CONNECTIONISM AND LINGUISTICS

Timo Honkela

1. INTRODUCTION

Much of the formal and computational study of language has centered around syntax, to the detriment of semantics and pragmatics. The reason for this might be that the methods available have been more suitable to the study of syntax. It seems that so called connectionist models offer a promising method for dealing especially with semantics and pragmatics. The most advanced connectionist systems are artificial neural networks which have e.g. learning capabilities. This learning can be applied to linguistic material such as corpora.

In the following, connectionist methods are compared to more traditional symbolic methods. Within the connectionist paradigm there are a number of different approaches. Two of them - backpropagation and self-organizing maps - are presented. Some examples of connectionist linguistic models are given in section 3. The further possibilities of connectionist models are analyzed in section 4.

2. ON CONNECTIONISM AND ITS RELATION TO TRADITIONAL METHODS

Although there are interesting analogies between present-day computers and human brains (e.g. memory), it must be remembered that there are significant differences. The following two are singled out by Koikkalainen (1992:17-19).[1]

---

[1] See e.g Dayhoff (1990), Hautamäki (1990), Hecht-Nielsen (1990), Kohonen (1988), Rumelhart and McClelland (1986), Seppälä (1992), Vadén (1992) and Weiss and Kulikowski (1991) as presentations of various aspects of connectionist models.

Firstly, many brain operations are not realizable in a sequential machine. In the brain parallelism is massive, there are about 10^{10} to 10^{11} processing elements, neurons, and each of them receives an average of 10^4 direct connections from other neurons. Secondly, what makes the brain really different from computers is that neurons as basic computing elements influence each other's response to stimuli. Hence a network of neurons can adapt and learn from input patterns. The exact mechanism of learning is unknown but the current opinion is that the information is stored in connections, synaptic weights, between the neurons.

Connectionist modelling is inspired by our knowledge of the nervous system. Certain kinds of connectionist networks are therefore called artificial neural networks. Also the phrase "parallel distributed processing" (PDP) is sometimes used.[2]

In the following traditional (symbolic) methods and connectionist models are compared. The comparison focuses on the following questions: What is the nature of representation? What kind of reasoning process is involved? What kind of possibilities are there to generalize automatically from examples?

2.1. Some traditional methods for representation and generalization

Semantic networks are one of the traditional ways of representing knowledge. A net consists of a set of nodes and directed links connecting the nodes. Nodes may refer to objects or properties, and links are used to represent relations. One might, for example, model the sentences *This is a red brick, It is also a toy* using a semantic net depicted in figure 1b.[3]

---

[2] An influential work in the connectionist enterprise has been Rumelhart and McClelland's two volumes using the phrase PDP: Rumelhart and McClelland (1986), McClelland and Rumelhart (1986).

[3] The knowledge in semantic nets can also be represented using predicate logic in the following manner: ∃x: Brick(x) & Toy(x) & Red(x) or, even ∃x: Brick(x) & Is_a(x, brick) & Is_a(x, toy) & Color(x, Red)

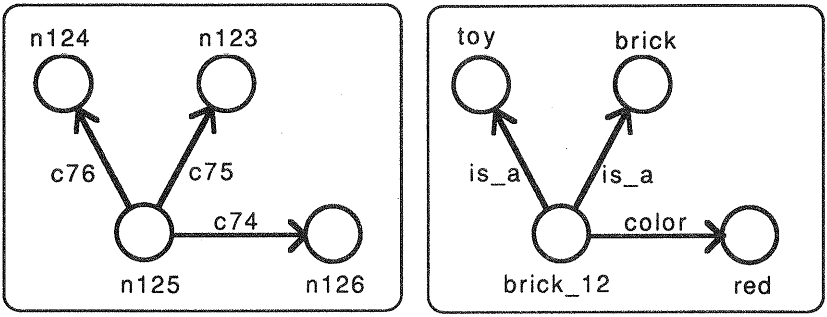

Figure 1. (a) A net and (b) a semantic net.

Semantic nets (figure 1b) are distinguished from ordinary nets (figure 1a) by their inclusion of semantics (Winston 1984:253). A semantic net is used to represent the reality explicitly. A meaning is associated both with the nodes and with the links of the network. Behind this kind of representational apparatus is the ontological view of reality as consisting of a set of discrete entities and a set of relations between them. The very same assumptions limit "the view of the world" of classical logic. Words in natural languages, however, are seldom entities with such precise meanings and, therefore, cannot be accurately modelled with symbolic logic. A problematic example familiar to linguists is that of mass nouns. Also, the meaning of a word like *big*, is not an entity with fixed boundaries precisely and constantly separating what is big from everything that is not big. Much more commonly, a meaning is fuzzy and changing, biased at any moment by the particular context. (Honkela and Vepsäläinen 1991: 897.)[4]

Explicitness similar to that of semantic networks can also be seen in tree-like representations of syntactic structures. A parse tree formed using a dependency grammar consists of nodes referring to the words of the parsed sentence and links denoting the dependency relations. In various formalisms nodes and links may refer to words, relations,

[4] This line of reasoning does not imply that external reality does not exist. It is only stated that an object-oriented way of modeling has it deficiences: the lack of means of dealing with e.g. continuous and chaotic phenomena of reality.

functions, constituents or other symbolic and explicit parts of syntactic analysis of a sentence. One may ask whether such nodes and links are real from the cognitive point of view.

