a complete, self-sufficient and non-hybrid alternative to the classical paradigm in accounting for the representational states and processes of cognitive systems, biological or otherwise (Smolensky 1988, Smolensky et al. 1992). However, just as the general term 'connectionism' continues to lack a clear definition as an alternative computational paradigm for cognitive science, there still remains fundamental unsolved problems for *distributed* connectionism to serve as such a general paradigm. These problems concern how to account for the constituent structure of thought and will be addressed in what follows. It is claimed that these problems can be solved both theoretically and in working connectionist simulations. Distributed connectionism, therefore, has a strong claim to being considered a second general computational paradigm for cognitive science.

Fodor and Pylyshyn (F&P,1988) have argued that thoughts or representational mental states, just like natural language, have *combinatorial syntactic and semantic structure* and that the utilisation of such structural properties is crucial to inference and reasoning. It is because mental representations have combinatorial structure that it is possible for mental operations or processes to apply to them by reference to their form. Thus, mental processes have *structure sensitivity*. Structures of expressions can have causal roles because structural relations are encoded (or implemented) by physical properties of brain states in appropriate ways. Localist connectionist networks, F&P claim, do not have combinatorial syntactic and semantic structure and the processes operating over them do not have structure sensitivity. Such networks, therefore, are at most vehicles for the implementation of cognition rather than accounts of cognition at the proper theoretical level which according to F&P is the level of complex symbol structures. Distributed representation networks are no better off in this respect than are localist networks, they claim. Briefly, the argument runs as follows with respect to localist networks: such networks need one set of elements to represent, e.g., the thought that John loves Mary and a different set of elements to represent the thought that Mary loves John. No one set of elements is able to represent the combinatorial syntax of the thought that John loves Mary for the simple reason that such a set of elements does not have combinatorial syntactic and semantic structure. In particular, the unit which fires whenever the system entertains the thought that John loves Mary does not have syntactic and semantic constituent structure. It is a simple, atomic and therefore unstructured unit which fires whenever the system has the thought that John loves Mary. And of course, processes involving this unit cannot be sensitive to a structure that the unit does not have.

It is possible to provide localist networks with constituent structure and variable binding (e.g., Ajjanagadde and Shastri 1991). We have done work on localist networks at CCI using a different approach and found such networks significantly more efficient than distributed networks solving the same problems in spatial cognition (Bernsen and Kopp 1993). However, localist networks are implausible from a cognitivist point of view and will not be discussed here. As for distributed networks, the long discussion fuelled by Foder and Pylyshyn's original paper still has not produced clear and stable solutions (Smolensky 1987, Fodor and McLaughlin 1990, Smolensky et al. 1992). The present paper attempts to identify some of the reasons why this is so.

The plan for the paper is as follows: Section 2 offers a semi-formal presentation of the systematicity challenge to connectionism. Section 3 defines a first set of criteria which may enable a specific distributed connectionist system to meet the challenge. The system itself, *System 1,* is described in Section 4. In Section 5, a number of objections to System 1 as a satisfactory solution to the systematicity challenge are considered. One of these objections leads to the construction of a new distributed connectionist system, *System 2*, with a rather different architecture and a cognitive task slightly different from that of System 1 (Section 6). The results of a recent simulation done with System 2 are discussed in Section 7 leading to the conclusion that the systematicity challenge has now been fully met. However, a further twist to the argument remains and is discussed in Section 8. The final conclusion (Sect. 9) is that there is now strong, if not entirely conclusive, reason to leave the systematicity issue alone and carry on with the more substantial issues of cognitive science. Protagonists of the classical position will have to make significant progress to keep the discussion alive.

2. The Systematicity Argument

In a semi-formal expression, F&P's original argument goes like this:

(a) Information-processing which produces intelligent behaviour has a natural sub-class which is called *cognition* and which includes thinking, reasoning, understanding and generating natural language, etc.

(b) Cognition involves *thoughts* as a form of mental representation.

(c) Thought has a small number of basic properties which we may call *the set P.* The set P includes properties such as:

- systematicity;
- compositionality;
- inferential systematicity;
- and, possibly, other properties which have yet to be discovered.

Comment: The properties included in the set P replace, for the sake of the systematicity argument, the properties of 'combinatorial syntactic and semantic structure' and 'structure sensitivity' noted in Sect. 1 above. The properties of systematicity and compositionality will be explained under (d) below. The property of (semantic) compositionality will not be discussed separately since nothing in F&P’s argument hinges on the possible differences between systematicity and compositionality. It will be evident from the argument below that if systematicity is not a problem for distributed connectionism, neither is compositionality. The property of (some measure of) compositionality follows from systematicity both on F&P's account and on the present alternative account. The property of inferential systematicity will not be discussed separately as it does not form part of F&P’s central argument. I shall simply conjecture that if distributed networks can handle systematicity, they can also handle inferential systematicity *using the very same cognitive mechanisms* as they use in handling systematicity.

(d) The units of (localist) connectionist systems have no internal structure. Two units representing the thoughts that John loves Mary and that Mary loves John, respectively, have no common structure (they *have* no structure). It follows that such systems might be able to have, e.g., the thought that John loves Mary but unable to have the thought that Mary loves John. This possibility shows that such systems are not characterised by systematicity, that is, their ability to think some thoughts is not *intrinsically* connected to their ability to think certain others, even though both sets of thought involve exactly the same concepts or lexical entries and are composed by using exactly the same compositional principles. However, this does ot preclude that localist connectionist systems can be crafted to implement classical architectures of cognition. Compositionality, in this context, is simply the fact that the systematically related thoughts that a system is able to have are not only related syntactically but also semantically. I.e., which thoughts are systematically related is not arbitrary from a semantic point of view.

(e) Such systems, with or without the capacity for natural language, which lack systematicity, do not exist in nature. In nature, behaviour, thought, output of cognitive modules, and language are *necessarily* systematic.

Comment: Note that the systematicity argument does not rest on systems' possession of linguistic capacity. We are dealing with a much more fundamental property of thought and mental representation. Non-linguistic animals and infraverbal cognition also demonstrate systematicity of thought and F&P claim that the inadequacy of connectionist models as cognitive theories follows quite straightforwardly from this empirical fact. Basically, the thesis of the systematicity of thought is a claim about the systematicity of representations underlying a great deal of the observable behaviour in humans and animals. If some human or animal is able to think certain thoughts (or have certain mental representations), they are necessarily also able to think certain other thoughts which can be seen to be systematically (or intrinsically) related to the former.

