A Connectionist Model of Reflective Reasoning Using Temporal Properties of Node Firing

Jacques Sougné

Department of Psychology
University of Liège
4000 Liège, Belgium
sougne@vm1.ulg.ac.be

Abstract

This paper presents a connectionist model of human reasoning that uses temporal relations between node firing. Temporal synchrony is used for representing variable binding and concepts. Temporal succession serves to represent rules by linking antecedent to consequent parts of the rule. The number of successive synchronies is affected by two well-known neurobiological parameters, the frequency of neural rythmic activity and the precision of neural synchronization. Reasoning is predicted to be constrained by these variables. An experiment manipulating the amount of successive synchronies is presented. Experimental results would seem to confirm the predictions.

Introduction

Shastri & Ajjanagadde (1993) described SHRUTI, a connectionist model of tractable reasoning. They intended to build a neurally plausible model based mainly on temporal properties of observed rythmic neural firing patterns in the brain. Certain neurons tend to oscillate in synchrony at a frequency of 30 to 80 Hz. (i.e. every 33 to 12 ms). The model of Shastri & Ajjanagadde used this property in the attempt to solve the binding problem. Variables and their respective contents fired in synchrony making appropriate bindings between roles and fillers. The use of this temporal property to solve the binding problem was first studied by Clossman (1988).

Their model is able to draw inferences with great efficiency. But reasoning is limited to what Shastri & Ajjanagadde call *reflexive reasoning.* This concept describes the type of reasoning that people do effortlessly, immediately, almost reflexively. It is contrasted with *reflective reasoning,* demanding more effort and attention and taking more time. Deductive reasoning that has been studied extensively in cognitive psychology pertains to the class of reflective reasoning. In this paper I will attempt to extend the properties of Shastri & Ajjanagadde model to allow it to also handle reflective reasoning. From the model, I will derive certain predictions and provide empirical evidence substantiating these predictions.

There is neurobiological evidence for considering synchronization as a binding mechanism (Gray & al. 1989, Singer 1995, Nelson 1995). The interval between two spikes of a neuron (π) is approximately 12 to 33 ms. Inside this interval a number of different synchronies can occur.

This number is constrained by π and by ω, the width of a window of synchrony. Nodes that fire within a lag of less than $\omega/2$ are considered to fire in synchrony (figure 1).
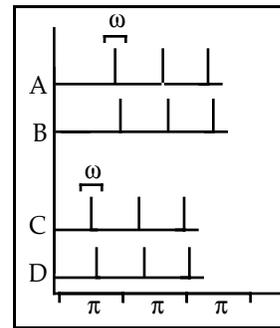

Figure 1: A representation of rhythmic activation and synchrony. Node B fires in synchrony with node A and node D fires in synchrony with node C (since the precision is less than $\omega/2$)

The precision of synchronization between neurons has been reported to be approximately 3 ms. (König & al., 1995). This means that ω should be around 6 ms. These values π and ω limit the number of different possible synchronies to around 10. This corresponds to the well known number 7 ± 2, the span of working memory (Miller, 1956, Lisman & Idiart, 1995). Within a particular window of synchrony, however, numerous concepts can fire at the same time. The relevant constraint is the length of the chain of reasoning. As its length increases, the probability for an error to occur increases.

A good example of reflective reasoning is deduction, which has been widely studied. Typically subjects take approximately 5 sec to draw a single deductive inference (Clark 1969). This delay is far longer than those described by the Shastri & al. model for reflexive inference. There are two ways to extend the Shastri & al. model to allow it to produce these inferences with these delays. Either deductive reasoning tasks require, a succession of different reflexive episodes, or they require learning to initially encode the problem, followed by the building of an appropriate chain of reasoning that will link the data to a more abstract rule that enables inference by appropriate bindings.

The first solution requires an explanation of the nature of the relationship between successive dynamic bindings. The

second requires a learning algorithm to build a new chain of reasoning and the presence of abstract knowledge at the end of the chain of reasoning, for example pragmatic reasoning schemas (Cheng & Holyoak 1985).

