Inference in Connectionist Networks

Lokendra Shastri

We are capable of drawing a variety of inferences effortlessly, spontaneously, and with remarkable efficiency — as though these inferences are a reflex response of our cognitive apparatus. This remarkable human ability poses a challenge for cognitive science and computational neuroscience: How can a network of slow neuron-like elements represent a large body of systematic knowledge and perform a wide range of inferences with such speed? We briefly review work in connectionist modeling that attempts to address this challenge by demonstrating how a neurally plausible network can encode a large body of semantic and episodic facts, systematic rules, and knowledge about entities and types, and yet perform a wide range of explanatory and predictive inferences within a few hundred milliseconds.

Keywords: reasoning, relational knowledge, connectionist networks, neural networks, semantic memory, temporal synchrony, binding problem

1. Introduction

Understanding written language is a complex task. It involves, among other things, recognizing words, accessing lexical items, disambiguating word senses, parsing, and carrying out inferences to establish referential and causal coherence, generate expectations, make predictions, and recognize the speaker's intent. Nevertheless we can understand language at the rate of several hundred words per minute. This rapid rate of language understanding suggests that we are capable of performing a wide range of inferences rapidly, spontaneously, and without conscious effort — as though they are a reflex response of our cognitive apparatus. In view of this, such reasoning may be described as reflexive reasoning (Shastri and Ajjanagadde, 1993).

Consider the following simple narrative: “John fell in the hallway. Tom had cleaned it. He got hurt.” Upon reading this narrative most of us would infer that Tom had cleaned the hallway, John fell because he slipped on the wet hallway floor, and John got hurt because of the fall. These inferences allow us to establish causal and referential coherence among the events and entities involved in the narrative. They help us explain John’s fall by making plausible inferences that the hallway floor was wet as a result of the cleaning, and John fell because he slipped on the wet floor. They help us causally link John’s hurt to his fall. They help us determine that the “it” in the second sentence refers to the hallway, and the “He” in the third sentence refers to John and not to Tom.

This remarkable human ability poses a challenge for cognitive science and computational
neuroscience: How can a system of simple and slow neuron-like elements represent a large body of systematic knowledge and perform a wide range of inferences with such speed?

Any system that attempts to explain our reflexive reasoning ability must possess a number of properties: First, such a system must be representationally adequate. It must be capable of encoding specific facts and events (e.g., John slipped in the hallway) as well as statistical summaries (e.g., It often rains in Seattle) and general rule-like regularities that capture the causal structure of the environment (e.g., If one falls down, one gets hurt). An important feature of such causal knowledge is that it is context-dependent and evidential in nature. Moreover, it is not tied to specific entities; rather, it applies to a range of entities of certain types. In formal terms, the expression of such general knowledge involves variables (see below). Second, the system should be inferentially adequate, that is, it should be capable of drawing inferences by combining evidence and arriving at coherent interpretations of perceptual or linguistic input. In order to do so, the system should be capable of establishing referential coherence. In particular, it should be able to posit the existence of appropriate entities, and it should be able to unify entities and events by recognizing that multiple designations refer to the same entity or event. Third, the system should be capable of learning and fine-tuning its causal model based on experience, instruction, and exploration. Finally, the system should be scalable and computationally effective; the causal model underlying human language understanding is extremely large, and hence, a system for establishing causal and referential coherence should be capable of encoding a large causal model and performing the requisite inferences within fractions of a second.

1.1 Relational knowledge and the binding problem

Since connectionist models compute by spreading activation, they are well-suited for performing evidential and parallel computations (Shastri 1989). But representing events and situations and reasoning about them require the representation of relational structures and a solution to the dynamic binding problem (Feldman, 1982; Smolensky 1990; Shastri and Aijanagadde, 1993).