(f) Distributed connectionist systems are not necessarily systematic, or at least have not yet been shown to be so, unless they merely implement classical physical symbol systems. On distributed connectionist principles, the systematicity of thought is a *mystery*, F&P claim. However, they do not claim to be able to prove that distributed connectionist systems are necessarily unable to exhibit systematicity without merely implementing classical cognitive architectures. Rather, the burden of proof is laid on distributed connectionism.

(g) The systematicity argument does not presuppose the physical symbol systems hypothesis (or the classical paradigm for cognitive science).

(h) It follows from (a)-(g) that connectionism cannot (in principle or at least at present) account for thought. Since thought is basic to cognition, connectionism does not constitute an alternative general paradigm for cognitive science. Now, since the physical symbol systems hypothesis *can* account for thought (or the set P), the hypothesis that classical constituents are tokened as part of the thought they syntactically constitute is the 'only game in town' as an account of thought and hence of cognition. Thought has combinatorial syntactic and semantic constituent structure in the sense of the physical symbol systems hypothesis. In particular, distributed connectionist systems are at best implementations of classical cognitive architectures.

I shall accept (a)-(e) and (g) for the sake of the argument, (c), (e) and (g) being accepted with no qualification other than a bracketing of the theoretical motivations behind F&P’s notion of thought. The crucial issue, therefore, is (f) to which we may now turn. If (f) is false, the conclusion (h) does not follow.

3. A First Reply to the Systematicity Argument

To show that (f) above is false, it must be demonstrated

- that distributed connectionist systems *are able to* produce systematic mental representations;
- the *cognitive* mechanism(s) which cause this to be *necessarily* the case must be made clear. These mechanisms, of course, cannot and should not be classical syntactic and semantic ones. Otherwise, doubts may remain as to whether what has been produced is merely a connectionist implementation of a classical cognitive architecture.

A crucial assumption in the argument below is that distributed systems do not have classical syntax (or *ditto* constituent structure). Given this assumption, if (f) is shown to be false in the manner just indicated, distributed connectionism can be considered a general paradigm for cognitive science. I shall come back to that assumption in Sect. 8 below. The claim we are considering is the claim that *only* classical syntax or the existence of an internal syntactic structure of mental representations can explain the systematicity of mental representation that is evident from much of human and animal behaviour.

Clearly, nothing hinges on representing the particular examples concerning John's love for Mary and Mary's love for John. Indeed, 'love' is a very difficult concept to represent, and not only in connectionist systems. We are thus free to choose a different example for experimental demonstration and subsequent interpretation. On the other hand, the issue does seem to be one of representing central cognition or linguistically expressible thought rather than possibly unconscious and linguistically inexpressible peripheral cognitive states.

So we need an example of a complex thought. It could be a two-place, asymmetrical relational thought just like the ones about John and Mary above. The thought should have systematicity in the sense that a system should not be able to have the thought that aRb (or $R(a,b)$) without, necessarily, being able to have the different but systematically related thought that bRa (or $R(b,a)$), a and b representing individuals. If a connectionist system with distributed representations is able to entertain such systematically related thoughts then this system demonstrates systematicity. It also demonstrates compositionality in the rather limited sense of Sect. 2 above. And, *ex hypothesis*, it does so without having syntactical representations in the classical sense of the term. Finally, we would like to be able to explain why systematicity necessarily obtains in the system.

An example meeting the above constraints is the relational thought partially caused by perception that something (e.g., a triangle, or John) is to the right of something else (e.g., a square, another triangle, or Mary). Let us apply some conditional reasoning to this example:

If a distributed connectionist system

- can learn from experience something which for the sake of the argument is sufficiently close to the concept of 'spatial object in general'; and
- can learn from experience something which for the sake of the argument is sufficiently close to the concept of 'right-of-ness in general' as correctly applying to two arbitrary spatial objects;

then

- it *does not matter* to the system whether it applies these concepts to R(a,b) which it has seen in the training set, or it applies them to R(b,a) which it has not seen in the training set;
- the system therefore *necessarily* realises systematicity through a distributed version of variable binding where:
- variable binding is achieved through processes of *abstraction* from experience, *generalisation* from experience, and *instantiation* to new, currently experienced instances which it has not seen before; and
- variable binding is dynamically achieved through patterns of weighted connections and without classical syntactic constituency and syntactic combinatorial structure.

Such a system will have other classical properties as well such as an infinite generative capacity (if it could only be tapped in some way), basic context-independence of its representations, and semantic compositionality in the sense of Sect. 2 above. In fact, the system is in many respects similar to a classical system which might be capable of solving the same cognitive task. The system processes 'real' semantic material which might have been produced by a machine vision front-end and it produces output which might be input for further processing by a natural language generation module. It is therefore consistent with the hypothesis of modular cognitive architecture. Furthermore, the system learns the concepts it has from experience which is undoubtedly the way biological systems come to have such concepts as the ones considered here.

Note that nothing has been said about syntax above. It is not claimed that the described system acquires, through experience and training, a syntactic representation of the form '*right-of(x,y)* '. It is not claimed that, on the basis of such a representation the system performs a formal syntactic operation of binding, through the operation of substitution which we call instantiation, the variables x and y to, say, John and Mary. Such representations and operations form part of *one particular* (i.e., syntactic) algorithmic way of describing what the system might be doing. As hypothesised above, our distributed connectionist system does not do things this way. It does learn the abstract concept that *something* is spatially related to *something else* through the relation 'right-of'. But instead of the variables x and y it has a pattern of weighted connections between its units of activation which perform *as if* they were variables like x and y, or, rather, which perform the same cognitive task as that performed by a syntactic system with variables x and y, but differently. And the distributed system does not formally bind the variables x and y (which it does not have) to the particular individual objects it perceives through the formal syntactic operation of substitution. Rather, the 'something' and 'something else' representations of the network (which are realised by its weighted connections and units of activation) become activated by input representing individual objects in space. This activation allows the network to determine whether or not those objects stand in the right spatial relationship for the relationship 'right-of' to obtain between them.

The system described does not, strictly speaking, know of formal logic and does not represent the world in terms of formal logic. But it does represent abstract concepts and knows how to apply them to individuals that it perceives in its world. It represents abstract concepts of two kinds. First, it represents the concept of a spatial object in general, more or less. Second, it represents the concept 'right-of' in general, more or less, since it is able to correctly describe objects in different positions as being or not being to the right of other objects. The powerful mechanisms of abstraction from experience and subsequent instantiation to experienced objects are what is responsible for these capabilities and thus for the system's mastery of systematicity (and compositionality). The crucial point is that the system does master (non-syntactic) combinatorial semantic constituent structure. In other words, if such a system can be built, it will realise systematicity at the cognitive level through algorithmic means that are basically different from those of classical syntactic systems.

