A Hybrid Symbolic/Connectionist Model for Noun Phrase Understanding

STEFAN WERMTER & WENDY G. LEHNERT

This paper describes a hybrid model which integrates symbolic and connectionist techniques for the analysis of noun phrases. Our model consists of three levels: (1) a distributed connectionist level, (2) a localist connectionist level, and (3) a symbolic level. While most current systems in natural language processing use techniques from only one of these three levels, our model takes advantage of the virtues of all three processing paradigms. The distributed connectionist level provides a learned semantic memory model. The localist connectionist level integrates semantic and syntactic constraints. The symbolic level is responsible for restricted syntactic analysis and concept extraction. We conclude that a hybrid model is potentially stronger than models that rely on only one processing paradigm.

KEYWORDS: Natural language processing, connectionism, hybrid models, parallel distributed processing, relaxation networks, backpropagation, connectionist/symbolic systems.

1. Introduction

In recent years there has been a growing interest in using connectionist techniques for natural language processing. While traditionally the analysis, representation, and generation of natural language were exclusively dominated by symbolic approaches, lately connectionist techniques have received increased attention because of their attractive properties including noise resistance, learning behavior, neural plausibility, associative retrieval, and knowledge integration.

There have been at least two main directions of work in connectionist artificial intelligence: implementation-oriented and task-oriented. Implementation-oriented connectionism tries to show how symbolic representations and computations can be implemented with connectionist techniques. Connectionist systems have been developed to implement semantic networks (Shastri, 1988), rule-based systems (Tour-

Stefan Wermter & Wendy G. Lehnert, Department of Computer and Information Science, University of Massachusetts at Amherst, Amherst, MA 01003, USA; Tel: (413) 545-3639. The first author is currently with the Fachbereich Informatik, University of Dortmund, 4600 Dortmund 50, West Germany. We would like to thank Claire Cardie, David Lewis, Ellen Riloff, and the anonymous reviewers for comments on earlier drafts of this paper. This research was supported by the Advanced Research Projects Agency of the Department of Defense, monitored by the Office of Naval Research under contract No. N00014-87-K-0238, the Office of Naval Research, under a University Research Initiative Grant, Contract No. N00014-86-K-0764 and NSF Presidential Young Investigators Award NSFIST-8351863.

etzky & Hinton, 1988; Shastri & Ajjanagadde, 1989), representation languages like KL-ONE (Derthick, 1988), hierarchies and tree-like structures (Hinton, 1988; Pollack, 1988). Other connectionist systems show how symbolic computations can be implemented, e.g. variable binding (Touretzky & Hinton, 1985), sequential processing (Jordan, 1986; Elman, 1988), and recursion (Pollack, 1989). This implementation-oriented research demonstrates that connectionist models can, at least to a certain extent, implement symbolic structures and computations.

Task-oriented connectionism tries to show how specific tasks can be modeled with connectionist techniques. Numerous tasks in natural language processing have been attacked in recent years, e.g. parsing (Fanty, 1985; Hanson & Kegl, 1987; Howells, 1988; Kwasny, 1988), word sense disambiguation (Cottrell & Small, 1983; Bookman, 1987), anaphor resolution (Allen, 1987), compound noun understanding (Wermter, 1989b), sentence generation (Gasser, 1988), script and concept understanding (Dolan & Dyer, 1988; Miikkulainen & Dyer, 1989), language acquisition (Rumelhart & McClelland, 1986), and role assignment (McClelland & Kawamoto, 1986; St John & McClelland, 1988). These approaches demonstrate that connectionist models are useful for certain restricted tasks in natural language processing.

Implementation-oriented and task-oriented connectionism both demonstrate several advantages and disadvantages of symbolic and connectionist processing techniques. Although purely connectionist systems (Waltz & Pollack, 1985; Sejnowski & Rosenberg, 1986; Hanson & Kegl, 1987) and purely symbolic systems (Charniak, 1983; Dyer, 1983; Riesbeck & Martin, 1986; Grosz *et al.*, 1987; Hirst, 1987) have both shown impressive results, it has become obvious that connectionist techniques and symbolic techniques exhibit complementary strengths (Dyer, 1988; Lehnert, 1988; Touretzky, 1988; Hendler, 1989). While symbolic processing has advantages in representing schemata, recursive structures, variable binding, inheritance hierarchies, and sequential control, connectionist processing has advantages in associative retrieval, noise resistance, knowledge integration, generalization, and learning. Because of these mutually complementary properties, hybrid symbolic/connectionist systems promise to be more powerful than systems operating within only one paradigm.

In this paper we present a hybrid model for understanding noun phrases. This model combines localist and distributed connectionist techniques with symbolic techniques. The model consists of three levels: (1) distributed connectionist networks are used to learn semantic relationships between nouns, (2) localist connectionist networks integrate semantic constraints and syntactic constraints, and (3) symbolic techniques provide a restricted syntactic analysis and concept extraction. In Section 2 we describe our domain and our noun representation, in Section 3 the distributed connectionist level, in Section 4 the localist connectionist level, and in Section 5 the symbolic level. We show how a hybrid model can be used to understand noun phrases from a scientific technical domain.

2. The Domain: Noun Phrases in Scientific and Technical Sublanguages

Noun phrases are the dominant source of information in scientific and technical sublanguages (Hirschman, 1986). Because noun phrases are so important, natural language processing systems in these domains need a powerful and flexible model for understanding noun phrases. To investigate such a model we chose noun phrases from the NPL (National Physics Laboratory) corpus (Sparck Jones & Van Rijsbergen, 1976) as our domain. The NPL corpus contains queries and titles of scientific articles from the physical sciences. For example:

- Effects of electromagnetic fields on turbulences in gases.
- Note on the cause of ionization in the F-region.
- Radio emission by plasma oscillations in nonuniform plasmas.
- Calculation of fields on plasma ions by collective coordinates.
- An iterative analogue computer for use with resistance network analogues.

Syntactic, semantic, contextual, and world knowledge are all necessary for understanding complex noun phrases containing multiple prepositional phrases. In the past, several techniques have been developed to describe the problem of attaching prepositional phrases to the correct constituents (Prepositional Phrase Attachment; for instance, Kimball, 1973; Frazier & Fodor, 1978; Ford *et al.*, 1982; Crain & Steedman, 1985; Wilks *et al.*, 1985; Dahlgren & McDowell, 1986; McClelland & Kawamoto, 1986; Schubert, 1986; Hirst, 1987; Lehnert, 1987). Our hybrid approach is different from these approaches because we integrated distributed connectionist networks, localist connectionist networks, and symbolic techniques for understanding noun phrases.