Consider the representation of the event (E1) "John gave Mary a book in the hallway." This event is an instance of a specific sort of interaction involving John, Mary and a book that occurs in a particular location (the hallway). John and Mary are performing specific "roles" in this interaction; John is the one who is doing the giving, and Mary is the one who is doing the receiving. Moreover, a book is the object being given by John to Mary. Consequently, this event cannot be represented by simply activating the conceptual roles giver, recipient, and given-object, and the entities John, Mary, and book. Such a representation would be identical to that of an event wherein Mary gave John a book in the hallway.

An unambiguous representation of E1 requires the representation of bindings between the roles of E1 (e.g., giver) and the entities that fill these roles in the event (e.g., John). Note that two levels of bindings are involved: (1) entities occurring in the event are bound to the respective roles they fill in the event, and (2) all of the role-entity bindings pertaining to the event are grouped together in order to distinguish them from role-entity bindings pertaining to other events.

In order to reason about events and situations, connectionist models must also be capable of propagating dynamic bindings. Consider a network that encodes the systematic rule-like knowledge "If someone gives a recipient an object then the recipient comes to own that object." If such a network's state of activity is initialized to express the event "John gave Mary a book," its state of activity should evolve rapidly to include the dynamic representation of the inferred fact "Mary owns a book." Generating such inferred facts
involves the systematic propagation of dynamic bindings in accordance with the rule-like knowledge encoded in the network. A rule specifies antecedent and consequent relational structures and a correspondence between the roles of these relational structures. For example, the rule-like knowledge “If someone gives a recipient an object then the recipient comes to own that object” specifies that a _give_ event results in an _own_ event wherein the _recipient_ of the _give_ event corresponds to the _owner_ of the _own_ event and the _give-object_ of the _give_ event corresponds to the _own-object_ of the _own_ event. A step of inference, therefore amounts to taking an instance(s) of the antecedent relational structure(s) and creating — dynamically — an instance(s) of the consequent relational structure(s), with the role bindings of the latter being determined by (i) the role bindings of the former and (ii) the correspondence between role bindings specified in the rule. Once the representation of an inferred fact is established, it may be used in conjunction with other rule-like knowledge to create other inferred facts. Such a chain of inference may lead to a proliferation of inferred facts via the propagation of bindings across relational structures.\(^2\)

### 1.2 Connectionist solutions to the binding problem

It is straightforward to represent role-entity bindings within connectionist networks using additional nodes and links (e.g., using binder nodes that encode conjunctions of roles and entities). But while it is feasible to use additional nodes and links to encode persistent long-term knowledge, it is implausible to assume that binder nodes can be recruited for representing large numbers of dynamic bindings arising rapidly during language understanding and visual processing. In standard computing within the von Neumann architecture, bindings are expressed using variables and pointers, but these techniques have no direct analogues in connectionist networks.

Connectionist modelers have devised several solutions to the binding problem. One of the earliest solutions is due to Feldman (1982) who showed how any element of a group of \(N\) entities could be dynamically associated with any element of another group of \(N\) entities using an interconnection network. Several researchers have proposed formal as well as computational models for representing bindings using various types of convolution, tensor product, matrix multiplication, and XOR operations (e.g., Smolensky, 1990; Plate, 1994). Touretzky and Hinton (1988) proposed a solution to the binding problem in the context of a distributed connectionist production system (DCPS). The above models, however, do not address how multiple bindings can be expressed and propagated simultaneously in order to support inference. For example, DCPS can only deal with rules involving a single role and its reasoning process can only apply a single rule at a time. Thus such models are not adequate for modeling reflexive reasoning (Shastri and Ajjanagadde, 1993).

Another solution to the dynamic binding problem (Barnden and Srinivas, 1991) makes use of the relative position of active nodes and the similarity of their firing patterns to encode bindings. A solution proposed by Lange and Dyer (1989) and Sun (1992) assigns a distinct activation pattern (a signature) to each entity and propagates these signatures to establish role-entity bindings.
tri, 1989; Shastri and Ajjanagadde, 1993) proposed a biologically plausible solution for expressing and propagating role-entity bindings; they suggested that a role-entity binding such as (John=giver) be expressed by the synchronous firing of the role and entity nodes (also see Park, Robertson and Stenning, 1995; Sougne, 1996; Hummel and Holyoak, 1997). While the use of synchrony for binding perceptual features during visual processing had been suggested earlier by von der Malsburg (1981), the SHRUTI model of Shastri and Ajjanagadde (1993) discussed below offered a detailed account of how synchronous activity can be harnessed to represent complex relational knowledge and carry out rapid inference with respect to such knowledge. The advantages of using temporal synchrony to solve the dynamic binding problem over other solutions to this problem are discussed in (Shastri and Ajjanagadde 1993; Shastri 1996).