I choose the second solution, which meant including in the model of Shastri & al. a mechanism for adding windows of synchrony during the reasoning process. In SHRUTI the form of the query must contain all windows of synchrony necessary for a conclusion to be drawn. Deductive reasoning tasks often require adding new windows of synchrony because the variables of the conclusion may be bound to a different content than those of the premises. For example, consider a conditional reasoning task with a rule: "If Tom gives a candy to Gus, then Tom has stolen money from Mary" and a premise: "Tom gives a candy to Gus". In order to generate the correct conclusion ("Tom has stolen money from Mary"), "candy" and "money" cannot fire in synchrony, nor can "Gus" and "Mary". Otherwise there would be confusion about exactly what is stolen or given and to whom it is stolen or given.

Shastri & Ajjanagadde (1993) describe a learning algorithm that enables nodes to fire in synchrony starting with a noisy temporal firing distribution. Learning is not only necessary to synchronize firing of appropriate nodes, but also to link the given rule to a more abstract rule that enables the generation of the conclusion.

Shastri & Grannes (1995) modified SHRUTI to enable it to perform negation and detection of inconsistencies. They added a negative collector node to every predicate. This option was not chosen in the present model since in their extension, the negation of each predicate must be specifically included. This seemed somewhat unrealistic since negation is a general concept and need not be specifically associated with each predicate.

The Shastri & al. model makes use of different kinds of nodes: fact nodes, collector nodes, enabler nodes, etc. According to Ajjanagadde (1994) this set of node types, as well as the distinction between roles and fillers, is unnecessary. He describes a system that uses a set of excitatory and inhibitory links, called a link bundle. These link bundles preserve the qualities of SHRUTI while simplifying the representation. I adapted this solution to the present model.

Description of the Model

Network Structure

The network is composed of a set of nodes which are not fully connected. Each node is either excitatory or inhibitory. Excitatory nodes can send only excitatory messages, and inhibitory nodes only inhibitory messages. While excitatory links bind two nodes, inhibitory links bind a node to a excitatory connection. This enables a temporary blocking of a specific excitatory transmission. (figure 2).

Each connection has a weight which has two functions: determining the probability of transmission of activation, and, in the event of transmission, the amount of activation transmitted. Connections also have a length which determines the delay for the activation to propagate from the afferent node to the efferent node.
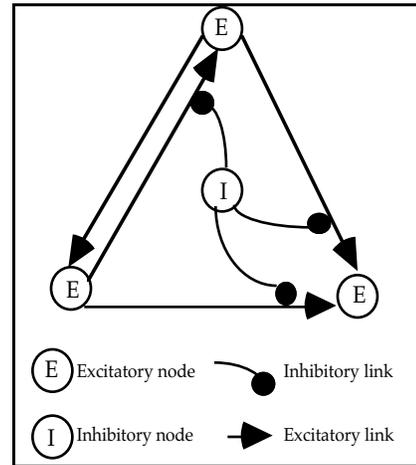

Figure 2: Structure of the net

A node represents a small cluster of neural units (akin to the Hebbian notion of cell assemblies) and a set of nodes firing in synchrony represents a concept or a set of bound concepts. In figure 4 the concept "Consequent" is composed of 4 nodes and is bound to "Lawrence" because they both fire in synchrony (figure 3). In this network architecture, unlike many connectionist architectures, there is no notion of layer.

Activation Propagation

Excitatory activation is stochastically transmitted along connections. Connection weights represent the strength of activation passing and the probability of propagation through the connections. According to their weight, connections are stochastically selected. Once selected, the afferent node receives an activation equal to the value of weight. If the sum of activations arriving at a node reachs a threshold, the node fires. Activation propagation along a connection takes time and the propagation speed is randomly selected from a particular interval which is modified by the learning algorithm. Inhibitory activation follows the same principles, except that activation propagates from nodes to excitatory connections.

Once a node fires, it tends to fire rhythmically with a frequency between 30 and 80 Hz (i.e. between 33 and 12 ms). Following Shastri & Ajjanagadde's (1993) notation, this inter-spike delay is called π. The width of windows of synchrony last 3 to 6 ms and this value is noted ω.

Synchronous firing of nodes either produces variable binding or associates units that participate in a concept representation. The temporal sequence of synchronous firing indicates relations between concepts, enabling the representation of rules. For example, in figure 3, "Antecedent" fires before "Consequent" and stands for the rule: "If *Antecedent* then *Consequent*".