Now we describe the representation of nouns in our domain of the physical sciences. We represent a noun as a binary vector of 16 features. This feature representation was developed as follows. First, we used thesaurus knowledge (EJC, 1967; NASA, 1985) for classifying the nouns occurring in the noun phrases. We categorized each noun according to the most general term in the hierarchy that describes the noun. This step abstracted specific nouns like 'carbon resistor', 'noise fluctuation', and 'transistor' to more general terms like 'resistor', 'variation', and 'semiconductor device'. Then, we grouped these most general thesaurus terms into 16 classes which form the basis of our feature representation. These 16 features describe the basic meaning of a noun in our domain.

For example, the term 'carbon resistor' is dominated by 'resistor' at the most general level and 'transistor' is dominated by 'semiconductor device'. 'Resistor' and 'semiconductor device' belong to the class (and therefore have the feature) ELECTRIC OBJECT. Each noun can have multiple features. For instance, the noun 'acceleration' has the features CHANGING-EVENT and ENERGY. Table I shows all features along with examples taken from the corpus. Having described the domain encoding, we now turn to a description of our three-part model.

Table I. Semantic features of the nouns and examples

Semantic features	Examples
MEASURING-EVENT	Observation, investigation, research
CHANGING-EVENT	Amplification, acceleration, loss
SCIENTIFIC-FIELD	Mechanics, ferromagnetics
PROPERTY	Intensity, viscosity, temperature
MECHANISM	Experiment, technique, theorem
ELECTRIC-OBJECT	Transistor, resistor, amplifier
PHYSICAL-OBJECT	Earth, crystal, vehicle, room
RELATION	Cause, dependence, interaction
ORGANIZATION-FORM	Layer, level, stratification, F-region
GAS	Air, oxygen, atmosphere, nitrogen
SPATIAL-LOCATION	Antarctic, earth, range, region, source
TIME	June, day, time, history
ENERGY	Radiation, ray, light, sound, current
MATERIAL	Aluminum, water, carbon, vapor
ABSTRACT-REPRESENTATION	Note, data, equation, term, parameter
EMPTY	Cavity, vacuum

3. Learning Semantic Prepositional Relationships in Distributed Connectionist Networks

In this section we describe semantic prepositional relationships and examine how they can be learned. Based on our feature representation, backpropagation networks learn underlying regularities of prepositional relationships from a corpus of training noun phrases. The learned regularities are used for attaching prepositional phrases to appropriate constituents within new noun phrases.

3.1. Semantic Prepositional Relationships

Within noun phrases, nouns can be connected with prepositions, as in 'Symposium on hydrodynamics in ionosphere'. Understanding these noun phrases relies on understanding prepositional relationships. A *prepositional relationship* is the semantic relationship between the features of two nouns which are connected by a preposition. Prepositional relationships can be either plausible or inplausible. *Plausible* prepositional relationships are possible relationships, such as 'symposium on hydrodynamics'. *Implausible* prepositional relationships are relationships which are not reasonable. 'Symposium in ionosphere' is implausible because symposiums do not take place in the outer atmosphere.

Knowing about the *plausible* prepositional relationships 'symposium on hydrodynamics' and 'hydrodynamics in ionosphere' and knowing about the *implausible* prepositional relationship 'symposium in ionosphere', we must interpret the noun phrase 'symposium on hydrodynamics in ionosphere' so that the prepositional phrase 'in ionosphere' attaches to 'hydrodynamics', but not to 'symposium'. Since knowledge about the plausibility of the prepositional relationship between two nouns can help to rule out implausible interpretations of the whole noun phrase, we have trained backpropagation networks to learn the plausibility of prepositional relationships.

3.2. Learning Semantic Prepositional Relationships with Backpropagation Networks

We use backpropagation networks (Rumelhart *et al.*, 1986) to learn the plausibility of prepositional relationships within noun phrases. For each preposition there is one backpropagation network that determines the plausibility of the prepositional relationships (see Figure 1). One network consists of three layers of units. The input layer consists of 32 binary units (values 0 and 1) representing 16 features for each of the two nouns. The single real-valued output unit determines whether the prepositional relationship is plausible (value 1) or implausible (value 0). Twelve real-valued hidden units encode the mapping from the input units to the output units from a training set. All levels in the backpropagation network are fully connected. We need one training set of prepositional relationships for each preposition.

First we concentrated on the three prepositions 'in', 'of', and 'on'. We randomly extracted 50 noun phrases from our corpus which contained only these three prepositions, for instance:

- Note on the cause of ionization in the F-region.
- International symposium on fluid mechanics in the ionosphere.

Based on these 50 noun phrases, we built one training set for each preposition. Each training example in the training set consists of two feature vectors for the two nouns together with the binary plausibility value for the prepositional relationship between these nouns. The plausibility value is set to 1 if the prepositional relationship in the

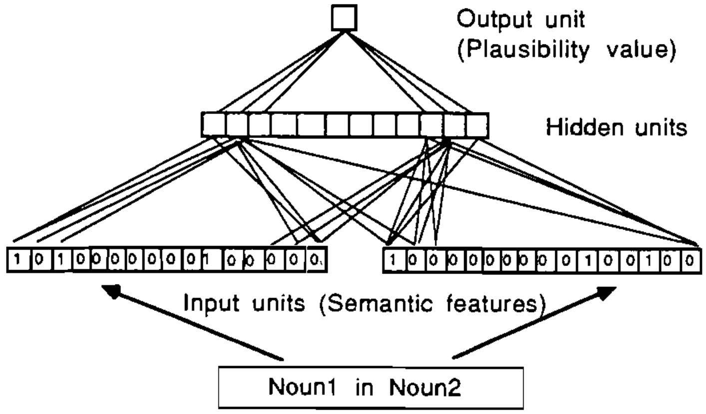

Figure 1. Backpropagation network for the prepositional relationships of 'in'.

training set is plausible and is set to 0 otherwise. From now on and where the context is clear, we will use the term prepositional relationship for both the semantic relationship between the two nouns and the representation of this semantic relationship as a training instance. Each noun in the 50 noun phrases is stored in a lexicon with its name and the associated 16 features. The following examples show two nouns with their features using the same feature order as in Table I.