2. The SHRUTI architecture

SHRUTI is a fully implemented, neurally plausible model that demonstrates how a network of simple nodes and links can encode a large body of semantic and episodic facts, systematic rules, and knowledge about entities and types, and yet perform a wide range of explanatory and predictive inferences within a few hundred milliseconds (Shastri and Ajjanagadde, 1993; Shastri, 1999; 2000; Mani and Shastri 1993; Shastri and Wendelken, 2000; Wendelken and Shastri 2000; 2002).

SHRUTI can represent and process probabilistic, relational knowledge and deal with incomplete and inconsistent beliefs (Shastri, 1999, Shastri and Grannes, 1996). At the same time, SHRUTI’s representational machinery can encode parameterized action schemas, operators, and reactive plans (Shastri, Grannes, Narayanan, and Feldman, 2002). These process-based representations can encode compositional and hierarchically organized actions and support control behaviors such as conditional execution, interruption, consumption and termination, partial ordering, concurrency, and iteration.

SHRUTI propagates both beliefs and utilities over a single underlying “causal structure” to make predictions, seek explanations, and identify actions that increase expected future utility. The computations performed by SHRUTI can also be seen as an approximation of probabilistic reasoning in belief-nets (Pearl, 1988) and of dynamic programming, wherein a multi-step approximation of the Bellman equation (Bellman, 1957) is computed using the causal domain model, prior probabilities, and expected future utilities associated with (partial) world states (Thompson and Cohen, 1999).

2.1 An Overview of SHRUTI’s structure

The encoding of relational information in SHRUTI is mediated by structured clusters of nodes, referred to as focal-clusters. Each relational schema (or frame) has an associated focal-cluster. Such a focal-cluster for the relation fall is depicted in Figure 1 as the dotted ellipse labeled FALL. Each label within this ellipse denotes a connectionist node that serves a specific function outlined below. As explained in (Shastri, 1999), each connectionist node within a focal-cluster is a computational abstraction and corresponds to an ensemble of neurons. In general, a relational focal-cluster consists of the following nodes:

- A role node for each role in the relational schema. For simplicity, Figure 1 assumes that fall has only two roles: patient and location. The synchronous firing of a role node (e.g., patient) with an entity node (e.g., John) encodes a dynamic role-filler binding (e.g., [fall:patient=John]). Such dynamic role-filler bindings pertaining to the roles of a given relational schema constitute the currently active instance of that relation.
- Positive (+) and negative (−) collector nodes whose activation levels signify the levels of belief and disbelief, respectively, in the currently active relational instance. This can
range from absolute belief (+ fully active; – inactive) to absolute disbelief (– fully active; + inactive); from ignorance (neither + nor – active) to contradiction (both + and – active); and various graded beliefs in between. The + and – collector nodes are mutually inhibitory (inhibitory links terminate in a filled circle).

- An enabler node (E) whose activation signifies a search for support or explanation for the currently active relational instance. This node’s level of activity signals the strength with which support for this relational instance is being sought.
- A pair of utility nodes, +$p$ and +$n$, associated with the positive collector, and a pair of utility nodes, –$p$ and –$n$, associated with the negative collector. The activation levels of +$p$ and –$p$ indicate the desirability (i.e., positive utility) of the occurrence and non-occurrence, respectively, of the currently active relational instance. Similarly, the activation levels of +$n$ and –$n$ indicate the undesirability (i.e., negative utility) of the occurrence and non-occurrence, respectively, of the currently active relational instance.