Inductive inference as learning

Karlgren (1990:97) motivates the study of machine learning in the following way:

> "One theme which I see as crucial in computational linguistics at this particular point of time is machine learning ... Modeling learning is interesting in itself but modeling language user's learning and adaptation also attacks one of the most salient features of natural languages and one of which so far is intriguing feature that human users understand utterances and texts by means of knowledge about the language system and that such knowledge is successively acquired from the utterances and texts we understand. To get a relevant model for human linguistic competence we must teach machines to learn: to update their grammar and lexicon from the very texts on which they apply them ... It is my belief that there are basic procedures, as yet poorly understood, which are common to language change over longer periods, language acquisition by an individual and the mutual adaptation between dialogue participants or the reader's adaptation to the author during and possibly merely for the purpose of the current dialogue or text."

The area of machine learning is diverse (see e.g. Honkela and Sandholm 1992) but the main emphasis has traditionally centered around inductive reasoning. Whereas deductive reasoning makes existing knowledge explicit, inductive reasoning is meant to create general laws from specific examples. An inductive conclusion has the following properties: (a) It is consistent with the examples, and (b) it explains the examples.

A system might look for general properties of English words. If there are two examples - *give* and *great* - there are several possible generalizations, for example:

a. All the words are English (no others are encountered),

b. words with letter e in them are English, or

c. words beginning with the letter g are English.

If the system is given the Swedish word *gata* as a negative example, it must ignore the hypotheses (a) and (c). [5]

It is important to remember that inductive conclusions are defeasible (see also Levinson 1983:114). How is this defeasibility dealt

[5] The example is simplified on purpose and is for illustration only.

with? Traditional methods often use a no-guessing principle: when there is doubt about what to learn, learn nothing (Winston 1984:395).[6]

2.2. Connectionist networks

One may ask whether there are any other ways of using networks than attaching explicit meanings to all the nodes and links to represent e.g. linguistic knowledge. Yes, there are, and these alternatives, connectionist networks, are the main theme of this article. Such networks can be characterized in the following way.

A connectionist network consist of nodes and connections between them where those connections do not have any individual and explicit semantic label associated to them.[7]

In a connectionist network each node has some degree of activation. Active nodes may excite or inhibit other nodes.

Nodes with explicit semantics

As an example we might examine a network used for word sense disambiguation (Veronis and Ide 1990). Each word in the input is represented by a word node connected by excitatory links to sense nodes representing the different possible senses for that word in the Collins English Dictionary (ibid. 391). Each sense node is in turn connected by excitatory links to word nodes representing the words in the definition of that sense. Inhibitory links are created between different meanings of the same word. Through this kind of process, a network with thousands of nodes is created. A part of this kind of network is shown in figure 2.

---

[6] As a thoroughful description of inductive reasoning and processes, see Holland et al. (1986).

[7] There are also more restrictive definitions of connectionist models. Koikkalainen (1992:43) makes a clear distinction between connectionist models and so called artificial neural networks: "Perhaps the most striking feature in connectionist models is that there are so called "grandmother" cells, neurons that have a symbolic label like 'table', 'apple' or 'green'.". Here a hierarchical relation is adopted: artificial neural networks are a special kind of connectionist networks.

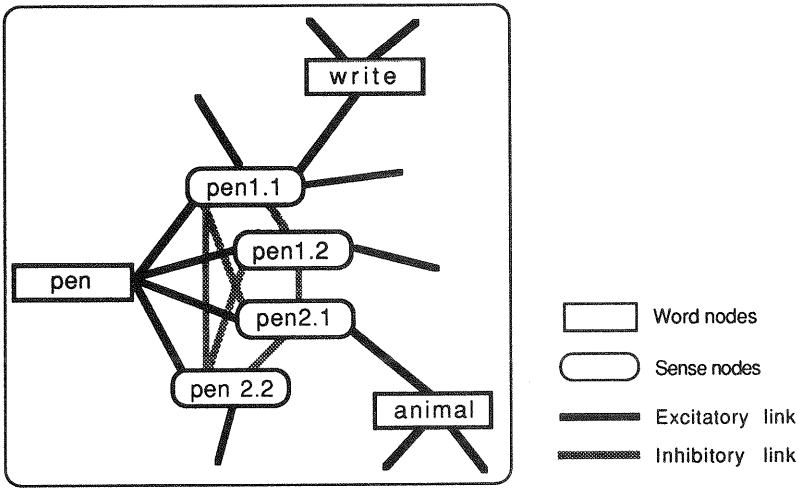

Figure 2. Connectionist network as a representation of a lexicon for disambiguation purposes (modified from Veronis and Ide 1990:392).

The use of the network is based on spreading activation (see next page).