F-region	(0 0 0 0 0 0 1 0 1 1 1 0 0 0 0 0)
Ionization	(0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0)

Now we will describe the training for the prepositional relationships for 'in'. There were 124 prepositional relationships for the preposition 'in' in the 50 noun phrases. Since most of these prepositional relationships in the 50 existing noun phrases are plausible prepositional relationships, most training examples would be plausible prepositional relationships.[1] We added the 124 inverse prepositional relationships to the 124 prepositional relationships so that the training set for 'in' consists of 248 prepositional relationships. An inverse prepositional relationship is a prepositional relationship in which the order of the two nouns is changed. Including inverse prepositional relationships in the training set prevents the network from being overloaded with too many plausible relationships since most of the inverse prepositional relationships are implausible. We illustrate the prepositional relationships and the inverse prepositional relationships for the preposition 'in' for our example 'Note on cause of ionization in F-region' together with their plausibility values in Table II.

3.2.1. Training results for the prepositional relationships for 'in'. Now we show the results for the training set with the 248 prepositional relationships for 'in'. We

Table II. Prepositional relationships for 'in' in the phrase 'Note on cause of ionization in F-region'

Prepositional relationships		Inverse prepositional relationships	
Note in F-region	0	F-region in note	0
Cause in F-region	1	F-region in cause	0
Ionization in F-region	1	F-region in ionization	0

conducted three runs training three backpropagation networks with the prepositional relationships for 'in'. The three different runs show that our training does not depend on a fortuitous initialization of the weights in the network. In each run the backpropagation network was trained for 1600 epochs (396 800 prepositional relationships) with the learning rate $\eta=0.01$ and weight change momentum $\alpha=0.9$. The weights in the backpropagation network were updated after each complete epoch.

After the training phase was completed, the trained networks were tested with the training set. To interpret the tests we introduce the terms 'error tolerance', 'error rate', and 'total error'. The *error tolerance* determines how much the actual outcome of the output unit could deviate from the desired outcome 0 for an implausible prepositional relationship and from the desired outcome 1 for a plausible prepositional relationship and still be considered correct. The *error rate* is the percentage of incorrectly classified prepositional relationships in the training set or in the test set. The *total error* is the total sum squared error on the complete training set as defined in Rumelhart *et al.* (1986, p. 323).

For the training set, the three networks of the three runs showed an error rate between 6.5% and 6.9% using an error tolerance of 0.49, and between 7.3% and 7.7% using an error tolerance of 0.3 (see Table III). A network which was not trained at all was tested with the training set and showed an error rate of 54.0% for the error tolerance 0.49, and an error rate of 73.4% for the error tolerance 0.3. These tests with the training examples demonstrate that an effective representation for prepositional relationships can be learned.

Table III. Test results for the training set for the prepositional relationships of 'in'

Run	1	2	3	No learning
Total error at the start of the training	77.8	62.5	70.1	—
Total error at the end of the training	6.6	7.1	8.1	—
Error rate for the training set for error tolerance 0.49	6.9	6.5	6.9	54.0
Error rate for the training set for error tolerance 0.30	7.7	7.3	7.7	73.4

After the networks had been tested with the 248 training examples, we tested the networks with 30 new test examples which were not part of the training set. For the test set we chose 15 plausible and 15 implausible prepositional relationships from our corpus with the only constraint that the prepositional relationships in the test set were not part of the training set. Examples from the test set are shown in Table IV.

Table IV. Examples of the test set for the prepositional relationships for 'in'

Plausible prepositional relationships	Implausible prepositional relationships
Effect in ferromagnetics	Japan in investigation
Distortion in amplifier	Power-supply in diode

The test results with 30 new prepositional relationships showed an error rate between 16.7% and 26.7% for the error tolerance 0.49 and between 20% and 30% for the error tolerance 0.3 (see Table V). The performance of the trained network on the new test examples can be demonstrated by comparing the described error rates with an untrained network. Training for 1600 epochs reduces the error rate for test examples which were not in the training set from 53.3% to 16.7% for an error tolerance of 0.49, and from 70.0% to 20.0% for an error tolerance of 0.3.

Table V. Test results for the test set for the prepositional relationships of 'in'

Run	1	2	3	No learning
Error rate for the test set for error tolerance 0.49	16.7	26.7	16.7	53.3
Error rate for the test set for error tolerance 0.30	20.0	30.0	20.0	70.0

To sum up the results for training the backpropagation networks with prepositional relationships for 'in', we have shown that for an error tolerance 0.49 trained networks can provide the plausibility value of a prepositional relationship correctly in about 93% of the prepositional relationships in the training set and in about 83% of the prepositional relationships in the test set.

3.3. Learned Internal Representations for the Prepositional Relationships for 'in'

After training had been completed we examined the internal representation in the backpropagation network. Figure 2 illustrates the activation values of the hidden units for 10 training examples. The first five rows show the hidden units for training examples with a plausible prepositional relationship, the last five rows show the hidden units for training examples with an implausible relationship.

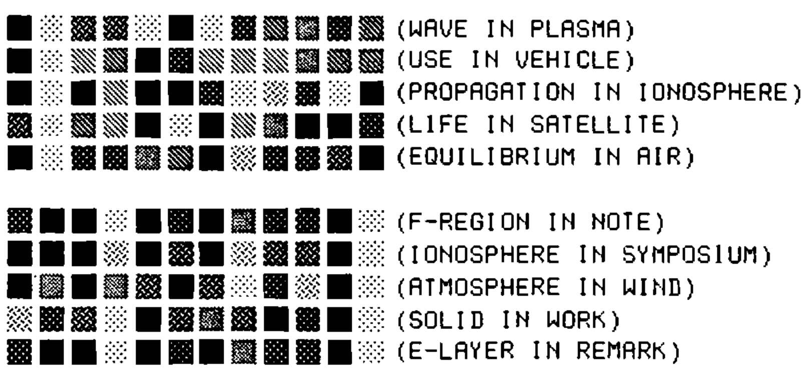

Figure 2. The hidden units for prepositional relationships from the training set.

Each row contains the 12 hidden units for one training instance. The hidden units have activation values between 0 (white) and 1 (black). Comparing the internal representations of the plausible prepositional relationships and the implausible prepositional relationships we found that plausible relationships correlate with a low value for hidden unit 2 and a high value for hidden unit 12. Implausible relationships correlate with a high value for hidden unit 2 and a low value for hidden unit 12. We do not claim that these two units are exclusively responsible for the distinction between plausible and implausible prepositional relationships. However, there is a strong tendency for these two units at least to play an important role in the internal representation of this distinction.