There exist links from a focal-cluster’s collector nodes to its enabler node. This link converts a dynamic assertion of a relational instance into a query about that assertion. Consequently, the system continually seeks an explanation for all active relational instances.

The focal-cluster associated with a relational schema acts as an anchor for encoding and attaching various kinds of knowledge about the relation. This includes motor and perceptual schemas associated with the relational schema,
causal connections between this relational schema and other relational schemas, lexical and naming information, and episodic and semantic facts involving this generic relation. Information pertaining to a relation converges on its focal-cluster, and this information can be accessed by fanning out from the focal-cluster. It has been argued in Shastri and Ajjanagadde (1993) that it may be essential to associate such a focal-cluster with each relational schema in order to process relational information without cross-talk and at speeds required by cognitive processing.

The focal-cluster for an entity, say John, consists of an enabler node, ?.John, and a collector node, +:John. Persistent information about various perceptual and semantic features of John, his relationship with other concepts, and the roles he fills in various events are encoded via links between the focal-cluster of John and appropriate circuits and focal-clusters representing sensory, perceptual, and semantic “knowledge” distributed across various neural structures and regions. The focal-cluster of a type, T, contains a pair of enabler nodes ?e:T and ?v:T, and a pair of collector nodes +e:T and +v:T. The activation levels of +:John, +v:T, and +e:T signify the degrees of belief that entity John, type T, and an instance of T, respectively, play a role in a currently active situation. The activation of an enabler node signifies a search for an explanation about the appropriate entity, type, or instance.

The active representation of the relational instance “a man fell in the hallway” (i.e., [fall:patient = a man], [location = hallway]) is encoded by the firing of the +:fall node together with the synchronous firing of the (role) node fall:patient and the (filler) node +e:Man and the synchronous firing of the (role) node fall:location and the (filler) node +v:.Hallway (see Figure 2).

Long-term facts are encoded as temporal pattern matching circuits, or ‘fact’ structures. All facts pertaining to a relational schema are attached to its focal-cluster. If the enabler of a focal-cluster is active, and if the dynamic role-filler bindings currently active in the focal-cluster match the (static) bindings encoded by the long-term fact, the fact structure becomes active and, in turn, activates the positive collector of the relation’s focal-cluster and re-activates the role-filler bindings of the matching fact. SHRUTI encodes four different types of facts. These are:

1. Episodic facts (E-facts) that record relational instances corresponding to specific events and situations (e.g., I saw John in the library on Tuesday);

2. Taxon facts (T-facts) that record statistical knowledge about a relation (e.g., Soccer moms own minivans) and can be viewed as generalizations defined over multiple E-facts. Both E-facts and T-facts respond to partial cues, but while E-facts distinguish between highly similar events, T-facts respond to similarities.

3. Reward facts (R-facts) that record rewards or punishments associated with specific events and situations. R-facts are analogous to E-facts. Activation of a positive (negative) reward fact indicates the real or imagined attainment of some reward (punishment).

4. Value facts (V-facts) that associate a reward or a punishment with a generic event or situation. V-facts are analogous to T-facts and hold a statistical summary of
past activity. Thus they can predict future reward or punishment.

The distinction between different types of facts is motivated by biological considerations. For example, converging evidence from neuropsychology and neuroimaging suggests that while E-facts are encoded in cortico-hippocampal circuits, T-facts are encoded in purely cortical circuits (Shastri, 2001; 2002a). A connectionist encoding of E-facts and T-facts is described in (Shastri, 1999) and that of R-facts and V-facts in (Wendelken and Shastri, 2002). An anatomically realistic and physiologically grounded model of one-shot learning that is fully compatible with SHRUTI, and that can rapidly transform a transient pattern-of-activity-based representation of an event into a persistent structure-based memory trace of the event is described in (Shastri 2001; 2002a). This model of one-shot learning based on synaptic long-term potentiation (Bliss and Collingridge, 1993; Shastri, 2002b) provides a biologically grounded explanation for the rapid memorization of E-facts and R-facts.