Nodes with no explicit meaning

In the network of (Veronis and Ide 1990) all nodes have an explicit meaning. A node is either a word node or a sense node. The discrete set of senses is determined using a dictionary. There are also connectionist models where some nodes do not have explicit meaning. To illustrate, let us first examine backpropagation network architecture. A back-propagation network consists of an input layer of nodes, a layer of hidden nodes and an output layer of nodes (figure 3).

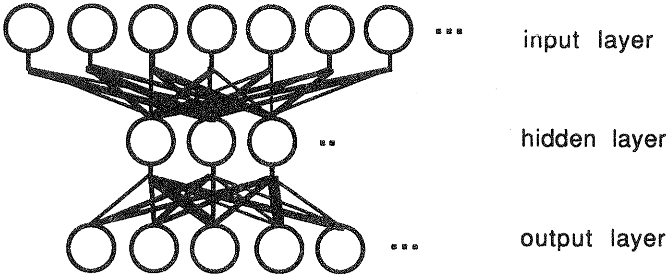

Figure 3. A backpropagation network architecture.

The important fact is that there is no semantic label attached to elements of the hidden layer. Their influence is determined by the learning process. The meaning of the input and the output elements depends on the application.[8]

Spreading activation

The basic idea behind spreading activation is that the nodes of a network influence each other through the connections. Each node has an activation level and each connection has a strength. Both activation levels and strengths are usually real numbers. The strength may - for example - be limited between -1 and +1. In the case of a strength of -1 there is maximal inhibition and accordingly, a strength of +1 means maximal excitation.[9]

Usually there are two kinds of nodes: those that can receive external input and those that are influenced only by the other nodes in the network. The latter ones are often called hidden nodes.

One task in designing a connectionist network is to determine which nodes are connected, i.e., the pattern of connectivity. There are two basic kinds of networks in this respect. Feedforward networks have unidirectional connections. Inputs are fed into one layer (input), and

---

[8] Nowadays the majority of the applications deal with pattern recognition, e.g. the analysis of pictorial images and speech.

[9] Bechtel and Abrahamsen (1991) outline these principles using examples aiming at a presentation for readers less familiar with mathematics. Hecht-Nielsen (1990) gives a detailed description of the connectionist computing techniques.

outputs are generated at the output layer as a result of the forward propagation of activation.[10] Interactive networks have connections which propagate activation to both directions.

Veronis and Ide (1990:392) describe the spreading of activation in their model for disambiguation in the following manner. When the network is run, the input word nodes are activated first. Then each input word node sends activation to its sense nodes, which in turn send activation to the word nodes to which they are connected, and so on throughout the network for a number of cycles. At each cycle, word and sense nodes receive feedback from connected nodes. Competing sense nodes send inhibition to one another. Feedback and inhibition cooperate in a winner-take-all strategy to activate increasingly related word and sense nodes and deactivate the unrelated or weakly related nodes. Eventually, after a few dozen cycles, the network stabilizes in a configuration where only the sense nodes with the strongest relations to other nodes in the network are activated. For example, given the sentence *The young page put the sheep in the pen* the network correctly chooses the correct senses of page ("a youth in personal service"), sheep and pen.

Learning in artificial neural networks

What is the distinction between connectionist systems and artificial neural networks? The conventions are still evolving but it seems reasonable to define that an artificial neural network includes all of the following:

- a network architecture with nodes and links, where at least links do not have explicit meaning,
- an activation principle and
- a learning principle.

This means that there are a number of connectionist systems which are not artificial neural networks because they do not learn.

---

[10] The flow of activation is determined by well-defined mathematical equations. Exact details vary but the basic ideas are the same for most of the models. Output for a node is straightforwardly same as its activation if the activation is over zero. Usually there are a number of input connections to a node (even thousands). The effect of all the inputs is computed as a sum of the single inputs from each incoming connection. A single input is usually computed as a product of the activation of the node and the strength of the connection. The input then effects the activation of the node.

There are a number of different artificial neural network models. In the following, two of them - backpropagation and self-organizing maps - are studied more closely.

Backpropagation

The backpropagation neural network is the most widely used network nowadays (Hecht-Nielsen 1990:125). The architecture described in the following is the basic one. Many variants of this basic form exist. In the general case, a backpropagation network consists of n layers, where n is usually 3 or greater. The following description is based on an architecture with three layers (see figure 3). The first layer is an input layer which simply takes the inputs in and distributes them, without modification, to all the nodes in the second layer. The second layer is usually called the hidden layer. Each node on hidden layer receives the output signal of each of the nodes of the input layer. The third layer is the output layer which in turn receives the output of the nodes of the hidden layer.

Teaching a backpropagation network is based on a set of examples. Each example has an input and the corresponding correct output. The network's operation during training consists of two sweeps through the network. The first sweep starts by giving the input to the nodes of the input layer. The forward spreading activation then reaches the output layer. The second sweep is ready to start. It is based on the deviations between the network's actual result and the desired result (error).[11] The error for each node is propagated back (hence the name) and the weights of the connections are modified so that the network is more likely to give the correct answer next time. This kind of process is continued until the network reaches a satisfactory level of performance, or until the user gives up.[12]

What are the input-output pairs presented to the network? The applications vary from recognition of handwritten characters to sentence processing. In the study by McClelland and Kawamoto (1986) the model

---

[11] One may e.g. think of a system which recognizes hand written characters. The first sweep might give 'E' as a result on the output layer, the correct output being 'C'. this deviance motivates the second sweeps, which tries to correct the behaviour of the network.