After looking at the hidden units with respect to the plausibility of a prepositional relationship we asked if the hidden units represent different word senses for plausible prepositional relationships. We used a simple clustering algorithm to cluster the vectors of 12 hidden units of the plausible prepositional relationships. This clustering algorithm takes a set of prototype vectors as its input and classifies all instances according to the minimal distance to the given prototype vectors. An instance is assigned to the class with the smallest distance to the prototype vector. This distance is computed as the sum of the squared differences between the feature vector of the current instance and the feature vector of the prototype. Although this simple clustering method relies on knowing 'good' prototype vectors, this method serves as a first approximation for a classification of the hidden units.

In Figure 3 we show examples of the internal representation for three clusters. The prototypes for the three clusters are the prepositional relationships 'effect in rectifier', 'radiation in atmosphere', and 'effect in beam'. These prototypes were chosen because they illustrate different interpretations of 'in': 'in a physical/electrical object', 'in a spatial location/gas', and 'in energy', respectively. In comparing the hidden units of these clusters, we found that units 6 and 9 essentially contain the information for differentiating these interpretations. In the first cluster, unit 6 has a low activation value and unit 9 has a high activation value; in the second cluster, unit 6 has a high activation value and unit 9 has a low activation value; and in the last cluster, both units 6 and 9 have low activation values. Although we found a few instances in the training set which use different units to distinguish between these clusters, most prepositional relationships in the three clusters can be differentiated solely based on the units 6 and 9.

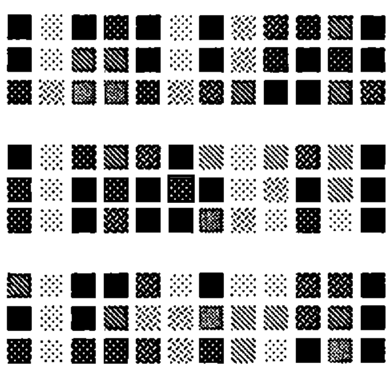

Figure 3. The hidden units for plausible prepositional relationships from three clusters.

To sum up, we have shown that specific hidden units in the distributed internal representation of the learned prepositional relationships are involved in encoding the plausibility of a relationship and in encoding specific interpretations for the prepositional relationship 'in'. Backpropagation networks which are trained for 1600 epochs can learn effective distributed representations of prepositional relationships. We demonstrated that these network representations currently reach a performance of about 93% (error rate about 7%) on the training set of prepositional relationships and about 83% (error rate about 17%) on the test set of prepositional relationships. Although we describe in detail only the training results for the prepositional relationships for 'in', we showed elsewhere (Wermter, 1989b) that other prepositional relationships behave very similarly. Experiments with semantic relationships for seven prepositions (by, for, from, in, of, on, with) demonstrate that the results for the prepositional relationships of 'in' hold for other prepositional relationships as well.

4. Integration of Semantic Relationships with Syntactic Constraints in Localist Connectionist Networks

While the previous section focused on learning semantic prepositional relationships with backpropagation networks, we now turn to a description of the localist network level. First, we briefly describe some syntactic constraints in noun phrases. Then, we show how simple syntactic constraints and learned semantic constraints can be integrated in a localist connectionist network for disambiguating noun phrases.

4.1. Syntactic Constraints

The two syntactic constraints we consider are the locality constraint and the no-crossing constraint. The *locality constraint* says that a prepositional phrase is more likely to attach to a close preceding noun than to a distant preceding noun. For instance, in the noun phrase 'Techniques for measurements in discharges' the prepositional phrase 'in discharges' might attach to 'measurements' or to 'techniques'. The locality constraint suggests that 'in discharges' attaches to 'measurements' because 'measurements' is closer than 'techniques'.

The *no-crossing constrant* (Tait, 1983) for noun phrases means that branches for attachment do not cross. The following (constructed) example shows a violated no-crossing constraint:

Influence of the temperature on the electrons in Fahrenheit

4.2. Localist Connectionist Networks for the Integration of Multiple Constraints

Localist networks have been used for a number of tasks to integrate multiple constraints in natural language processing, for instance for sentence understanding (Waltz & Pollack, 1985; Lehnert, 1987, 1988), for word sense disambiguation (Bookman, 1987), and for lexical access (Cottrell, 1988). It has been demonstrated elsewhere that localist networks are useful for integrating semantic and syntactic constraints for noun phrase disambiguation (Wermter, 1989a). In this section we describe the most important properties of an efficient localist network that performs noun phrase disambiguation with fewer nodes (see Figure 4).

Our localist network consists of three types of nodes; *noun nodes* represent the nouns in a noun phrase, *semantic nodes* represent the plausibility of prepositional relationships between nouns, and *locality nodes* represent the distance between two nouns in a noun phrase. Each node has an activation potential between 0 and 10. A semantic node in the localist network is initialized with the plausibility value of the output unit of the appropriate backpropagation network (multiplied by a factor of 10

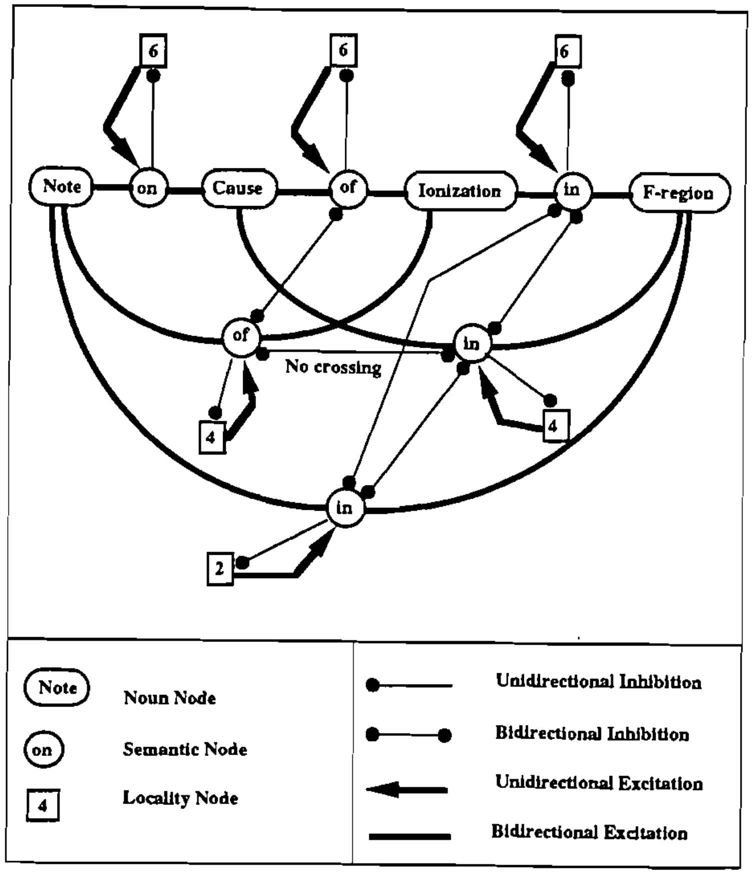

Figure 4. Localist network for the integration of multiple constraints.