A rule is encoded in SHRUTI as follows (refer to Figure 3):

- By linking appropriate (+ or −) collectors of the antecedent focal-clusters to appropriate collectors of the consequent focal-clusters. These links facilitate predictive (forward) inferences.
- By linking enablers of the consequent focal-clusters to enablers of the antecedent focal-clusters. These links facilitate explanatory (backward) inferences.
- By linking the corresponding roles of the antecedent and consequent focal-clusters (in both directions). If roles a and b are linked, the firing of a induces synchronous firing in b, thereby propagating role-entity bindings between inferentially related focal-clusters.
- Links from the appropriate utility nodes of the consequent focal-clusters to the appro-
priate utility nodes of the antecedent focal-clusters. These links facilitate the propagation of utility.

Links between antecedent and consequent focal-clusters are realized through an intervening focal-cluster termed the mediator focal-cluster (shown as a parallelogram in Figure 3)\(^3\). Type restrictions and the instantiation of unbound variables are handled via connections between the rule mediator and the type hierarchy. These connections are not shown in Figure 3. The weights of links encoding a rule are determined by the probabilistic or evidential strengths of the associated rule. Additional connectionist machinery exists to enable evidence combination and explaining away. Details of the above mechanisms are described in (Shastri, 1999; Shastri and Wendelken, 2000; Wendelken and Shastri, 2000; 2002).

2.2 Inference in SHRUTI

With reference to the network shown in Figure 3, the dynamic assertion of the fact “John fell in the hallway” corresponds to the firing of the node +:fall together with the synchronous firing of the nodes fall:patient and +:John and the synchronous firing of the nodes fall:location and +:Hallway. Given the network connectivity, this pattern of activity rapidly evolves so that the role hurt:patient starts firing in synchrony with the role fall:patient (and hence, with +:John), and the role hurt:location starts firing in synchrony with the role fall:location (and hence, with +:Hallway). The resulting firing pattern represents not only the event “John fell in the hallway,” but also the inferred event “John got hurt in the hallway.” Thus SHRUTI infers that John got hurt in the hallway given that John fell in the hallway. Furthermore, as a result of activity propagation in the entity and type structure (see Figure 3), SHRUTI also draws inferences such as “a man fell in the hallway” and “a person fell in some location.”

Similarly, the pattern of activity encoding “John got hurt in the hallway” automatically evolves to a pattern of activity that also encodes “Did John fall in the hallway?” At this time, the T-fact encoding the prior probability of people falling down provides a path for activity to flow from +:fall to +:fall, leading to the inference that John may have fallen down. This, in turn, leads to the activation of +:hurt and sets up a reverberatory loop of activity involving the focal-clusters for fall and hurt signaling that falling down is a coherent explanation for John getting hurt. The strength of this reverberatory activity would depend on the link weights encoding the rule (if you fall you get hurt) and the T-fact F1 (people are likely to fall).

In addition to the propagation of beliefs described above, the activation of +:hurt activates the V-fact V1 which encodes the negative utility associated with getting hurt. Note that the output of V1 is linked to +:$n$, and this signifies that getting hurt has a negative utility. Next V1 activates +:$n$:hurt, and this in turn activates +:$n$:fall signaling that negative utility is associated with falling.

As illustrated by this simple example, the assertion of any relational instance automatically leads to predictive and explanatory inferences and the evaluation of utility via the propagation of rhythmic activity between connected focal-clusters. The activation levels of the collector and enabler nodes of a relation are the result of the activation incident on them from various focal-clusters and the evidence combination functions operational at these nodes.

SHRUTI combines predictive inferences with explanatory (or abductive) inferences, exhibits priming, instantiates new entities during inference (if John ran, there must be a path along which he ran), unifies multiple entities by merging their phases of firing (if John got hurt and if there is a man who got hurt, then the man is likely to be John), and allows multiple explanations to compete with one another to identify the

\(^3\) The inclusion of a mediator was motivated, in part, by discussions the author had with Jerry Hobbs.
most promising explanation (explaining away).