[12] A detailed description of backpropagation is given in numerous sources. The description in Hecht-Nielsen (1990) was used here, though strongly shortened.

consists of two sets of units: one for representing the surface structure of the sentence and one for representing its case structure.

Lexicon

item	corresponding vector
cats	1 0 0 0
dogs	0 1 0 0
hate	0 0 1 0
love	0 0 0 1

"cats hate dogs" $\longrightarrow$ $\underbrace{\boxed{1}\boxed{0}\boxed{0}\boxed{0}}\underbrace{\boxed{0}\boxed{0}\boxed{1}\boxed{0}}\underbrace{\boxed{0}\boxed{1}\boxed{0}\boxed{0}}$

Figure 4. Simple example of mapping a task to neural networks: preprocessing of a three word sentence[13]

Because neural networks take numerical data as input, one has to preprocess symbolic data. A simplified example of coding sentences is presented in figure 4.

Self-organizing maps

The learning strategy of the backpropagation networks is supervised: for each example input there must also be "a right answer" as a correct output. The system then learns according to these input-output pairs. The task is not trivial, though, while after the learning period the network is able to deal also with inputs which were not present in the learning phase. This possibility is enabled by the generalization capabilities of the network.

---

[13] It must be emphasized that the input values for a network need not be binary (i.e. 0 or 1).

Kohonen (1982) has developed the self-organizing map (SOM) neural network paradigm. A SOM network need not be given any "right answers". The cells of the network become specifically tuned to various classes of patterns through a learning process. In the basic version, only one cell of a local group of cells at a time gives the active response to the current input. The locations of the responses tend to become ordered as if some meaningful coordinate system for different input features were being created over the network. The coordinates of a cell in the network then correspond to a particular domain of input patterns. (Kohonen 1990.)

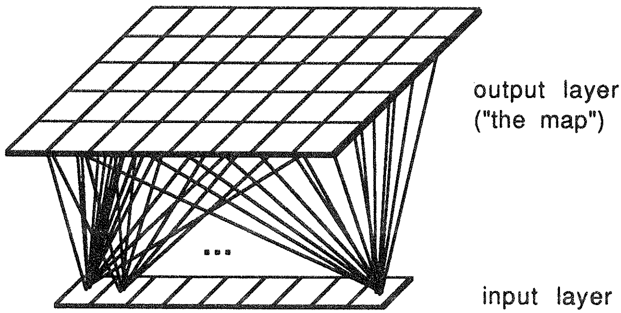

Figure 5. The basic architecture of a self-organizing map with a two-dimensional grid of cells on the output layer.

The ordering process has been shown to give meaningful results in various areas of use. One might, for example, input the network a series of pictures. A SOM in a sense looks for similarities between the pictures taking into account the statistical properties. An illustration of an ordered map is given in figure 6.

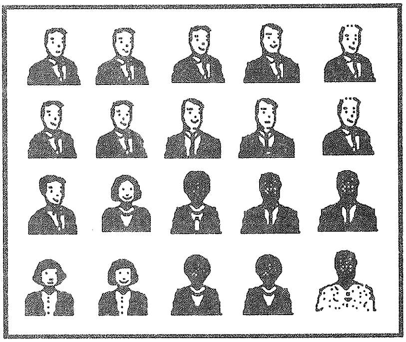

Figure 6. An illustration of an ordered map.

The use of unsupervised learning is grounded especially in the cases where no correct outputs are available by practical reasons or even "by definition" (matters of subjectivity).

3. EXAMPLES OF CONNECTIONIST LINGUISTIC MODELS

There are a number of experiments in which a connectionist model has been used to model a particular linguistic phenomenon. In the following, some of those studies are presented in two sections according to the linguistic level of the approach (structure versus content).

3.1. Models of morphology and syntax

It may be concluded that much of the connectionist linguistic study concentrates on syntax. Many artificial neural network models have been developed for speech recognition (see e.g. Kangas 1992) but a minority of the research is linguistically motivated. At the level of morphology, Koskenniemi (1983:134-136) discusses the relation between finite state automata (in the two-level model) and neural networks. A number of experiments have been made in disambiguation (e.g. Cottrell 1985, Veronis and Ide 1990). The use of neural networks for disambiguation has similarities with the use of statistical models. Connectionist disambiguation is based on the idea that a network is taught by giving it a number of examples in which the correct interpretation of an ambiguous word or expression has been given. It is crucial that enough context has been given.

Much of connectionist research concerning syntax relies on the traditional framework of well-known grammars (as examples Faisal and Kwasny 1990, Kamimura 1991, Nakamura et al. 1990, Schnelle and Wilkens 1990). It is also possible to apply a more radical approach and use implicit categories or try to build a network which autonomously creates categories. It has also been questioned whether any symbolic categories are needed.