to get values between 0 and 10). The higher the plausibility value of a prepositional relationship, the higher the initialization value for the semantic node. The initialization of the locality nodes is based on the relative distance between the nouns. The closer two nouns are in a noun phrase, the higher the initialization value for the locality node between these nouns. It is important to point out that the initialization of the locality nodes is fairly independent of specific values (e.g. 6, 4, and 2 in Figure 4). Other initialization values (e.g. 3, 2, 1 or 8, 4, 2) work as well as long as there is a decreasing relationship for the distance between the nouns, and as long as all the values are not too close to the upper and lower bounds of the nodes. Noun nodes are initialized with 0 activation since they serve only as the framework to which semantic nodes and locality nodes connect. The semantic constraints are encoded as the semantic nodes, the locality constraints as the locality nodes, and the no-crossing constraints as specific inhibitory connections between semantic nodes in crossing attachment links (see Figure 4).

All nodes are connected via inhibitory and excitatory connections, as Figure 4 shows for a network with three prepositions. Each noun node has excitatory attachment links to each noun node of preceding nouns. The semantic nodes in competing attachment links are inhibitorily connected. The locality nodes provide excitation to semantic nodes depending on the distance of the attachment. The inhibitory connection from a semantic node to a locality node prevents the locality node from sending too much excitation to the semantic node. Networks for noun phrases with a different number of prepositions are built in exactly the same systematic manner. The input nodes are initialized and activation spreads through the network according to a standard relaxation algorithm (Feldman & Ballard, 1982). After about 20 to 30 cycles the localist network settles in a global interpretation, and the semantic nodes with the highest activation values determine the preferred structural interpretation. Examples of this process and for the interaction of the localist network with the symbolic structures are given in the following sections.

5. Symbolic Level

The previous two sections explained the connectionist networks for learning semantic relationships and for integrating semantic and syntactic constraints. In this section we describe the symbolic level of our model for noun phrase understanding.

In the last two sections we have assumed that the noun phrases for the connectionist networks only consist of nouns and prepositions. However, as our examples in Section 2 showed, noun phrases often contain other parts of speech as well, including adjectives, adverbs, and determiners. Although these parts of speech might contain significant information they are usually less important for the representation of the *essential* concept of a noun phrase and the structural disambiguation of the noun phrase. Therefore, the purpose of the symbolic level is to extract the essential sequence of nouns and prepositions from the complete noun phrase. Then, this essential reduced noun phrase has the canonical form of nouns and prepositions required for our connectionist levels.

The first mechanism is an analysis of the noun phrase with respect to its syntactic constituents. This restricted syntactic analysis is provided by a subsystem of CIRCUS (Lehnert, 1988), which uses a stack-based architecture to recognize simple syntactic constituents. Since we want to extract the essential reduced noun phrase from the complete noun phrase and since we do not need complete parse trees for this extraction, the restricted syntactic analysis in CIRCUS is sufficient for our purpose.

This subsystem uses a syntactic dictionary and syntactic predictions for identifying constituents. The syntactic predictions are encoded as requests and the request packet mechanism of MCELI (Schank & Riesbeck, 1981) is used to process the predicted next constituents. Before we begin to analyze a noun phrase, an initial syntactic prediction for the head noun will be on top of the stack. This prediction allows us to skip possible intervening constituents like adjectives, adverbs, and determiners and stores the head noun in a global buffer. At this point the current request is removed from the stack. If a preposition follows, then a new request is pushed on the stack for the following prepositional phrase. As soon as the next noun is identified it is stored in another global buffer for this prepositional phrase. This process of adding syntactic predictions, removing the predictions, finding the desired constituents (prepositions and nouns), and storing them in global buffers is continued until the noun phrase is completely processed.

Although it might seem that a simple pattern matching algorithm which identifies nouns and prepositions using a syntactic dictionary might be sufficient, such a simple approach does not account for more complicated noun phrases with associated subclauses or participle constructions. For instance, for the noun phrase 'the man in the satellite which blinked in the sun' we only want to extract 'man in satellite' and not 'man in satellite in sun', which would be constructed in a simple pattern matching approach based solely on parts of speech. Using a syntactic prediction for a new subclause associated with the relative pronoun 'which', it is possible to detect and skip this subclause so that only syntactically desired constituents are extracted.

While the restricted syntactic analysis transforms a noun phrase into a reduced noun phrase based on syntactic predictions, a second mechanism can extract the essential part of this reduced noun phrase based on semantic predictions. The semantic predictions are associated with words in the semantic dictionary. Semantic predictions are fulfilled if the current part of speech in the noun phrase is considered essential. The question of what is considered essential depends on the application and the domain. For instance, in an information retrieval context we might have queries like:

- Information on papers about turbulences in gas.

In this domain it is not wise to include the nouns 'information' and 'papers' in a concept representation since they do not contribute any important *distinguishing* information. For this information retrieval task only the nouns that are important for the domain can fulfill the semantic predictions. In our example these nouns are 'turbulence' and 'gas', but not 'information' and 'papers'. Therefore 'turbulence in gas' is extracted as the essential part of the noun phrase in this application. In general, these semantic predictions are fulfilled if the current constituent is an essential part of the noun phrase or if a preceding constituent was identified as an essential part. Both semantic and syntactic predictions allow us to extract the essential part of a noun phrase. Specific examples of this level are given in the next section.

6. Operation of the System

In this section we describe the operation of our whole system and show some examples of its performance. First, we focus on how a typical noun phrase is processed in detail:

- Note on a new cause of increasing ionization in the F-region.