With all the above machinery in place, SHRUTI can rapidly draw inferences such as “Tom had cleaned the hallway,” “John fell because he slipped on the wet hallway floor,” and “John got hurt because of the fall” in response to the input “John fell in the hallway. Tom had cleaned it. He got hurt.” Moreover, SHRUTI draws these inferences rapidly; the time taken to draw an inference is simply \( t \cdot \alpha \), where \( t \) is the length of the chain of inference, and \( \alpha \) is the time required for connected nodes to synchronize. Assuming that (i) gamma band activity underlines the propagation of bindings, \( \alpha \) is around 25-50 milliseconds. As discussed in (Shastri and Ajjanagadde, 1993), the speed of reflexive reasoning in SHRUTI satisfies the demands of real-time language understanding.

The estimate of utility at some antecedent predicate is based on both the utility value of its consequent and the probability that it will be reached from the consequent. This structure has the effect that the activation of a particular goal - via activation of a utility node - automatically leads to the assertion of its potential causes as subgoals, via spreading activation along a causal chain. Belief in a relational instance that is related to an active goal via a causal chain leads to an internal reward or punishment (via the activation of an R-fact) or to the recognition that such a reward is likely (via activation of a V-fact).

It has been shown that with an appropriate assignment of link weights, inference in SHRUTI is close to the probabilistic norm (Wendelken and Shastri, 2000). Moreover, it has been shown that approximations of these probabilistic weights can be learned using a neurally plausible unsupervised learning mechanism, Causal Hebbian Learning (Wendelken and Shastri, 2000) wherein link weights are updated depending on whether the presynaptic cell fires before or after the postsynaptic cell.

### 2.3 Predictions

Dynamic bindings, and hence, active relational instances (i.e., active facts) are represented in SHRUTI as a rhythmic pattern of activity over nodes in the long-term memory (LTM) network. In functional terms, this transient state of activation holds information temporarily during an episode of reflexive reasoning and corresponds to the **working memory underlying reflexive reasoning** (WMRR). Note that WMRR is just the state of activity of the LTM network and not a separate buffer\(^4\).

The use of temporal synchrony for encoding dynamic bindings leads to the prediction that during reflexive reasoning, a large number of facts may be active at the same time and a large number of rules may fire simultaneously, provided the number of distinct entities occurring as role-fillers in these items remains small (ca. 7). This constraint is motivated by biological considerations - each entity participating in dynamic bindings occupies a distinct phase, and hence, the number of distinct entities that can occur as role-fillers in dynamic facts cannot exceed \( \pi_{max}/\omega \). Here \( \pi_{max} \) is the maximum delay between consecutive firings of synchronous cell-clusters (about 25 milliseconds assuming that gamma band activity underlies the encoding of dynamic bindings), and \( \omega \) equals the allowable jitter in synchrony (ca. \pm 2 milliseconds).

Most proposals characterizing the capacity of the working memory underlying cognitive processing have not made adequate attention to the structure of items in the working memory and their role in processing. For example, proposals such as Just and Carpenter (1992) characterize working memory capacity in terms of “total activation.” In contrast, the constraints on working memory capacity predicted by SHRUTI

---

\(^4\) The active facts represented in the WMRR during an episode of reflexive reasoning should not be confused with the small number of short-term facts an agent may overtly keep track of during reflexive processing and problem solving. WMRR should also not be equated with the short-term memory implicated in various memory span tasks (Baddeley, 1986).
depend not on total activation but rather on the maximum number of distinct entities that can participate in dynamic bindings simultaneously. This prediction is consistent with the notion that cognitive tasks performed without deliberate thought tend to involve only a small number of distinct entities at a time — though of course, these entities may occur in multiple situations and relationships.

SHRUTI’s demonstration that a large number of rules — even those containing variables — may fire in parallel as long as the number of distinct entities referenced by the active facts remains small (ca. 7) may be contrasted with Newell’s suggestion (1980) that while “productions” (i.e., rules) without variables can be executed in parallel, productions with variables may have to be executed in a serial fashion. Thus SHRUTI suggests that neurally plausible architectures can support a high degree of parallelism — even when dealing with complex knowledge involving variables.