The abstract representation which our hypothesised system has 'that some spatial object is to the right of some other spatial object' would clearly seem to count as a semantically *complex* representation. Being abstract, this representation is, at least in principle, able to *generate* infinitely many different instantiations. In virtue of its abstractness, it is also to a large extent *context-independent* (contrast Smolensky 1988). We obtain these classical properties without having to assume a syntactic level of representation tokening atomic symbols and complex symbols having atomic symbols as their parts.

It turns out not to be too difficult to build a distributed connectionist system with these capabilities. The system is a kind of micro-world animal, but in contrast to the animal described by F&P (1988) this animal masters systematicity: 'Such an animal would be, as it were, *aRb* sighted but *bRa* blind, since presumably, the representational capacities of its mind affect not just what an organism can think, but also what it can perceive. In consequence, such animals would be able to learn to respond selectively to *aRb* situations but quite *un*able to learn to respond selectively to *bRa* situations. (So that, though you could teach the creature to choose the picture with the square larger than the triangle, you couldn't for the life of you teach it to choose the picture with the triangle larger than the square).'

4. System 1

We have built a network with distributed representations having the properties described above (Bernsen and Ulbaek 1992a). The simulated network learned how to apply the concept 'to the right of' through being trained on pictures of discriminably different 2-D objects. A semantic difficulty had to be overcome. The concept 'to the right of' is closer to perception than is 'loves' and has a simpler and less exciting semantics, but its semantics is not that simple either since it has an asymmetrical trajectory-landmark structure (Langacker 1987). When an object, *a*, is said to be to the right of another object, *b*, then object *b* acts as landmark for the trajector *a*. To capture this property, we placed one object at the centre of the 2-D array whenever the presentation contained a landmark object. This method of placement gave landmark status to the object without the need for separate labels for any of the objects used in the simulation in addition to their different visual appearances and positions. It may be assumed that the concept 'to the right of' is normally learned only by creatures which have independent concepts of the objects perceived in the scene. The setup described circumvents this difficulty without giving way on the crucial issue of systematicity. It might be objected that the system does not learn the completely general concept 'to the right of' but only learns the concept 'to the right of a fixed landmark'. This is true, but we did not consider the objection serious with respect to the principles we wanted to demonstrate. As a matter of fact, our common 'to the right of' -concept is even more complicated than that since it also allows us to change coordinate systems from a viewer-dependent coordinate system to an object-centered coordinate system. Again, this does not affect the central point of the demonstration.

The system also had to learn that 'to the right of' is a two-place predicate. When there is only one object in the scene, or when there are more than two objects, the question whether 'this object is to the right of that object (the landmark)' either does not make sense or is ambiguous. In such cases, the system answered 'no' to the question posed to it. On all other presentations one object was placed at the landmark site. A second object was then placed in one of four different positions around the landmark object (right, left, above, or below). We did not teach the network to discriminate among all those positions but simply to respond with a 'yes' if and only if the trajector was positioned to the right of the landmark, and to respond with a 'no' otherwise. In this way, the network was answering the question: 'Is the trajector to the right of the landmark?' If there was no landmark, it responded with a 'no' and if there were three or more objects present, it also responded with a 'no'.

The network was a standard one-layer backpropagation network with graphics facilities for the display of presented objects and running on a PC. The training tolerance for output was 0.1, which means that on a scale from 0 to 1 the network would count 0.9 as correct and stop training when all exemplars in the training set perform above 0.9. The testing tolerance was 0.4 which is sufficient for mechanically distinguishing success and failure. The 2-D picture array measured 8 x 20 (160 input units). The hidden layer had 30 units and the output layer had 2 units for 'yes' and 'no', respectively. The training set consisted of 6 different objects which were placed in different numbers, positions and combinations and sometimes as landmark, sometimes as trajector. The landmark site and each trajector site consisted of a field of 4 units. The different objects occupied different numbers and combinations of units at a site. The test object set included 3 objects different from the 6 in the training set (see Fig. 1).

[Insert Figure 1 around here]

Figure 1. (a) shows the objects in the training set. (b) shows the new objects in the test set. (c) shows an input example.

To demonstrate that the network could handle the systematicity of *aRb* and *bRa*, we only trained the network on one of these relations for a given pair of objects while saving the second relation between the pair for the test. Thus, (1) if the network had been trained on '*a* is to the right of *b* ', it was not trained on '*b* is to the right of *a* '. In the test phase, the network was shown already familiar objects in combinations it had not encountered before. In addition (2), the network was shown objects it had not encountered before in order to verify that it was able to abstract a sufficiently general concept of '2-D spatial object'. Taken together, (1) and (2) offered sufficient evidence that the network was able to master systematicity from *aRb* to *bRa*; abstraction to the 'right of' concept which we are used to representing as *xRy*; and abstraction over all possible objects in its world, thus successfully taking the set [*a*,*b*,.....,*i*] as instantiations of *x* and *y* or as legitimate arguments of the relation *R*.

The training file consisted of 84 training exemplars. The network converged on the desired output in 24 epochs with the mentioned training tolerance of 0.1. The test file consisted of 65 test exemplars. The network was able to generalise successfully within the testing tolerance of 0.4. In other words, systematicity is so simple that a mouse could probably achieve it if its cognitive architecture consists of distributed connectionist networks. It is, therefore, we concluded, *no* mystery why nature contrives to produce only systematic minds.

5. Objections to System 1

System 1 did not, however, persuade our colleagues that we had fully met F&P's systematicity challenge to distributed connectionism. This section reviews some of their objections.

Objection 1: 'The system is merely a pattern-matcher.'

The simplest reply to this objection is: So what? It is not clear why this is an objection as long as the difference between 'mere' pattern-matching and full-fledged systematicity has not been defined. And it is not clear that this distinction *has* been defined.

Objection 2: 'System 1 does not process semantically complex thoughts.'

Here is a simple counter-claim: It does ! The trouble is perhaps that we both lack an agreed upon measure of levels of 'semantical complexity' and a clear definition of 'thought' which are not influenced by controversial theoretical prejudice. Intuitively, at least, the thought that something is to the right of something else is not a semantically simple one. This objection leads to the following

Objection 3: 'The system only recognises *that some 2-D spatial object is to the right of some other 2-D spatial object.* This is not systematicity.'