Learning

Learning modifies weights and connection delays on the basis of information from the external world. It is assumed that facts perceived are composed of a series of concepts. The activation of these concepts has a particular temporal order. The task of the learning algorithm is to reproduce this temporal order and to link concepts to the rest of the knowledge base. External activation should therefore be distinguished from internal autonomous activation of the net. Modification of weights and delays depends on the overlap of external and internal activation. In the current model, when activations overlap, the weights of excitatory connections increase, and the delay intervals shift to focus around the most recently selected delay duration. In addition, the weights of connections inhibiting excitatory connections also decrease. When activations do not overlap the process is inverted. Delay learning can rapidly focus on a particular value, enabling nodes to change their partners of synchrony and therefore permitting plasticity in variable binding.

Predictions of the Model

This model predicts that reasoning will take more time and be more difficult as the number of synchronies increases. The number of possible temporal synchronies depends on the frequency of rhythmic pattern or the inter-spike delay (π) and the width of windows of synchrony (ω). When the number of synchronies increases, the frequency must decrease (i.e., π becomes larger) and/or ω must decrease. If this last value decreases too much, the probability of confusion between synchronies increases and the probability of errors increases. In addition, the time for the network to reach a stable state will increase.

Experiment 1

To test the last prediction we ran a reasoning experiment comparing two situations. The first situation requires six windows of synchrony (6ω), the second requires eight windows of synchrony (8ω). Thirty subjects received the following instructions: "A rule written in blue will appear on the screen, you must read it and tell me when you have understood it. Afterward, a statement, written in red will appear on the screen. You should relate it to the rule and tell me what you conclude". Subjects were given a series of similar arithmetic exercises to familiarize them with the procedure. Subjects were randomly assigned to one of the two groups.

In the first group (6ω), subjects received this rule: "If Tom gives a candy to Gus then Tom has given a candy to Lawrence" and the questions appearing in a random order were: "Tom gives a candy to Gus", "Tom does not give a candy to Gus", "Tom has given a candy to Lawrence" and "Tom has not given a candy to Lawrence". This task of conditional reasoning makes reference to material implication. Subjects' inferences are of four types: Modus Ponens $\frac{p \supset q, p}{q}$, Modus Tollens $\frac{p \supset q, \sim q}{\sim p}$ (sound inferences) Denying the antecedent $\frac{p \equiv q, \sim p}{\sim q}$ and Affirming the consequent

$\frac{p \equiv q, q}{p}$ (sound only in material equivalence). Table 1 shows these inferences.

	Modus Ponens	Denying the Antecedent	Affirming the Consequent	Modus Tollens
Group 1 6ω	from *Tom gives a candy to Gus* infer *Tom has given a candy to Lawrence*	from *Tom does not give a candy to Gus* infer *Tom has not given a candy to Lawrence*	from *Tom has given a candy to Lawrence* infer *Tom gives a candy to Gus*	from *Tom has not given a candy to Lawrence* infer *Tom does not give a candy to Gus*
Group 2 8ω	from *Tom gives a candy to Gus* infer *Tom has stolen money from Mary*	from *Tom does not give a candy to Gus* infer *Tom has not stolen money from Mary*	from *Tom has stolen money from Mary* infer *Tom gives a candy to Gus*	from *Tom has not stolen money from Mary* infer *Tom does not give a candy to Gus*

Table 1: instantiation of inferences

Starting with the query: "Tom gives a candy to Gus", the model predicts a rhythmic pattern with 6 windows of synchrony that could be represented by figure 3 and a network state enabling sound inferences as in figure 4. With 6 windows of synchrony, if ω is 3 ms., π must be ≥ 18 ms. (the frequency must be inferior to 56 Hz.); if ω is 6 ms, π must be ≥ 36 ms. (the frequency must be inferior to 28 Hz.). In figure 3, π is equal to 24 ms. (42 Hz.) and ω is equal to 4 ms.
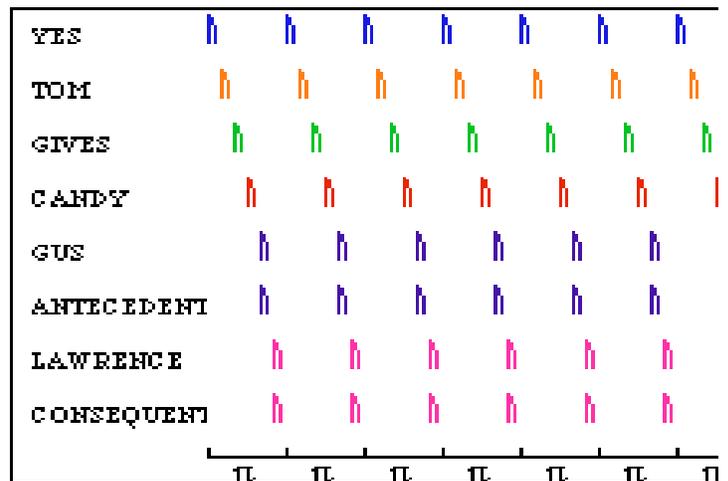