Connectionist approaches have been criticized by claiming that a proper linguistic method should have a possibility of representing constituent structures (Fodor and Pylyshin 1988). As an answer to the criticism, Niklasson and Sharkey (1992) have developed a connectionist model which implements non-concatenative compositionality by using the Recurrent Auto-Associative Memory (RAAM) neural network model devised by Pollack (1990). The presentation of a complex expression like *NP1 & (V & NP2)* could be generated in the way shown in figure 7. Each of the constituents "V", "&" and "NP2" is represented by n nodes. Each of these constituents is presented to the network. Then, the distributed non-symbolic representation at hidden layer of the expression is combined with the representations for "&" and "NP1".

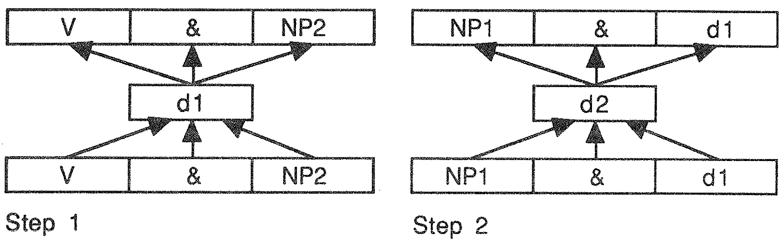

Figure 7. Generation of complex expressions (adapted from Niklasson and Sharkey 1992). Nodes "d1" and "d2" are distributed representations of complex constituents.

The RAAM architecture provides the means for generating complex representations which consists of constituents that themselves are either complex or atomic. Niklasson and Sharkey (1992) also show how to train a network to make transformations on the distributed non-symbolic representations of the expression generated by RAAM.

3.2. Modelling semantics using self-organizing maps

The first difficulty of connectionist linguistic modelling is encountered when trying to find metric distance relations between symbolic items. It can not be assumed that encodings of symbols in general have any relationship with the observable characteristics of the corresponding items. As a solution to the problem, it is possible to present the symbol in context during the learning process. In linguistic representations, context might mean adjacent words. Similarity between items would be reflected through the similarity of the contexts. (Kohonen 1990.)

Ritter and Kohonen (1989) have presented in their work a self-organizing system which creates representations of lexical relationships. A semantic map is formed during a self-organizing process. Ritter and Kohonen used two kinds of input materials. Firstly, they trained the network using simple sentences where a word was presented in its context. In the other experiment they used discrete attributes attached to a set of words. Both experiments were successful. Nouns, verbs and adverbs are automatically segregated into different domains on the map. Within each domain a further grouping according to aspects of meaning is discernible (Kohonen 1990:1476).

The self-organizing map has been used also by Scholtes (1991) Schyns (1990) and Honkela and Vepsäläinen (1991) to model various phenomena related to semantics.

4. THE POTENTIAL OF CONNECTIONIST MODELLING IN SEMANTICS AND PRAGMATICS

There are tasks in which reality or "pictures of reality" are mapped into linguistic expressions. Finding "entities" from a picture is not a trivial task, as revealed by attempts to give computers such pattern recognition abilities. Attempts to specify the features of an entity have usually succeeded only with highly constrained unnatural stimuli. A similar problem exists in the expression of natural languages. Through a gradual process of learning, people develop exquisite skills for dealing with words despite their imprecision and contextual dependency. People are fairly good at mapping continuous parameters (e.g. size) into apparently discrete expressions (*tiny*, *big*, etc.).

A person understands that there may be subjective differences (*big* may mean something different to a child than to a adult), strong contextual influences (*big* in *big city* has different connotations than in *big fly*), and imprecision (in a given context, a person may reliably call one stimulus *moderate* and another *big*, but in between is a gray range of stimuli not clearly one or the other). A person also reacts to the "surplus meanings" and associations of a word. E.g. *large* is a more sophisticated, less childish word than *big* and thus more likely to be used in scientific writing (*a large difference between groups*) or advertising aimed at adults (*a large automobile*). All these shades of meaning are dealt with accurately and indeed employed usefully by most adults in their language usage and understanding. (Honkela and Vepsäläinen 1991:897-898.)

4.1. Representation of imprecise concepts

Providing a natural representation of a large set of concepts requires some soft constraints or, more specifically, the use of membership functions - like those in fuzzy set theory (Zadeh 1983) - and statistical descriptions. The following illustrates the need for such devices.

In traditional syntactic analysis, various categorizations are used. One may compare the inclusion of a group of words into a category ("... are verbs") and the use of the categories in abstract rules ("a verb may ..."). It may seem that the abstract rules are precise, but when they are applied, it is to be noted that discrepancies exist between the rules and the linguistic phenomena. A rule may be seen to be incorrect in various ways:

(1) A rule may be overtly generalized (like "all English nouns are preceded by an article"). This kind of situation should lead to the refinement of the rule[14]. The more thorough the test for the rules is, the more likely it is that there are cases in which a rule does not work.