This is an interesting objection because of the need to explain *why* the representations of System 1 do not exhibit systematicity. The interesting point is that it is far from clear that such an explanation can be provided. Until such an explanation may be forthcoming, let us consider a slightly weakened version of Objection 3:

Objection 4: 'Well, at least it was not the kind of systematicity we had in mind. In fact, the kind of systematicity demonstrated (if any) does *not* satisfy the R(a,b) requirement. It only satisfies the following, weaker requirement: ∃x & ∃y such that x and y are numerically different spatial objects & R(x,y). Basically, your system cannot distinguish between R(a,b) and R(b,a). It will respond with 'yes' when either of these spatial relations obtain in the scene.'

This is certainly true and to the point. The following reply does not really counter Objection 4: Since System 1 works with real semantic material (objects in visual scenes) it is very probable that its pattern of activation is different when it sees right-of(triangle, square) and right-of(square, triangle).

Objection 4 continued: 'This won't suffice. Here are the crucial points:

(1) Thought at least involves *the possibility* of expression of distinctions made, whether in motor behaviour, language or system (or module) output. A difference in patterns of activation is not sufficient for different thoughts to occur.

(2) Thought requires *general concepts*. A dog perceiving a visual scene containing, i.a., a radio, does not necessarily see (that object as) a radio. Dogs probably lack the concept of a radio, as betrayed by their behaviour.'

The upshot of this discussion is that systematicity, in some unanticipated sense, may well have been demonstrated by System 1 but that a more complex form of systematicity, one formal expression of which is the relationship between R(a,b) and R(b,a), has yet to be demonstrated by distributed connectionist systems. Just as importantly, however, this will not change the basic structure of the argument of Sect. 4 above. We already have systematicity and complex thoughts.

6. System 2

The requirements for the new system, System 2, are straightforward in view of what has been said above. The distributed connectionist System 2

- should produce different output for R(a,b) and R(b,a) or, e.g., for right-of(triangle, square) and right-of(square, triangle);
- should necessarily have systematicity; and
- the cognitive mechanisms behind systematicity should be made clear at least to the extent that its necessity has been explained.

For the System 2 simulation a recurrent net was used in order to (1) maintain strict classical syntax in output expression, and (2) fit the hypothesis that thought of a certain complexity may be temporally extended. Point (1) ensures a clean interface between the cognitive module which entertains the thought based on visual perception that, e.g., R(a,b), and a language generation module. Point (2) at least begins to address the topic of the two first objections in Sect. 5 above. That is, there may actually be differences in complexity between thoughts so that some thoughts need a temporal dimension for their representation whereas other, simpler, thoughts do not. However, this is a hypothesis so far and it is quite possible that ordinary backpropagation networks such as the one used for System 1 might solve the problem addressed by System 2, only using a less clean output syntax. It may be mentioned that a first version of System 2 using recurrent nets and exhibiting the feature of visual attention in order to distinguish between the performance of different cognitive tasks has been described in Bernsen and Ulbaek (1992b). However, the capacity for generalisation of that version was not tested and hence it never demonstrated necessary systematicity.

System 2 is a Tlearn network (due to Jeff Elman, UCSD) running on a SUN. When the network is in recurrent mode there is a copy layer in addition to the hidden layer. The copy layer is used to copy the activity of the hidden layer at time t1. At time t2 the activity of the hidden layer at t1 is fed back into the hidden layer from the copy layer. In this way the network is sensitive to earlier input activity and is able to produce dynamic, temporally discrete output based on static input. System 2 has 100 input units corresponding to a 10 by 10 visual matrix or scene on which 2-D spatial objects of various shapes are shown to it. The hidden layer and the copy layer have 100 units each, and the output layer has 6 units each one of which, after training, permanently corresponds to one particular named object in the scene. 6 different objects, called A, B, C, D, E and F, respectivly, were shown to the system, each occupying part of a 4 by 4 field within the matrix (see Fig. 2). 6 objects allow 30 different right-of orderings between them. 24 of these were used for training and the remaining 6 orderings were used in the test phase. The 6 orderings used in the test phase were AB ('B is to the right of A'), BC, CD, DE, EF and FA. Each static input scene is presented to the network in two consecutive time slices. It is the output which is time-dependent or coded serially. The coding for the presence of, e.g., a anywhere in the scene is:

t1: 1 0 0 0 0 0
t2: 0 0 0 0 0 0

Coding for, e.g., a is to the right of b (or BA) is:

t1: 0 1 0 0 0 0
t2: 1 0 0 0 0 0

Coding for b is to the right of a (or AB) is:

t1: 1 0 0 0 0 0
t2: 0 1 0 0 0 0

In the simulation, no output unit was introduced for expressing the *right of* relation itself. Since the network just recognises one type of spatial relation such an extra output unit is unnecessary, the temporal ordering of the output unit activations performing the distinction between, e.g., right-of(a,b) and right-of(b,a). If the network were to be able to recognise more than one kind of spatial relation (e.g., *above(x,y)* as well), units identifying the type of spatial relationship currently attended to would have to be introduced.

[Insert Figure 2 around here.]

Figure 2. (a) shows the 2-D objects used. (b) shows a training input situation in which A is to the right of D.

System 2 was trained and tested on two successive cognitive tasks:

- In *Task 1,* the system learns to identify discriminably different individual objects present anywhere in the visual array; e.g., when object *a* is present, the *a* output unit fires. This way, an *individual object recognition task* is performed requiring the system to develop concepts for each of the spatial objects or object types presented to it.
- In *Task 2,* the system learns to identify right-of-ness as obtaining between identified individual objects present anywhere in the visual scene. For instance, when right-of(a,b) obtains in the visual scene, the output unit representing object b fires *before* the output unit representing object a; when right-of(b,a) obtains, the output unit representing object a fires *before* the output unit representing object b. This way, a *right-of task* is learned which is different from, and more complex than that performed by System 1. Subsequently, System 2 was tested on instances of right-of(x,y) which it had not encountered during training to verify that the system had succeeded in generalising to the concept right-of(x,y) and was able to instantiate to arbitrary combinations of the individual objects known to it.

Recent results from simulations using the version of System 2 described above demonstrate, in my view convincingly, that a distributed connectionist system can solve the systematicity problem posed to System 2. Final work on the simulation will be done in order to obtain a 100% clean test score but does not appear to have any theoretical significance otherwise. The same lack of theoretical significance, it is claimed, would characterise attempts to make the system handle additional semantical complexity to the right-of-ness problem discussed here, of which there are many. The simplest one is the following. The current system knows of 6 different visual objects (or object types) and demonstrates necessary systematicity in its ability to identify right-of-ness with respect to those. There is little doubt that the system, if it were to be taught the identities of new objects, would be able to demonstrate necessary systematicity with respect to those as well. Further semantic complexity abounds even in as simple a concept as right-of(x,y) as indicated elsewhere in this paper. System 2 implicitly masters the trajectory-landmark structure described above just as biological systems do. Other aspects of conceptual complexity are not mastered by System 2. These should be handled eventually by distributed connectionist systems, of course, but not as part of a discussion of the systematicity problem. Rather, they should be investigated as part of a research programme in spatial cognition and in the linking of visual and linguistic processing of information.