Figure 3: Concepts firing following the query: "Tom gives a candy to Gus" in the situation 1. Gus and Antecedent are firing in synchrony as well as Lawrence and Consequent

Connections represented in figure 4 enable sound material implication inferences. When the premise is "Yes, Tom gives a candy to Gus", activation is propagated from "Gus" to "Antecedent", from "Antecedent" to "Consequent" and from "Consequent" to "Lawrence". When the premise is "No, Tom gives a candy to Gus", "No" inhibits the connection from "Antecedent" to "Consequent" thus prevents any conclusion. When the premise is "Yes, Tom gives a candy to Lawrence", "Lawrence" activates "consequent" but "Yes" inhibits the connection from "Consequent" to "Antecedent" thus prevents any conclusion. When the

premise is "No, Tom gives a candy to Lawrence" no inhibition occurs and activation is propagated to enable "Gus" nodes to fire.
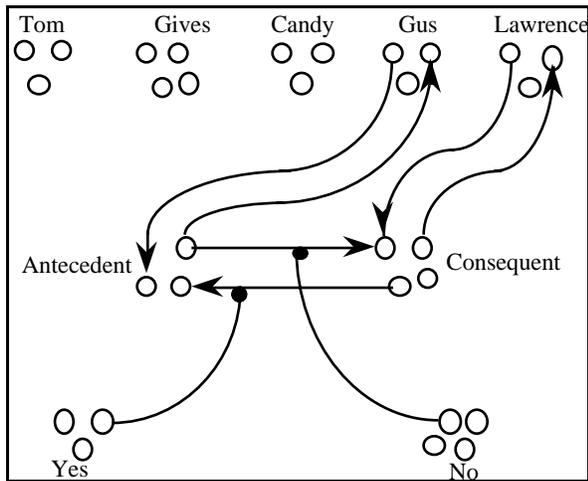

Figure 4: A network configuration that enables sound infe-rences with regard to material implication. Each concept is represented by a set of interconnected nodes although it is not showed. (Note that some nodes could be used by more than one concept, and that only a small number of connections are represented). Affirmation and negation nodes are necessary for sound material implication inference through inhibition mechanism.

Figure 5: Concepts firing following the query: "Tom gives a candy to Gus" in the situation 2. Gives and Antecedent are firing in synchrony as well as Has_stolen and Consequent

For the second group (8ω), subjects received this rule: "If Tom gives a candy to Gus then Tom has stolen money from Mary" and the questions were: "Tom gives a candy to Gus", "Tom does not give a candy to Gus", "Tom has stolen money from Mary" and "Tom has not stolen money from Mary" appearing in a random order. The required 8 windows of synchrony can be represented as in figure 5 and the state of the network as in figure 6. With 8 windows of synchrony, if ω is 3 ms, π must be ≥ 24 ms. (the frequency must be inferior to 42 Hz.), if ω is 6 ms, π must be ≥ 48 ms. (the frequency must be inferior to 21 Hz.) which is out of realistic range. In Figure 5, π is equal to 28 ms. (36 Hz.) and ω is equal to 3 ms.