The *symbolic level* extracts the sequence of nouns and prepositions from this noun phrase and provides the following noun phrase:

- Note on cause of ionization in F-region

This symbolic level could also skip relative clauses as in 'ionization in the F-region which is close to the Antarctic' or participle constructions as in 'ionization in the F-region surrounding the Antarctic'. Then all possible prepositional relationships for the reduced noun phrase are computed:

- Note on cause
- Cause of ionization
- Note of ionization
- Ionization in F-region
- Cause in F-region
- Note in F-region

The feature representation of each noun in the prepositional relationships is looked up in the lexicon. Based on these features, backpropagation networks at the *distributed level* are initialized for each prepositional relationship. The output of the backpropagation networks are the plausibility values in the network at the *localist level.* Locality nodes and noun nodes are initialized as well. Then the localist network starts processing, integrates the syntactic and semantic constraints, and stabilizes in a global interpretation of the noun phrase. The activation of the semantic nodes in the localist network determines the preferred structural interpretation of the noun phrase. In our example there are three semantic nodes that have high activation values after the relaxation. These nodes correspond to the following interpretation:

Note on cause of ionization in F-region

In the following we show more examples of noun phrases and their structural interpretation:

(1) Effect of field on turbulence in gas hair →
CONCEPT: effect
OF-REL: field
ON-REL: turbulence
IN-REL: gas

(2) Dependence of amplification in phosphor on intensity →
CONCEPT: dependence
OF-REL: amplification
IN-REL: phosphor
ON-REL: intensity

(3) Distortion in amplifier on satellite in Van-Allen-belt →
CONCEPT: distortion
IN-REL: amplifier
ON-REL: satellite
IN-REL: Van-Allen-belt

(4) Experiment on diffraction of ray in layer →
CONCEPT: experiment
ON-REL: diffraction
ON-REL: ray
IN-REL: layer

The last example (4) shows that not *all* attachments are necessarily wrong for an interpretation to be considered incorrect. The first two attachments are correct but 'in layer' should attach to 'diffraction' rather than to 'ray'. Nevertheless, we consider structural interpretations with at least one wrong attachment to be incorrect. Using this strict and conservative evaluation we tested our system with 80 noun phrases containing up to three prepositions. A correct structural interpretation was assigned for 88% of 50 noun phrases that contained prepositional relationships from the training set and for 77% of 30 noun phrases that contained prepositional relationships which were not in the training set. Prepositional phrases can attach to several nouns if only semantic constraints are considered. The overall strategy is to prefer semantic constraints over syntactic constraints (locality and no-crossing of branches) and to use syntactic constraints to favor one of several possible semantic interpretations.

7. Discussion

In this section we first compare our hybrid model with other symbolic models for structural noun phrase disambiguation. Then we focus on the single levels of our model and explain why we chose a hybrid three-level model.

Recently there has been a lot of interest in attacking the problem of structural ambiguity, especially in prepositional phrase attachment (Wilks *et al.*, 1985; Schubert, 1986; Dahlgren & McDowell, 1986; Jensen & Binot, 1987; St John & McClelland, 1988). All these approaches focus on attaching a *single* prepositional phrase within a sentence of the form ⟨NP⟩ ⟨VP⟩ ⟨NP⟩ ⟨PP⟩. Our approach focuses on attaching *multiple* prepositional phrases within noun phrases. Attaching multiple prepositional phrases in noun phrases is a much harder problem since we cannot rely on predictive verbal knowledge alone.

Most previous work on prepositional phrase attachment relies on an intuitive development of symbolic heuristic rules (Wilks *et al.*, 1985; Schubert, 1986; Dahlgren & McDowell, 1986; Hirst, 1987). Since prepositional phrase attachment cannot reasonably be attacked without semantic knowledge, these rules have to encode the semantic knowledge and have to be redesigned for new domains. Our model tackles this problem by learning and generalizing over semantic constraints and eliminating knowledge which has to be handcoded.

Another approach for reducing the amount of knowledge engineering can be found in Jensen & Binot (1987). This approach attacks the problem of acquiring semantic knowledge for attachments by using definitions in an on-line dictionary. Although this symbolic approach was shown to attach correctly a single prepositional phrase in some sentences, this method depends on suitable definitions in the lexicon. While using on-line dictionaries is a very reasonable attempt, it appears that much more work is required in standardizing semantic knowledge in on-line dictionaries before we can use them to support disambiguation in a general manner.

Recent work on symbolic prepositional phrase attachment (Dahlgren, 1988) reports a success rate above 93% for the attachment of *single* prepositional phrases. These results were obtained by hand-testing intuitively developed rules on several small corpora. Our approach reaches 88% on the training set and 77% on the test set of new noun phrases. Although our results might be even better with further training we believe that our current results already demonstrate the effectiveness of our approach for two reasons. First, multiple prepositional phrase attachment is a much harder problem than Dahlgren's single prepositional phrase attachment. In our experiments we considered noun phrases with up to three prepositional phrases. Second, our model did not rely on intuitively developed rules but learned part of its knowledge. In general, we believe that our hybrid model has a lot of potential compared with traditional, purely symbolic methods since our hybrid model attacks a much harder problem, acquires part of its knowledge by learning, and already comes close to the best performance of purely symbolic approaches that attack a significantly simpler problem.

We now turn to the discussion of the three levels in our hybrid model and give reasons for the design of each individual level. The *symbolic level* performs a restricted syntactic analysis and extracts the essential concept (the sequence of nouns and prepositions) of a noun phrase for the attachment decision. A symbolic approach is more suitable for this level since the extraction of the essential concept based on syntactic and semantic predictions is a sequential control problem—it has to be decided which constituents to process. In a symbolic mechanism, syntactic and semantic predictions for this extraction can be formulated easily. In contrast, in a connectionist framework localist networks would have to be designed or distributed networks trained to perform the extraction. Although there has been some success using recurrent networks for processing restricted sequential structures (Elman, 1988; Pollack, 1988; St John & McClelland, 1988), these recurrent networks do not seem to be powerful enough for this complex sequential control problem. Since symbolic techniques are particularly suitable for dealing with sequentiality and control, they are more useful for dealing with the variety of constituents as they occur in noun phrases in real-world examples.