2.4 Learning SHRUTI structures

The learning of facts and rule-like mappings involving different types of focal-clusters presents a formidable challenge. However, an analysis of the problem and the results of ongoing research suggest that it is feasible to acquire these focal-clusters, facts, and mappings using biologically plausible learning rules, if one assumes an appropriate form of prior network structure (Shastri, 2001; Shastri & Wendelken, 2003). An excellent example of this may be found in the idiosyncratic architecture and local circuitry of the hippocampal formation and the nature of cortico-hippocampal interactions, which have been shown to be ideally suited for supporting the rapid encoding of episodic facts (Shastri, 2001; 2002a). Computational modeling of genetically plausible developmental processes (Marcus, 2001) also suggests that it is plausible to assume the pre-existence of recurring patterns of circuitry required for learning structures such as focal-clusters.

Specifically, ongoing work suggests that, given the appropriate prior network structure, functional cell clusters and circuits required to encode knowledge in SHRUTI can be learned via recruitment learning (Feldman, 1982; Shastri, 2002b) and Causal Hebbian Learning (Wendelken & Shastri, 2000). Recruitment learning can be described as follows: Learning occurs within a partially structured network. Recruited (free) nodes are nodes that have acquired distinct functionality by virtue of their strong interconnections to other recruited and sensorimotor nodes. Unrecruited (free) nodes are connected via weak links to a large number of free, recruited, and sensorimotor nodes. Free nodes form a pool of nodes from which suitably connected nodes are recruited for representing new functional units. Recruitment learning can be grounded in long-term potentiation and, over time, it can transform a quasi-structured network into nodes and circuits with specific functionalities (Shastri, 2001; Shastri, 2002b). In Causal Hebbian Learning (CHL), synaptic strength (weight) updates depend on the relative timing of pre- and post-synaptic firing (cf. spike-timing dependent plasticity (Bi & Poo, 2001)). It has been shown that simple forms of causal relationships (rules) can be learned using CHL (Wendelken & Shastri, 2000).

3. Conclusion

Connectionist models provide a natural and computationally effective framework for encoding complex evidential interactions. Over the past decades, these networks have grown in their sophistication and representational power to encompass not only perceptual and associative phenomena, but also high-level cognitive activities such as reasoning, language understanding, and problem solving. A promising development has been the development of connectionist models that are explicitly guided by behavioural, biological, and computational constraints. Having resolved some difficult representational problems, the focus of the field is shifting toward
more complex problems such as the learning of relational schemas and context-dependent mappings between schemas (rules), the development of sophisticated representation of actions, events, and time, the modeling of deliberative decision-making and planning, and the study of structured adaptive networks grounded in perception and action.

Acknowledgments

The preparation of this article was partially funded by grants from Infineon Technologies, Qualcomm Inc., and ICSI general funds. Thanks to V. Ajjanagadde, C. Wendelken, D.R. Mani, B. Thompson, M. Cohen, Massimiliano Garagnani, and D.G. Grannes and for their contributions to the development of the SHRUTI model.

References


(Received 25 Dec. 2002)

(Accepted 13 Jan. 2003)
Lokendra Shastri

Lokendra Shastri is a Senior Research Scientist and Principal Investigator at the International Computer Science Institute (ICSI). Before moving to ICSI, he was on the faculty of the Computer and Information Science Department, University of Pennsylvania. Shastri received a Bachelor of Engineering degree in Electronics from the Birla Institute of Technology and Science, India in 1977, an M.S. in Computer Science from the Indian Institute of Technology, Madras in 1981 and a Ph.D. in Computer Science from the University of Rochester in 1986. He is interested in neurally motivated computational models of learning, memory, inference, and language understanding; the role of temporal synchrony in neural information processing; and spatio-temporal pattern recognition. Shastri is on the advisory board of Neural Computing Surveys and the editorial board of Connection Science. He has served as a consultant for the National Science Foundation and the United Nations Development Programme.