The connections shown in figure 6 enable sound material implication inferences. For each of the four premises, the activation is correctly propagated.
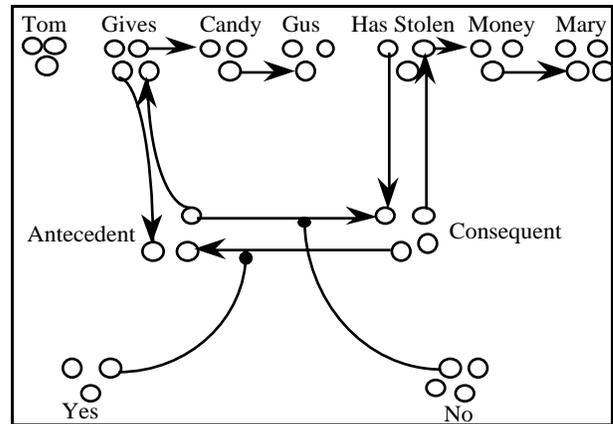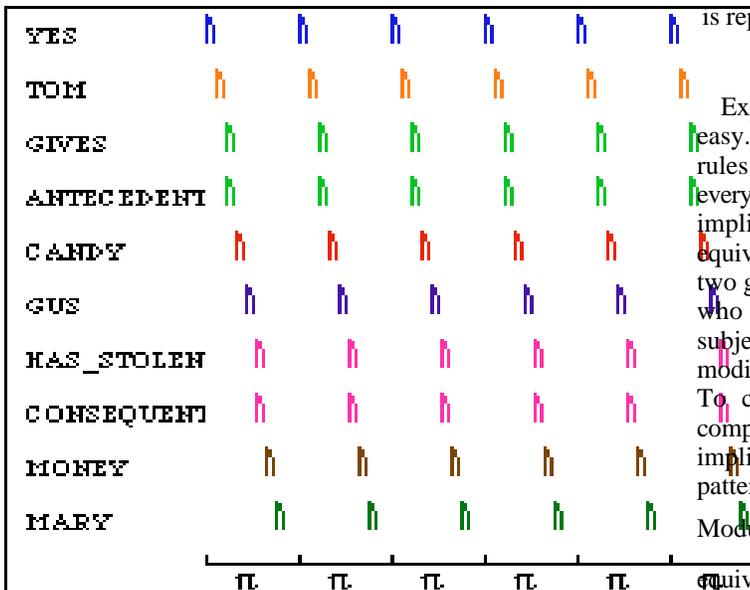

Figure 6: A network configuration that enables sound inferences with regard to material implication. Each concept is represented by a set of interconnected nodes although it is not showed.

Experimentally evaluating the difficulty of the task is not easy. It is well known that subjects often do not follow rules related to material implication. In some situations of everyday life, sound inferences related to material implication are less adequate than these of material equivalence. But we cannot separate these types of people in two groups, those who accept material implication and those who accept material equivalence. A number of biases affect subjects responses (Evans 1989). Content and context also modify the pattern of responses (Cheng & Holyoak, 1985). To compare the performance between groups, we could compare response patterns that are sound for material implication or material equivalence to other types of patterns. The sound inferences for material implication are Modus Ponens $\frac{p \supset q, p}{q}$ and Modus Tollens $\frac{p \supset q, \sim q}{\sim p}$, for material equivalence, Modus ponens $\frac{p \equiv q, p}{q}$, Denying the antecedent $\frac{p \equiv q, \sim p}{\sim q}$, Affirming the consequent $\frac{p \equiv q, q}{p}$ and Modus Tollens $\frac{p \equiv q, \sim q}{\sim p}$. We compare subjects with consistent responses to subjects who contradict themselves from one inference to another.

As the figure 7 shows, the proportion of consistent patterns of responses is significantly higher in the group where the presented rule require six windows of synchrony (χ^2 (1)= 5.129 p < .05).
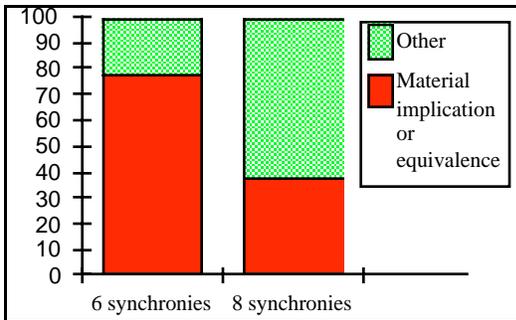
Figure 7: Percentage of consistent and inconsistent patterns of inferences for the two groups.

Even though this criterion measures consistency among successive inferences, it might be criticized as being simply a measure of task difficulty. The time taken by subjects to draw an inference can also be used. The time from the appearance of the question to the response was recorded, giving the averages indicated in Table 2.