The training set consisted of the 6 individual objects and the 24 training right-of orderings located at two thirds of the possible locations in the scene. The input consisted of 1260 time sequences corresponding to 630 different scenes and the network was trained for 2.000.000 epochs. After training the error measure (the total sum of squares) was < 0.05. This is hardly sufficient for a 'clean' simulation and explains the remaining errors in the test phase (see below). However, it suffices for the point to be demonstrated here. Nothing was done to make the simulation run quickly and efficiently. The test set consisted of 168 time sequences corresponding to 84 different scenes showing 6 different right-of orderings in two thirds of the possible locations in the scene.

a	b	c	d	e	f
0.997	0.018	0.000	0.000	0.000	0.000
0.023	0.997	0.000	0.000	0.000	0.000
0.374	0.030	0.000	0.000	0.000	0.002
0.000	0.546	0.000	0.002	0.000	0.101
0.950	0.006	0.000	0.000	0.000	0.000
0.002	0.998	0.000	0.000	0.000	0.215
0.992	0.422	0.000	0.000	0.000	0.000
0.000	1.000	0.000	0.000	0.000	0.005
1.000	0.002	0.000	0.000	0.000	0.000
0.000	0.969	0.000	0.005	0.111	0.000
0.974	0.000	0.000	0.000	0.000	0.000
0.001	0.999	0.000	0.000	0.000	0.000
0.998	0.000	0.000	0.000	0.000	0.000
0.009	0.627	0.000	0.000	0.000	0.004
0.001	0.998	0.000	0.000	0.000	0.000
0.001	0.947	0.000	0.001	0.002	0.004
0.980	0.003	0.000	0.000	0.000	0.001
0.002	1.000	0.000	0.000	0.000	0.003
1.000	0.022	0.002	0.000	0.000	0.002
0.000	0.994	0.000	0.000	0.003	0.006
0.837	0.000	0.000	0.000	0.000	0.000
0.075	0.638	0.000	0.001	0.072	0.003
1.000	0.000	0.000	0.000	0.000	0.000
0.632	1.000	0.000	0.000	0.000	0.001
0.990	0.016	0.000	0.000	0.000	0.011

0.000	**0.995**	**0.000**	**0.091**	**0.000**	**0.000**
1.000	**0.000**	**0.018**	**0.000**	**0.000**	**0.000**
0.002	**1.000**	**0.000**	**0.000**	**0.000**	**0.000**

Table 1. Test output for System 2 perceiving that b is to the right of a at 14 different scene locations. Two consecutive rows show output corresponding to one static scene.

In the test phase, System 2 successfully solved the right-of problem in 57 out of 84 cases if one uses the criterion for success that output from the units corresponding to the observed objects in the proper temporal sequence has to be > 0.5 and output from all other units has to be < 0.5. This is, however, just one way in which correct results can be mechanically separated from incorrect results. Using the much sharper criterion for success that output from the correct units in correct temporal sequence has to be > 0.9 and output from all other units has to be < 0.1, System 2 successfully solved the right-of problem in 36 out of 84 cases. However, using an equally mechanically applicable winner-takes-all strategy for success according to which only the two most active output units are selected, System 2 successfully solved the right-of problem in 73 out of 84 cases. Output for the b is to the right of a -problem is shown in Table 1.

7. Discussion of System 2

It has been established that distributed connectionist nets can learn the following from experience:

- concepts of different 2-D spatial objects or object types (a, b, c, etc.) (System 2);
- the concept of 2-D spatial-object(x);
- the concept of right-of(x,y) as obtaining between unidentified individual 2-D spatial objects (System 1);
- the concept of right-of(x,y) as obtaining between identified individual 2-D spatial objects and, in consequence, the distinction between, e.g., right-of(a,b) and right-of(b,a) (System 2).

It seems justified to conclude that no further simulation experiments are needed to demonstrate that distributed connectionist nets can exhibit necessary systematicity in the case of a class of two-place asymmetrical predicates. The precise identification of the class of predicates for which distributed connectionist systems are able to exhibit necessary systematicity is not relevant to the systematicity debate.

Moreover, the cognitive mechanisms producing systematicity in System 2 are the same as in System 1. The mechanisms are distributed connectionist versions of abstraction and generalisation from experience, and instantiation to experience. It has been known for a long time that distributed connectionist systems are good at performing these conceptual operations.

It is not surprising that System 2 produces necessary systematicity using the very same cognitive mechanisms as does System 1, for the following reason. The spatial input scheme for right-of-ness (R(a,b); R(a,c); R(c,b); etc.) is *itself systematic*. That is, it involves the same tokens/types, the same spatial relation, and two different types of spatial order due to the asymmetry of the predicate involved. Hence this input scheme can be subjected to the basic cognitive mechanisms of abstraction, generalisation and instantiation. It is therefore tempting to propose the general hypothesis that, given a systematic input problem, distributed connectionist systems will necessarily exhibit systematicity. Only the complexity, in some sense to be discovered, of some class of systematicity problems may limit the scope of the hypothesis just stated. However, this complexity issue is only accidentally related to the systematicity problem discussed here.

Given the performance of System 2, it appears highly likely that a system of this kind might exhibit necessary inferential systematicity (cf. Sect. 2 above). Inferential systematicity would be found in a system which, necessarily, if it were able to conclude *a* from, e.g., *a&b* and *a&b&c&d&e,* would also be able to conclude *a* from, e.g., *a&b&c* (Fodor and Pylyshyn 1988). What a system with inferential systematicity has to learn is, e.g., an abstract schema for propositional inference from conjunctions. The abstraction is the following. It does not matter how many conjuncts (or members) you have in a set or which they are: you are always allowed to infer a subset of the conjuncts from the set. We have already seen the general character of the distributed connectionist mechanisms needed for this cognitive task.

It would appear that the systematicity problem has been solved to the advantage of distributed connectionism. However, there is a final twist to the issue which will be addressed in the next section.

8. Mere Implementation?

If distributed connectionist systems produce necessary systematicity, there seems to be only line of objection left for maintaining that the classical symbolic paradigm is the only general computational paradigm for cognitive science. This line of objection is to claim that the demonstration provided above, just like all other purported demonstrations of systematicity in connectionism to date, is just another demonstration that distributed connectionism may implement classical cognitive architectures (cf. the assumption made in Sect. 3 above). In other words, despite all that has been said so far, Systems 1 and 2 merely implement classical syntax and that is why they demonstrate necessary systematicity.