(2) It may be found out that some of the words or structures in a category tend to behave in a distinct manner in a certain context. Therefore, it may be better to create a new category for those exceptional

---

[14] In inductive reasoning and machine learning the processes involved here are called specializing and generalizing (see e.g. Winston 1984:385-394).

words or structures rather than try to take the exceptions into account in the rule level.[15]

(3) The reason for a failure of a syntactic rule when tested against some real data may be on another level. It is usual that a syntactic rule is too general to take into account semantic or pragmatic distinctions.The use of a linguistic structure may be guided by the context dynamically. Sometimes it is even possible that the speakers create some "rules of their own" to last only during that particular discussion.

There are several possibilities to deal with these difficulties:

- The rules and categories are refined to match the actual phenomena as closely as possible.
- The conditions for the success of a linguistic description are explicated as precisely as possible. This may include restrictions concerning style etc.
- Some statistical measures are connected to the rules. One may test the rules using large corpora, and then attach a probability of success for each rule using the results. (see e.g. Ejerhed 1990)

One important problem is how to acquire the descriptions of impreciseness (e.g. membership functions in fuzzy sets). The use of unsupervised connectionist learning can be seen as a potential solution to the problem. The activity level in the output of an artificial neural network might be interpreted (in proper conditions) as a degree of membership in a fuzzy set or even as a fuzzy truth value of a proposition.

The learning process can be based on material which consists of words or phrases in accordance with a textual context (Ritter and Kohonen 1989), symbolic features (ibid), continuous values of some parameters (Honkela and Vepsäläinen 1991), or even pictorial images.

Among others Smolensky (1986) and Cussins (1990) have studied the possibility and the nature of connectionist concepts (see also Vadén 1991, Itkonen 1992 and Vadén 1992). Also Wildgen and Mottron (1987) have analyzed the possibility of linguistically oriented self-organizing processes.

---

[15] A pessimistic view into this process would be that finally - after a thorough modelling and testing process - practically all words have a category of their own.

4.2. Pragmatics

In his analysis of delimiting the area of pragmatics Levinson (1983:21-22) draws attention to work in artificial intelligence. There the term language understanding is used because of the fact that understanding an utterance involves more than knowing the meanings of the words uttered and the grammatical relations between them.

What are the possibilities of connectionism concerning pragmatics? The study of this area is in its very beginning. One might list some possibilities:

- modelling conversational aspects and
- modelling mutual knowledge, subjectivity and intersubjectivity.

In a traditional approach one might model a conversational situation where the speaker and the listener know/believe certain propositions. It has been difficult to model situations where the persons differ in their assessment of truth value (or degree of truthfulness), or the persons do not share a similar view on the meanings of the linguistic expressions.

Consider, for example, a boy who tells his mother *I'll be home at two o'clock* but does not arrive until about three. The mother may be very angry, saying *You never come back when you promise.* But in another version of the same story, the mother might be delighted. What does *at two o'clock* mean? One possibility is that of complete ambiguity: the expression means to the other 2 pm and to the other 2 am or different days. This kind of phenomenon is easily dealt with symbolic, discrete descriptions. The more challenging and possibly more common source of misunderstanding is the possible impreciseness of the expression *at two o'clock.* It may mean to someone an interval from one to three and to someone else an interval from 10 to three to three o'clock.

The interpretation of an expression is often context-dependent on various ways: depending on the utterer, the listener and the situation. The interpretation tends to be narrower if the utterer and listener are not familiar to each other. The interpretation depends on the formality of the situation (business, family, holiday etc.) and possible activities related to the time expression: *I'll come back at two o'clock* is taken more precisely if there is a mutual knowledge of a meeting, a tennis hour or a train leaving - to show some examples. In summary, any simple time expression has numerous interpretations which are determined by the context. The context is very complicated, and there is an interval or more precisely, a subjective probability distribution involved concerning the interpretation. These aspects are very difficult - or even impossible - to model using traditional formal symbolic methods.

Another crucial aspect in accordance with context-dependency concerning many conversational situations is the adaptation or learning involved. Learning during a discussion may have to do with

- the subject matter (e.g. A starts to tell to B what a certain computer is and why it is good), or
- the interpretation of expressions by the other subjects (*Oh, that's your conception of goodness. I can understand your personal view, but it's not relevant to me, because ...*)

In a long process people learn to interpret natural language expressions and also learn to understand at least some of the differences in the interpretation between other people.[16]

4.3. Contextuality

Pragmatics may also be defined to be the study of the ability of language users to pair sentences with the contexts in which they would be appropriate (Levinson 1983:24). Artificial neural networks (ANN) could be used to learn such pairings. Important in this respect is the possibility to enlarge the input and output vectors of ANNs. One can - in principle - easily take into account various aspects of the context. Practical problems are caused by (1) the amount of "experience" needed (how to collect all the data), and (2) the present day limitations concerning the size of the ANNs.[17]

There are a number of experiments where an ANN is taught to recognize the grammatically correct sentences. One might also try to teach various other aspects where much more knowledge of the context is needed. Here one can see a solution to the problem of the requirement for a fundamental idealization of a culturally homogeneous speech community or, alternatively, the construction on n pragmatic theories for each language (ibid. 25). The appropriateness conditions could be modelled

---

[16] These phenomena are significant in many areas of life (not to mention the questions of war and peace...)