Group 1 (6ω)	Group 2 (8ω)
4894 ms	6528 ms

Table 2: Average time in ms. between the question presentation and the response. Student t (103) = -3.046 p < .005

These data show a highly significant difference between the two groups. Subjects doing the task requiring eight windows of synchrony take more time to make an inference than subjects doing the other task. As predicted, when the number of windows of synchrony increases, the likelihood of confusion between synchronies increases, encoding or learning time increases, and the subject's response is delayed.

Psychological data about multiple instantiation tend to show that people deal more efficiently with a problem in which predicates are instantiated more than once. For example, De Soto & al. (1965) showed that subjects make less errors when inferring "A is better than C" from the premises "A is better than B" and "B is better than C" than from the premises "B is better than C" and "B is worse than A". My hypothesis involves an abstraction, a replacement of the multiple premises by a chunked summary that describes the situation. In this example, the predicate extends the number of its possible arguments to combine those of additional instantiated predicates. When it receives the second premise "B is better than C" of the first situation, the system merges the first premise to include C and may transform a representation "Better than: A, B" to "Better than: A, B, C". The second situation requires an additional process, that of using the opposite predicate to transform "B is worse than A" into "A is better than C".

Experiment 2

The above hypothesis concerning multiple instantiations was tested experimentally. However, for the moment, the present model cannot do multiple instantiation. I used two situations involving the same number of instantiations of the same predicate, but involving a different number of transformations in order to make a summarized conclusion. The first group of subjects received these premises: "Allan is in love with Mary", "Mary is in love with Allan", "Peter is in love with Barbara", "Barbara is in love with Peter". The second group received these premises; "Peter is in love with Mary", "Barbara is in love with Allan", "Allan is in love with Mary" and "Mary is in love with Peter". The four premises were presented in a random order to both groups. After reading these premises, subject had to infer which people were happy (i.e. where their love was reciprocated). A higher proportion of subjects gave the correct answer in the first group where relations are all reciprocal (.92), than in the second group (.66). Response times were also shorter for the first group (mean: 3874 ms) than for the second (mean: 10262 ms) Student t (18) = -3.352 p < .005.

These data show that multiple instantiation can be treated rapidly when only a small number of transformations are required to get a summarized representation. Subjects are far less efficient as the number of required transformations increases. However, there should be a mechanism leading to a summarized representation that maintains a trace of previous instantiations while the current instantiations are active.

Discussion

The present system attempts to model not only certain low-level neurobiological facts about synchrony of neural firing but also higher level psychological data on deductive reasoning. The reader can refer to Shastri & Ajjanagadde (1993) for a discussion of the neurobiological plausibility of this type of model.

We know from Johnson-Laird's work (e.g. Johnson-Laird & Byrne, 1991) that reasoning is highly constrained by working memory capacity and we know from Cheng & Holyoak's work (e.g. Cheng & Holyoak, 1985, Holyoak & Cheng, 1995) that reasoning is also constrained by pragmatic principles which are the result of the acquisition and generalization of knowledge. This model attempts to incorporate both levels in a single system. Working memory is conceived as the synchronous and rhythmic firing of nodes. This process is the consequence of the content of long term memory since working memory is viewed as the current activation of long term memory.

The parallel activation of nodes in this model gives it the theoretical possibility of avoiding problems of combinatorial explosion in a scaled-up version. This is also one of the most interesting feature of the Shastri & al. model (Shastri, 1993). Many of the sequential models of deductive reasoning have been shown to be intractable in scaled-up versions (Oaksford & Chater, 1995).

Among the limitations of this model are the problem of the multiple instantiation and catastrophic forgetting in the learning process. Multiple instantiation has been treated by Mani & Shastri (1993). For n instantiations of a concept, the amount of nodes required in their model is n^2. There is no mechanism in the present model for dealing with

multiple instantiation. But experiment 2 could guide our research. The number of simultaneous instantiations seems to be bounded around 2. We could imagine that nodes pertaining to a doubly instantiated predicate would fire at a rythm twice as fast as the others. As rythm is bounded by the refractory period of neurons, multiple instantiation is also bounded. This hypothesis seems reasonable since the frequency of oscillations has been found to be dependent on the amount of activity that a neuron receives.

Catastrophic forgetting is characterized by an interference of the newly learned content on previously learned information making this information lost. In my computational simulation, this limitation means that the problem must be encoded at the same time as the long term knowledge is stored. A separation of the architecture into a "long term memory" component and a "medium term memory" component might be an effective way of dealing with this problem (French, 1994).