On the face of it, this claim may appear quite unreasonable. So far, we have had no reason whatsoever to believe that distributed connectionist systems have classical syntax. On the contrary, their difference from classical systems is precisely that they do not have classical syntax. They take the same (systematic) input material as do classical systems and they have the same (systematic and, by design, syntactic) output performance as have classical systems. But what happens in between input and output is done by processing algorithms that are profoundly different from classical algorithms. Moreover, the entire explanation provided above of why distributed systems produce necessary systematicity has been offered at *the cognitive level via* conceptual operations such as abstraction, generalisation and instantiation scoping concepts and relations. So, again, the objection we are now considering would appear preposterous. Why is this objection still to be anticipated, all appearance to the contrary notwithstanding? In attempting to answer this question, I shall distinguish between three different issues. Once this has been done, it may have become clear why the burden of proof is now on the classicalists if the systematicity issue is to be kept alive.

1. Information processed and the processing of that information: The question whether distributed connectionist systems merely implement classical syntax is, in fact, ambiguous. The ambiguity lies in the *task (or input) domains* addressed by those systems. Here are the possibilities. We may be dealing with either:

(a) distributed connectionist systems working within the task domain of explicit symbols and classical syntax, possibly modified by the addition of statistical linguistic material in accordance with current trends in computational linguistics. One example among many is the parallel distributed processing of a combination of hard symbolic and syntactic rules and soft symbolic linguistic rules which incorporate corpus-derived numerical values. Of course, even in this case adherents of distributed connectionism maintain that no classical syntax is involved in the actual processing of input domain information. Such distributed connectionist systems may realise a universal Turing (LISP) machine which has no complexity boundary to its computations whatsoever (e.g., Smolensky et al. 1992);

or with

(b) distributed connectionist systems working within task domains that contain no explicit symbols and no explicit syntax. The example at hand is the parallel distributed processing of 'real' semantic information (i.e., visual representations) done by Systems 1 and 2 above. Again it is claimed that no classical syntax is involved in the actual processing of input domain information. Of course, the input information is itself *systematic* (cf. above), but that is an obvious requirement for the systems to be able to address the systematicity problem in the first place.

The relevant difference between cases (a) and (b) is the following: In (a), symbolic and syntactic information is *literally being processed* albeit in a distributed connectionist manner. This information is out there in the input. It is therefore not surprising if Smolensky et al. are still in for the criticism that they merely implement classical syntax. They are actually close to admitting as much themselves (Smolensky et al. 1992). In (b), on the other hand, there is no explicit symbolic information (rules or symbolic representations) anywhere, except in the output. In this case, symbolic and syntactic information is not literally being processed. It does not make sense to speak of 'mere implementation' here because we have an explanation at the right, i.e., cognitive level involving concepts, relations, order, abstraction, generalisation and instantiation. So the ambiguity mentioned is one between two completely different issues: (1) *what* (input) information is literally being processed? (2) *how* is the information being processed?

It is not the task of the present author to advise Smolensky and other distributed connectionists on anything. But it does seem odd, if not downright absurd, if the purportedly fundamental problem of systematicity can be solved, for certain classes of input domains, at least, such as the literal processing of symbolic and syntactic information, by merely looking at the task or input domain addressed by a distributed connectionist system, as follows: Is your distributed system processing explicit linguistic information? 'Yes.' But then your system is a mere implementation of a classical architecture! Some may still want to pursue this line of argument against one class of distributed connectionist simulations purporting to demonstrate systematicity but that will not be done here. In any case, such a strategy of argument would not be a general one because there are important classes of input domains relevant to the systematicity problem which do not involve explicit linguistic information, such as those of the Systems 1 and 2 above. The two remaining issues to be dealt with in this section address the general case.

2. The cognitive level and the algorithmic level: The story to be told about System 2 at *the cognitive level* goes as follows. System 2 learns the concepts of individual objects (or object types) which can enter into the spatial relation right-of. This way, System 2 is well under way to acquiring a general concept of 2-D spatial object(x) as denoting the kind of entities which may enter into this relationship. Furthermore, System 2 learns the general concept of the ordered relation right-of(x,y) in such a way that, given two individual 2-D spatial objects (or object types) known to it, it is capable of deciding which object is to the right of which other individual object. I see no reason why classicalists and distributed connectionists might not agree on this story about how any cognitive system identifies a right-of relationship between two different spatial objects known to it. They may disagree on the nature of the concepts and relations involved, but that is a different matter to be considered below.

At *the algorithmic level,* however, System 2 involves no classical syntax and no classical constituents since the algorithms work at the level of individual processing units in a distributed system of representation. When System 2 perceives *aRb* and outputs correctly, it must clearly activate its representations of *a*, *b* and *right-of-ness* and create the correct representation *aRb*. Given the distributed nature of its processing of information, there is every reason to assume that what happens is not that the representations of *a* and *b* are separately activated and, as such, with no further changes to their activation patterns enter into a relation with an equally separately activated representation of *right-of-ness.* Distributed connectionist systems do not work that way since they do not have classical constituents, but classical systems do. Similarly, when System 2 perceives *bRa* and outputs correctly, it must activate its representations of *a*, *b* and *right-of-ness* and create the correct representation *bRa.* It is possible that more or less the same set of units become active in this case as in the case of *aRb,* but the resulting vector which determines the output has to be different in the two cases.

3. Implicit and explicit internal systematicity: If this is still not the way to solve the systematicity problem, we are back to the intriguing possibility that System 2 above is a mere implementation of a classical and syntactical cognitive architecture. How could this be decided? The point is that *we don't have a clue.* We do not have the faintest idea of how, by looking into the way that System 2 processes information, we were to decide that, all appearances to the contrary notwithstanding, System 2 is nevertheless a mere implementation of a classical architecture (Brian McLaughlin has admitted as much in personal communication. This admission may be temporary, of course.).

So here are the options: We either deem the systematicity problem solved along the lines indicated or we wait until we get a clue as to how it may be decided whether, e.g., System 2 really is an implementation of a classical cognitive architecture. *If* such a clue or, rather, set of criteria, come forward; *if* they are operational (i.e., intersubjectively applicable to real systems); and *if* by using them it turns out that System 2 is an implementation of a classical architecture, then the current author would have to confess to being a classicalist after all. Meanwhile, there are good reasons to conclude that the existence proof of just one distributed connectionist system, whatever be its input domain, which is able to exhibit necessary systematicity is all it takes to refute the classicalist position. For the same good reasons, if, at the end of the day, this turns out not to be true, it is quite possible that the classicalist position will have become so rarefied in the process of further development that nobody would object to it.