[17] It must remembered that our linguistic capabilities especially in the area of pragmatics rely on a vast experience gathered during decades. This is one practical reason why it is unreasonable to expect artificial systems to compete with human beings in all the areas of natural language use.

using a connectionist network which adapts to the fine-grained varieties of the context and which may also adapt to take into account the developments in the conditions. It is also to be noted that ANNs have generalization capabilities which ensure that the situations which can be successfully dealt with can be different from any of those met before. [18]

4.4. Change and diachronic linguistics

There are several classical paradoxes which are related to the sameness of entities and change. Pylkkö (1989) analyses some of those paradoxes (concerning e.g. Shakespeare's identity in various situations) and ends up with the claim of physical objects to be cognitive fictions. Von Foerster (1981) has made same kind of conclusions:

- The logical properties of invariance and change are those of representations. If this is ignored, paradoxes arise.
- Objects and events are not primitive experiences. Objects and events are representations of relations.
- Operationally, the computation of a specific relation is a representation of this relation.

A pattern recognizing neural network does this kind of computation: it looks for objects from a scene.

4.5. Subjectivity

The use of connectionist models makes it possible to model imprecise boundaries between concepts and their contextual dependency. Unsupervised learning can be used to model aspects of individual differences in the natural language interpretation, i.e. subjectivity of meaning.

The activity patterns which result from an input vector vary according to the examples presented to the network ("experience"). The input may contain a word or expression for which one wishes to see the

[18] An interesting task would be to teach an ANN to recognize irony. The experiment should be focused on a certain area of subject matter. One might give a rule for irony: if A utters an expression which (1) B knows to be false and (2) B knows that (it is at least likely that) A knows that the expression is false, then B can suppose that A uttered an ironic expression. The problem for B is to check whether A really knows that the expression is false. Sometimes there are multiple sources (sound, facial expressions).

interpretation. One might also give a representation of the situation (of the context) to select the expression with strongest response.

The model for subjectivity includes differences in interpretation between an expert and a novice, an adult and a child, or a native speaker and a foreigner. An expert tends to use more specific and precise terms than a novice. In a multidimensional description generated by an artificial neural network the pattern of use of an expert is likely to be more complex.

Mutual understanding in conversation depends on the selection of words and expressions. Understanding is based on the intersubjective agreement on the meanings of the expressions. The activation patterns could be used to model the degree of this agreement. If the activation patterns of two persons are similar enough, a ground for mutual understanding exists. In some cases the background of persons gives rise to varying interpretations of expressions. The risk lies in the fact that often people do not have the possibility to check the interpretation of the utterer or the listener. (Honkela 1992.)

5. CONCLUSIONS

Linguists can respond to connectionism in at least two ways: they can take it as a challenge, or as an ally. The following is Bechtel and Abrahamsen's (1991:295-295) analysis of these two positions.[19]

1. Connectionism can be seen as a challenger to the traditional linguistics. It is possible to view as a challenge the approximationist claim that that explicit linguistic rules need not be mentally represented, and that rules merely approximate the more detailed representation provided by connectionist models. If one requires that linguistic analyses should conform to psychological processing, the connectionism, if successful, would have dramatic consequences for linguistic analyses in the Chomskian tradition.

2. Adherents of cognitive linguistics have welcomed connectionism as an ally in their psychologically-oriented alternative to Chomskian linguistics. Among others, Langacker (1987) denies the autonomy and primacy of syntactic analysis; instead, semantics is regarded as funda-

---

[19] Bechtel and Abrahamsen theselves state that they would be inclined to regard analyses of cognitive linguistics in a connectionist framework as a psycholinguistic rather than linguistic theory, leaving a gap at the most abstract level of analysis.

mental. In cognitive linguistics a subjectivist or conceptualist analysis of language is advocated. Both the grammar and meaning of expressions are seen to be founded on the body of knowledge that speakers possess, the mental models they build, and the mappings they make between domains of knowledge.

This article has presented the relationship of linguistics and connectionism in a rather optimistic vein. It remains to be seen how fruitful the connection between linguistics and connectionism is.

REFERENCES

Bechtel, W. and Abrahamsen, A. 1991. Connectionism and the Mind. Cambridge, Massachusetts: Basil Blackwell.

Churchland, P.M. 1989. A Neurocomputational Perspective: The Nature of Mind and the Structure of Science. Cambridge, Massachusetts: MIT Press.

Cottrell, G.W. 1985. A connectionist approach to word sense disambiguation. Technical Report TR 154, New York: University of Rochester, Department of Computer Science.

Cussins, A. 1990. The connectionist construction of concepts. The Philosophy of Artificial Intelligence, Boden, M.A. (ed.), Oxford University Press. 368-440.

Dayhoff, J. 1990. Neural Network Architectures; An Introduction. New York: Van Nostrand Reinhold.

Ejerhed, E. 1990. On Corpora and Lexica. The 1990 Yearbook of the Linguistic Association of Finland. Helsinki: Suomen kielitieteellinen yhdistys. 77-96.