Conclusions

A good model of human reasoning should explain the mechanisms by which people solve reasoning problems, the cause of systematic biases, and how the content of a problem affects performance. In this paper, I have focused on a possible low-level mechanism — namely, synchrony — that enables (and perturbs) deductive reasoning. However, I did not focus on systematic biases as described in Evans (1989). It is known that a problem's content affects the reasoning process. The present model attempts to give an emergent explanation for this phenomenon, rather than proposing a specific context-independent mechanism for it, we could incorporate the Pragmatic Reasoning Schemas (Cheng & Holyoak, 1985 Holyoak & Cheng, 1995) or possibly Social Exchange (Cosmides 1989). The model presented here attempts to combine neurobiological and psychological plausibility. In addition, experimental data confirmed a number of predictions of the model.

Acknowledgments

References

Ajjanagadde, V. (1994). Unclear Distinctions lead to Unnecessary Shortcomings: Examining the rule vs fact, role vs filler, and type vs predicate distinctions from a connectionist representation and reasoning perspective. *Proceedings of the AAAI-94.*

Cheng, P. W. & Holyoak, K. J. (1985). Pragmatic Reasoning Schemas. *Cognitive Psychology, 17,* 391-416.

Clark, H. H. (1969). Linguistic processes in deductive reasoning. *Psychological Review, 76,* 387-404.

Clossman, G. (1988). *A Model of Categorization and Learning in a Connectionist Broadcast System.* Unpublished Ph. D. thesis. Indiana University.

Cosmides, L. (1989). The Logic of Social Exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31,* 187-276.

De Soto, C. B., London, M. & Handel, S. (1965). Social reasoning and spatial paralogic. *Journal of Personality an Social Psychology, 2,* 513-521.

Evans, J. St B. T. (1989). *Bias in Human Reasoning: Causes and Consequences..* Hove: Lawrence-Erlbaum Ass.

French, R. M. (1994). Catastrophic Forgetting in Connectionist Networks, can it be prevented? In J. Cowan, G. Tesaro & J. Alspector (Eds.) *Advances in Neuro-Processing Systems 6.* San Francisco Ca.: Morgan-Kaufmann.

Gray, C. M., König, A., Engel, A. K. & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar`synchronization which reflects global stimulus properties. *Nature, 338,* 334-337.

Holyoak, K. J. & Cheng, P. W. (1995). Pragmatic Reasoning With a Point of View. *Thinking and Reasoning, 1,* 289-313.

Johnson-Laird, P. N., & Byrne, M. J. (1991). *Deduction.* London: Lawrence-Erlbaum Ass.

König, P., Engel, A. K., Roelfsema, P. R., & Singer, W. (1995). How Precise is Neuronal Synchronization? *Neural Computation, 7,* 469-485.

Lisman, J. E. & Idiart, M. A. P. (1995). Storage of 7 ± 2 Short-Term Memories in Oscillatory Subcycles. *Science, 267,* 1512-1515.

Mani, D. R. & Shastri, L. (1993). Reflexive Reasoning with Multiple Instanciation in a Connectionist Reasoning System with a Type Hierarchy. *Connection Science, 5,* 205-242.

Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two. *Psychological Review, 63,* 81-97.

Nelson, J. I. (1995). Binding in the Visual System. In M. A. Arbib (Ed.) *The Handbook of Brain Theory and Neural Networks.* Cambridge Ma.: MIT Press.

Oakford, M. & Chater, N. (1995). Theories of Reasoning and the Computational Explanation of Everyday Inference. *Thinking and Reasoning,* 1, 121-152.

Shastri, L. (1993). A Computational Model of Tractable Reasoning: taking inspiration from cognition. *Proceedings of the Thirteen International Joint Conference on Artificial Intelligence.*

Shastri, L. & Ajjanagadde, V. (1993). From Simple Associations to Systematic Reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences, 16,* 417-494.

Shastri, L. & Grannes, D. J. (1995). *Dealing with negated knowledge and inconsistency in a neurally motivated model of memory and reflexive reasoning.* International Computer Science Institute TR-95-041.

Singer, W. (1995). Synchronization of Neuronal Response as a Putative Binding Mechanism. In M. A. Arbib (Ed.) *The Handbook of Brain Theory and Neural Networks.* Cambridge Ma.: MIT Press.