Let us look at the foothold, as it were, of the classicalist scenario just outlined. It is that connectionism will have solved the systematicity issue only when it has provided a so far missing level of account of distributed processing of information:

> '..[connectionism must] explain the existence of [systematicity] without assuming that mental processes are causally sensitive to the constituent structure of mental representations.' (Fodor and McLaughlin 1990).

It is not clear why connectionism should accept this challenge. Otherwise, the classicalist position runs the risk of becoming, once again, trivially true, which it certainly is not, given its conception of constituent structure. The story told above clearly does assume that the mental processes generating systematicity are causally sensitive to the constituent structure of mental representations. Consider the following simple argument: Systematicity in both the input and the output requires systematicity in between. If this argument is true, then there's got to be causal sensitivity in distributed connectionist systems to the constituent structure of mental representations (i.e., of concepts, relations, order, predicates, etc.). System 2 above, for instance, satisfies the condition that there is systematicity in both the input and the output even though its task domain is not linguistic.

It may now be pointed out that a look-up table, for instance, might satisfy the condition that there is systematicity in both the input and the output. However, the *internal* systematicity in between might be said to be at most *implicit;* it is not *explicit* as in classical systems. System 2 is not a look-up table. This is proved by its capacities for generalisation and generation. The (fortunate) fact that System 2 is not merely a look-up table can hardly by itself constitute a proof of the classicalist position. However, from this fact one might go on to conjecture that if the internal systematicity in between input and output in distributed connectionist systems is explicit rather than implicit, then such systems are implementations of classical architectures after all. It follows that as long as it is not known whether distributed systems have explicit internal systematicity or not, connectionism has not answered the systematicity challenge:

> 'We still don't have a substantive [constituent level] connectionist account of systematicity.'
> (Fodor and McLaughlin 1990).

Instead of using the term 'explicit internal systematicity', Fodor and McLaughlin point out that, on the classicalist view, the representations *aRb* and *bRa* literally have the same parts and call this 'real constituency'. Does System 2 have real constituency? If the criterion for real constituency is that System 2 does represent, in both cases, and at the cognitive level, the external objects *a* and *b* and the relation *R* between them, then the answer is clearly affirmative. If the criterion is the syntactic symbol system hypothesis, then the answer would *seem* to be just as clearly negative: both System 1 and System 2 just behave as if they know of classical syntax but they don’t know of classical syntax. The very same units, more or less, may be active in both the *aRb* and *bRa* cases, but so far this has not been stated as a criterion for satisfying the classicalist position. If exactly the same units are active in the two cases, should they be equally active? Or should they be equally active within a certain margin of tolerance? Which margin? Why? If not exactly the same units are active in both cases, which is possible and even likely, then what? Are there principled margins of tolerance to be observed? Why? In other words, we simply don't know what 'real constituency' means in distributed connectionist terms. The same argument *seems* to apply to the question whether System 2 has 'combinatorial constituent structure'. It has - at the cognitive level. At the algorithmic level we have no idea of what that might mean in distributed connectionist terms.

Finally, the claim *seems* clearly justified that the structure of the complex representation which System 2 has (e.g., *aRb* or *bRa*) has a causal role in the generation of its behaviour: the system responds systematically differently depending on which one of these complex representations it has. In F&P's terminology (1988), when the system's representations of *a, b,* and *R* are simultaneously active and the system has the complex representation *aRb,* then we also have to admit that the system's representations of *a*, *b*, and *R* enter into a specific kind of 'construction' with each other. This construction differs from the construction among *a, b,* and *R* when the system has the complex representation *bRa*. So, the constituency relations are themselves semantically significant at the cognitive level as F&P claim they should be. At the algorithmic level, however, we do not know what the kind of 'construction' required is.

In summary, then, and combining the cognitive and algorithmic levels descriptions, the distributed connectionist representation *a is to the right of b* therefore *seems* to be a *non*-atomic mental representation having *non*-syntactic structure. It therefore *seems* to be just false to maintain that

> '... we cannot have both a combinatorial representational system and a connectionist architecture *at the cognitive level.'*

The classicalist position seems confused because the cognitive and algorithmic levels are not kept distinct in the systematicity argument. Once they are kept distinct, it becomes obvious that what is so far an uncontroversial cognitive level story is being combined with an algorithmic level story to which there exists an equally valid alternative.

It is true that things are not always what they seem. The trouble is, however, that we don't know what to look for in order to decide whether, e.g., System 2 is after all an implementation of a classical architecture. We don't have a clue as to what constitutes a generic concept of 'explicit internal systematicity' such that:

(1) this concept might subsume not only well-known instances of classical syntactic processing of information but also instances of distributed connectionist processing of information characterised by systematicity; and
(2) if this concept were applicable to distributed connectionist systems exhibiting necessary systematicity then these would count as implementations of classical architectures rather than alternatives to them.

Classicalist might want to argue that this concept of explicit internal systematicity is not generic at all: it is simply the classical concept of constituent structure. But this simply won't do. As demonstrated, we have no idea as to how the concept of explicit internal systematicity should be applied to the algorithmic processing of systematic distributed connectionist systems in order to determine whether they implement classical syntax or constituent structure or not. The classicalist position has so far not been geared to answer this question and it is perhaps not likely that it will be able to do so.

9. Conclusion

Several times above I have mentioned the agreement between classicalists and distributed connectionists about what takes place at the cognitive level of description in systems mastering systematicity. There are limits to this agreement, however, and this may turn out to have important implications for the *algorithmic* analysis of systems handling systematicity in 'real' semantic domains. It should not be forgotten that classicalists and distributed connectionists tend to disagree on what are the 'constituents' of thought. This problem *is* a substantial unsolved problem of cognitive science as witnessed by the fact that even within the core domain of the classicalist position, i.e., that of linguistic syntax, it is becoming increasingly clear that expert syntacticians disagree profoundly and perhaps incurably about the correct parsing of ordinary sentences. If the constitutents of thought are to a considerable extent non-classical, i.e., irreducible to the scheme of rules and (atomic and molecular) representations and their tokenings as causal syntactic entities, as distributed connectionists maintain, this would provide substantial evidence against the classicalist algorithmic story of what goes on in systems mastering systematicity and therefore against the classicalist position as a whole.