The *localist level* performs the integration of syntactic and semantic constraints. The links in the localist network implement the possible attachments and their mutual competition. The localist model considered here is more efficient in the number of nodes than the localist model for attachment in Wermter (1989a). That model used one structure node for each possible structural interpretation of a noun phrase. While that architecture was useful for short noun phrases, the number of structure nodes increased exponentially with the length of the noun phrase. In our present model the total number of nodes in the network increases only quadratically with the length of the noun phrases.

A similar localist network for prepositional phrase attachment in noun phrases can be found in Touretzky (1989). While Touretzky uses a similar attachment architecture, he implements locality constraints only in a restricted way by reducing the specific threshold for the unit representing the 'nearest neighbor' noun. In our model we implemented locality constraints explicitly in a more general way with locality nodes for the relationship to every preceding noun. For the locality nodes, we found empirically that initialization values should decrease with the length of the attachment and they should be well under the upper threshold for the nodes.[2]

The initialization of the semantic nodes is based on the *distributed level*. The distributed backpropagation networks learn and generalize the semantic prepositional relationships and provide a semantic memory model for the initialization of the localist network. Other work on learning relationships between constituents (Hinton, 1986; Cosic & Munro, 1988) cannot be directly used to provide this memory model. Hinton attacks a completion task that finds a specific relative given a person and a family relationship. Cosic & Munro tackle a completion task which determines the meaning of a preposition based on the lexical item of the preposition and two nouns. Although this work deals with learning relationships, these architectures cannot be directly used to support the initialization of single nodes in localist networks.

Furthermore, both architectures (Hinton, 1986; Cosic & Munro, 1988) have all constituents and all relationships encoded in one backpropagation network. While this might be sufficient for small applications, one huge network cannot be expected to be efficient in terms of training time and generalization behavior for scaling up to bigger applications. Therefore, we have one backpropagation network for each preposition. Apart from less training time and better generalization, this modular architecture also allows us the modification and addition of individual prepositions without retraining the whole network.

Another interesting design issue is the number of units in the backpropagation networks. The number of input units was determined by our choice of 16 features for representing each noun. There is one output unit for the plausibility value. More interesting is our choice of 12 for the number of hidden units. Increasing the number of 12 hidden units led to better performance on the training set, but worse performance on the test set. Decreasing the number of hidden units decreased the performance on the training set and test set. Apparently, there is a tradeoff between memorization and generalization and we found our best results with a hidden layer that had slightly less than half the number of input units.

8. Conclusion

We have described a hybrid symbolic/connectionist system for noun phrase disambiguation. The symbolic level supplies input for the connectionist networks by extracting the sequence of nouns and prepositions from a noun phrase. The localist connectionist network integrates semantic and syntactic constraints for noun phrase disambiguation and computes a preferred structural interpretation. Distributed connectionist networks learn semantic relationships between nouns, allow for generalizations of the learned relationships, and provide a semantic memory model for initializing nodes in the localist connectionist networks. This hybrid three-level model of distributed connectionist networks, localist connectionist networks, and symbolic concepts allows for the combination of learning and generalization, the integration of competing constraints, and the symbolic extraction of concepts and makes this hybrid model potentially stronger than models relying on techniques from only one of the three processing paradigms.

Notes

1. Implausible prepositional relationships like 'symposium in ionosphere' in the noun phrase 'symposium on hydrodynamics in ionosphere' occur less frequently in existing noun phrases than plausible prepositional relationships.
2. For example, for a noun phrase with three prepositions initialization values of 3, 2, 1 for the different attachments implement a small syntactic locality effect. The values 6, 4, 2 implement a moderate syntactic influence. If the values are too high, e.g. 10, 9, 8, the network is overloaded with excitation from the locality nodes.

References

Allen, R.B. (1987) Several studies on natural language and back-propagation. *Proceedings of the First Annual International Conference on Neural Networks.*

Bookman, L.A. (1987) A microfeature based scheme for modelling semantics. *Proceedings of the International Joint Conference on Artificial Intelligence.*

Charniak, E. (1983) A parser with something for everyone. In M. King (Ed.) *Parsing Natural Language.* London: Academic Press.

Cosic, C. & Munro, P. (1988) Learning to represent and understand locative prepositional phrases. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society.*

Cottrell, G.W. & Small, S.L. (1983) A connectionist scheme for modelling word sense disambiguation. *Cognition and Brain Theory*, **6**.

Cottrell, G.W. (1988) A model of lexical access of ambiguous words. In S. I. Small, G. W. Cottrell & M. K. Tanenhaus (Eds) *Lexical Ambiguity Resolution.* San Mateo, CA: Morgan Kaufmann.

Crain, S. & Steedman, M. (1985) On not being led up the garden path: the use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen & A. M. Zwicky (Eds) *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives.* Cambridge: Cambridge University Press.

Dahlgren, K. (1988) *Naive Semantics for Natural Language Understanding.* Dordrecht: Kluwer.

Dahlgren, K. & McDowell, J. (1986) Using commonsense knowledge to disambiguate prepositional phrase modifiers. *Proceedings of the National Conference on Artificial Intelligence.*

Derthick, M. (1988) Mundane reasoning by parallel constraint satisfaction, Technical Report CMU-CS-88-182. Carnegie-Mellon University.

Dolan, C.P. & Dyer, M.G. (1988) *Parallel retrieval and application of conceptual knowlege*, Technical Report UCLA-AI-88-3. Los Angeles, CA: University of California.

Dyer, M.G. (1983) *In-depth Understanding.* Cambridge, MA: MIT Press.

Dyer, M.G. (1988) Symbolic neuroengineering for natural language processing: a multilevel research approach. In J. Barnden & J. Pollack (Eds) *Advances in Connectionist and Neural Computation Theory.* Norwood, NJ: Ablex Publishers (in press) (also available as Technical Report UCLA-AI-88-14, University of California, Los Angeles).

EJC (1967) *Thesaurus of Engineering and Scientific Terms.* New York: Engineers Joint Council.

Elman, J.L. (1988) *Finding structure in time*, Technical Report CRL-8801. San Diego, CA: University of California.

Fanty, M. (1985) *Context-free parsing in connectionist networks*, Technical Report 174. Rochester, NY: University of Rochester.

Feldman, J.A. & Ballard, D.H. (1982) Connectionist models and their properties. *Cognitive Science*, **6**.

Ford, M., Bresnan, J.W. & Kaplan, R.M. (1982) A competence based theory of syntactic closure. In J. W. Bresnan (Ed.) *The Mental Representation of Grammatical Relations.*Cambridge, MA: MIT Press.