The conclusion reached in Sect. 7 above would therefore seem to stand up for the time being. Some might want to take this opportunity to leave the increasingly thin philosophical air zone of the systematicity problem and do more substantial cognitive science instead. Cognitive science still needs to provide an acceptable account of concepts, relations and predicates at both the cognitive, algorithmic and implementational levels. It is a qualified guess that this account will incorporate aspects which have been stressed by classicalists as well as aspects which have been stressed by distributed connectionists. Obviously, the demonstration that distributed connectionist systems can exhibit necessary systematicity in a way different from classical systems does not, by itself, give distributed connectionist cognitive architectures any advantage over classical architectures.

Acknowledgements: The research was carried out under grants from the Danish Research Councils for the Natural Sciences and for the Humanities. Their support is gratefully acknowledged. I am indebted to the continued help from Ib Ulbaek and Peter Wolff with the System 2 simulations described in this paper.

References

Ajjanagadde, V. and Shastri, L. 1991. 'Rules and variables in neural nets', Neural Computation 3, 121-34

Bernsen, N.O. and Ulbæk, I. 1992a. 'Two games in town. Systematicity in distributed connectionist systems', Artificial Intelligence and Simulation of Behaviour Quarterly, Special Issue on Hybrid Models of Cognition Part 2, No. 79, 25-30

Bernsen, N.O. and Ulbæk, I. 1992b. 'Systematicity, thought and attention in a distributed connectionist system', Working Papers in Cognitive Science WPCS-92-2, Centre of Cognitive Science, Roskilde University

Bernsen, N.O. and Kopp, L. 1993. 'A connectionist architecture for spatial cognition', (in preparation)

Fodor, J. A. and McLaughlin, B. P. 1990. 'Connectionism and the problem of systematicity: Why Smolensky's solution doesn't work', Cognition 35, 183-204

Fodor, J.A. and Pylyshyn, Z.W. 1988. 'Connectionism and Cognitive Architecture: A Critical Analysis', Cognition 28, 3-71

Langacker, R.W. 1987. Foundations of Cognitive Grammar. Stanford CA: Stanford University Press

Pylyshyn, Z.W. 1984. Computation and Cognition: Toward a Foundation for Cognitive Science. Cambridge MA: MIT Press

Smolensky, P. 1987. 'The constituent structure of connectionist mental states: A reply to Fodor and Pylyshyn', Southern Journal of Philosophy 26, Supplement, 137-161

Smolensky, P. 1988. 'On the proper treatment of connectionism', Behavioral and Brain Sciences 11, 1-74

Smolensky, P., Legendre, G. and Miyata, Y. 1992. 'Principles for an Integrated Connectionist/Symbolic Theory of Higher Cognition', Report 92-08, Institute for Cognitive Science, University of Colorado at Boulder

THOUGHT

A cognitive system's thought concerning what is visually perceived consists of:

(a) visual array information in focus of attention (produced by (an)other cognitive module(s) and processed by the module we are dealing with);

- some of this information is systematically related to other information (same tokens/types and same concepts involved, same relation, possibly different types of order of elements) and hence can be subjected to abstraction, generalisation and instantiation;

- the information is rather discrete and can be handled by basically context-independent predicates;

- the decision to apply a certain predicate to what is perceived is often, but not always, discrete (fuzziness, competition with other relevant predicates and other kinds of uncertain information);

(b) distributed representation and processing of that information;
(c) the possibility of providing discrete (and possibly temporally sequenced) output of that information to other cognitive modules (could have continuous dimensions from, e.g., competition with other predicates).

The necessary systematicity of thought is established and preserved throughout by well-known mechanisms of distributed connectionist systems.

Already S1 has thoughts. It also has systematicity, although (somehow) at a lower level than S2.

If we say that S1 doesn't have systematicity, then there are thoughts which have no (in that case undefined) 'systematicity'.

If distributed connectionist systems are 'mere implementations' of systematic information and its output expression in behaviour, language or other ways, then so are classical syntactic systems.

MODULARITY

S1 and S2 fit the modularity idea:

- (a) visual module(s) produce(s) scene information;

- a cognitive task module produces thoughts about the visual scene;

- a module (which could be the cognitive task module itself) produces focus of attention on information of a certain kind in the scene;

- since the (S 2) output from the cognitive task module is basically temporal and discrete, it is appropriate for a syntactic-semantic language generation module (for subsequent speech production).

COGNITIVE TASKS

The cognitive tasks above are all perceptual recognition tasks (of recognising objects and their relations).

The goal of such tasks is whether to apply a certain predicate or not (many kinds of uncertainty and continuity may be involved in specific cases).

Such tasks require a focus of attention -mechanism because perception is information-rich (could be hard-wired/innate).

Distributed connectionist systems can't help being systematic in solving them (when they do solve them).

COGNITIVE TASK-LEVEL ACCOUNTS

We have an account at the cognitive task level which makes systematicity a necessary product of distributed connectionist systems.

Such accounts answer the following questions:

- what has to be represented, what is the semantics of what has to be represented?
- which cognitive tasks does the system have to perform?
- what representations (thoughts) does the system need?
- which cognitive processes can enable the system to acquire and apply those representations while ensuring the necessary production of systematicity?
- what should be the nature of the system's output?
- which types of connectionist systems can do the tasks?

INFERENTIAL TASKS AND INFERENTIAL SYSTEMATICITY

Example: When trained on deducing a from, e.g., a&b, a&b&c&d, etc. in the training set, deduce a from a&b&c in the test set.

Hypothesis: There is no obvious reason why distributed connectionist systems could not necessarily exhibit inferential systematicity through applying the mechanisms of abstraction, generalisation and instantiation.

6. Concluding remarks

Perhaps the most interesting point about systematicity is that it is a function of the abilities to abstract general concepts from the perception of particular instances and to instantiate those concepts to new (or old) instances. Whether or not the network had actually met with a training case of *bRa* having already been shown a case of *aRb*, is of much less importance than its ability to perform concept abstraction since it is the latter which achieves systematicity and compositionality. As for compositionality, the necessary capability to handle *bRa* once *aRb* can be handled follows directly from the mastery of abstraction and instantiation since it is one and the same abstract concept which is being applied in both cases. And the mechanisms in distributed connectionist networks for handling abstraction and instantiation provide alternatives to the syntactic mechanisms of variables and variable binding.

However, our solution to the representational systematicity of thought might already have provided the mechanism needed for inferential systematicity. What a system with inferential systematicity has to learn is an abstract schema for propositional inference from conjunctions. The abstraction is the following. It doesn't matter how many conjuncts you have in a set or which they are: you are always allowed to infer a subset of the conjuncts from the set. We have already seen the general character of a distributed connectionist mechanism needed for this purpose. We speculate that this mechanism will also work for inferential systematicity.