Frazier, L. & Fodor, J.D. (1978) The sausage machine: a new two-stage parsing model. *Cognition*, **6**.

Gasser, M.E. (1988) *A connectionist model of sentence generation in a first and second language*, Technical Report UCLA-AI-88-13. Los Angeles, CA: University of California.

Grosz, B.J., Appelt, D.E., Martin, P.A. & Pereira, F.C.N. (1987) TEAM: an experiment in the design of transportable natural-language interfaces. *Artificial Intelligence*, **32**.

Hanson, S.J. & Kegl, J. (1987) PARSNIP: a connectionist network that learns natural language grammar from exposure to natural language sentences. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*.

Hendler, J.A. (1989) Marker passing over microfeatures: towards a hybrid symbolic/connectionist model. *Cognitive Science*, **13**.

Hinton, G.E. (1986) Learning distributed representations of concepts. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*.

Hinton, G.E. (1988) Representing part-whole hierarchies in connectionist networks. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*.

Hirschman, L. (1986) Discovering sublanguage structures. In R. Grishman & R. Kittredge (Eds) *Analyzing Language in Restricted Domains: Sublanguage Description and Processing*.

Hirst, G. (1987) *Semantic Interpretation and the Resolution of Ambiguity*. Cambridge: Cambridge University Press.

Howells, T. (1988) VITAL—A connectionist parser. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*.

Jensen, K. & Binot, J.L. (1987) Disambiguating prepositional phrase attachments by using on-line dictionary definitions. *Computational Linguistics*, **13**.

Jordan, M.I. (1986) Attractor dynamics and parallelism in a connectionist sequential machine. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*.

Kimball, J. (1973) Seven principles of surface structure parsing in natural language. *Cognition*, **2**.

Kwasny, S.C. (1988) *A parallel distributed approach to parsing natural language deterministically*, Technical Report WUCS-88-21. St Louis, WA: Washington University.

Lehnert, W.G. (1987) Learning to integrate syntax and semantics. *Proceedings of the Fourth International Workshop on Machine Learning*.

Lehnert, W.G. (1988) Symbolic/subsymbolic sentence analysis: exploiting the best of two worlds. In J. Barnden & J. Pollack (Eds) *Advances in Connectionist and Neural Computation Theory*. Norwood, NJ: Ablex (in press) (also available as COINS Technical Report 88-99, Department of Computer and Information Science, University of Massachusetts, Amherst, MA).

NASA (1985) *NASA Thesaurus* (National Aeronautics and Space Administration).

McClelland, J.L. & Kawamoto, A.H. (1986) Mechanisms of sentence processing: assigning roles to constituents. In J. L. McClelland & D. E. Rumelhart (Eds) *Parallel Distributed Processing*, Vol. 2. Cambridge, MA: MIT Press.

Miikulainen, R. & Dyer, M.G. (1989) *A modular neural network architecture for sequential paraphrasing of script-based stories*, Technical Report UCLA-AI-89-02. Los Angeles, CA: University of California.

Pollack, J. (1988) Recursive auto-associative memory: devising compositional distributed representations. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*.

Pollack, J. (1989) Implications of recursive distributed representations. *Conference on Neural Information Processing Systems* (submitted).

Riesbeck, C. & Martin, C. (1986) Direct memory access parsing. In C. Riesbeck & J. Kolodner (Eds) *Experience, Memory and Reasoning*. Hillsdale, NJ: Lawrence Erlbaum.

Rumelhart, D.E., Hinton, G.E. & Williams, R.J. (1986) Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds) *Parallel Distributed Processing*, Vol. 1. Cambridge, MA: MIT Press.

Rumelhart, D.E. & McClelland, J.L. (1986) On learning the past tenses of English verbs. In J. L. McClelland & D. E. Rumelhart (Eds) *Parallel Distributed Processing*, Vol. 2. Cambridge, MA: MIT Press.

Schank, R.D. & Riesbeck, C. (1981) *Inside Computer Understanding*. Hillsdale, NJ: Lawrence Erlbaum.

Schubert, L.K. (1986) Are there preference tradeoffs in attachment decisions? *Proceedings of the National Conference on Artificial Intelligence*.

Sejnowski, T.J. & Rosenberg, C.R. (1986) *NETtalk: a parallel network that learns to read aloud*, Technical Report JHU/EECS-86/01. Baltimore, MD: Johns Hopkins University Press.

Shastri, L. (1988) A connectionist approach to knowledge representation and limited inference. *Cognitive Science*, **12**.

Shastri, L. & Ajjanagadde, V. (1989) *A connectionist system for rule based reasoning with multi-place predicates and variables*, Technical Report MS-CIS-89-06. University of Pennsylvania.

Sparck Jones, K. & Rijsbergen, C.J. (1976) Information retrieval test collections. *Journal of Documentation*, **32**.

St John, M.F. & McClelland, J.L. (1988) Applying contextual constraints in sentence comprehension. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*.

Tait, J.I. (1983) Semantic parsing and syntactic constraints (Mark IV). In K. Sparck Jones & Y. Wilks (Eds) *Automated natural language processing*. Chichester: Ellis Horwood.

Touretzky, D.S. (1988) Beyond associative memory: connectionists must search for other cognitive primitives. *Proceedings of the AAAI Spring Symposium Series: Parallel Models of Intelligence: How Can Slow Components Think so Fast?*

Touretzky, D.S. (1989) *Connectionism and compositional semantics*, Technical Report CMU-CS-89-147. Carnegie Mellon University.

Touretzky, D.S. & Hinton, G.E. (1985) Symbols among the neurons: details of a connectionist inference architecture. *Proceedings of the International Joint Conference on Artificial Intelligence*.

Touretzky, D.S. & Hinton, G.E. (1988) A distributed connectionist production system. *Cognitive Science*, **12**.

Waltz, D.L. & Pollack, J.B. (1985) Massively parallel parsing: a strongly interactive model of natural language interpretation. *Cognitive Science*, **9**.

Wermter, S. (1989a) Integration of semantic and syntactic constraints for structural noun phrase disambiguation. *Proceedings of the International Joint Conference on Artificial Intelligence*.

Wermter, S. (1989b) Learning semantic relationships in compound nouns with connectionist networks. *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*.

Wilks, Y., Huang, X. & Fass, D. (1985) Syntax, preference and right attachment. *Proceedings of the International Joint Conference on Artificial Intelligence*.