Faisal, K.A. and Kwasny, S.C. 1990. Design of a Hybrid Deterministic Parser. Coling-90, Vol 1. 11-16.

Fodor, J.A. and Pylyshin, Z.W. 1988. Connectionism and cognitive architecture: A critical analysis. In: Pinker, S. and Mehler, J.: Connections and Symbols. Cambridge, Massachusetts: MIT Press. 3-71.

Foerster, H. von. 1981. Notes on an epistemology for living things. In: Observing systems, Intersystems publications. 258-271. (Originally published as "Notes pour une epistemologie des objects vivants" in

L'Unite de L'Homme: Invariants Biologiques et Universaux Culturel, Morin, E. and Piattelli-Palmerini, M. (eds.), Editions du Seul, 1974)

Hautamäki, A. 1990. Hermoverkot: Periaatteet, tekniikka ja filosofinen tulkinta. In: Tekoäly ja filosofia, Marjomaa, E. (ed.), University of Tampere, Tampere, 1990.

Hecht-Nielsen, R. 1990. Neurocomputing. Reading, Massachusetts: Addison-Wesley.

Holland, J.H., Holyoak, K.J., Nisbett, R.E. and Thagard, P.R. 1986. Induction; Processes of Inference, Learning, and Discovery. Cambridge, Massachusetts: MIT Press.

Honkela, T. and Vepsäläinen, A.M. 1991. Interpreting imprecise expressions: experiments with Kohonen's self-organizing maps and associative memory. Proceedings of the International Conference on Artificial Neural Networks (ICANN-91). Elsevier Science Publishers. 897-902.

Honkela, T. 1992. Connectionist semantics - remarks on subjectivity, continuity and change. A paper presented at International Conference on Cognition, Connectionism and Semiotics, June 1-3, Tampere, Finland.

Honkela, T. and Sandholm, T. 1992. Koneoppiminen. In: Tekoälyn ensyklopedia, Hyvönen, E. et al (eds.), Finnish Artificial Intelligence Society (in preparation), 1992, 15 p.

Itkonen, E. 1992. The Mental Representation of Natural Language. Proceedings of STeP-92, vol. 2. Helsinki: Finnish Artificial Intelligence Society. 60-66.

Kamimura, R. 1991. Acquisition of the grammatical competence with recurrent neural network. Proceedings of the International Conference on Artificial Neural Networks (ICANN-91), Elsevier Science Publishers. 891-896.

Kangas, J. 1992. Neurotietokone puheenkäsittelyssä. Hautamäki, A. and Nyman, G. (eds.): Kognitiotiede ja koneäly. Helsinki: Suomen tekoälyseura. 139-146.

Karlgren, H. 1990. Computational Linguistics in 1990. Coling-90, Vol 1. 97-99.

Kohonen, T. 1982. Self-organized formation of topologically correct feature maps. Biological Cybernetics, 43. 59-69.

Kohonen, T. 1988b. An Introduction to Neural Computing. Neural Networks, vol. 1, no. 1. 3-16.

Kohonen, T. 1990. The Self-Organizing Map. Proceedings of IEEE, vol. 78, no. 9. 1464-1480.

Koikkalainen, P. 1992. Neurocomputing Systems: Formal Modeling and Software Implementation. Lappeenranta University of Technology, Research papers 23 (Diss.).

Koskenniemi, K. 1983. Two-Level Morphology: A General Computational Model for Word-Form Recognition and Production. University of Helsinki, Department of General Linguistics, Publications, no. 11.

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

Levinson, S.C. Pragmatics. 1983. Cambridge: Cambridge University Press.

McClelland, J.L. and Kawamoto, A.H. 1986. Mechanisms of Sentence Processing: Assigning Roles to Constituents. In (McClelland and Rumelhart 1986). 272-325.

McClelland, J.D. and Rumelhart, D.E. (eds.). 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 2: Psychological and Biological Models. MIT Press. 272-325.

Nakamura, M., Maruyama, K., Kawabata, T. and Shikano, K. 1990. Neural Network Approach to Word Category Prediction for English Texts. Coling-90, Vol 3. 213-218.

Niklasson, L. & Sharkey, N.E. 1992. The miracle mind model. The First Swedish National Conference on Connectionism, Advance Proceedings. Skövde: University of Skövde.

Pollack, J.B. 1990. Recursive Distributed Representations. Artificial Intelligence, 46. 77-105.

Pylkkö, P. 1992. Connectionism and Associative Naming. Proceedings of STeP-92, vol. 2. Helsinki: Finnish Artificial Intelligence Society.

Ritter, H. and Kohonen, T. 1989. Self-Organizing Semantic Maps. Biological Cybernetics, 61. 241-254.

Rumelhart, D.E. and McClelland, J.L. (eds.) 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations. MIT Press.

Schnelle, H. and Wilkens, R. 1990. The translation of constituent structure grammars into connectionist networks. Coling-90, Vol 1. 53-55.

Scholtes, J.C. 1991. Kohonen Feature Maps in Natural Language Processing. Amsterdam: University of Amsterdam, Department of Computational Linguistics.

Schyns, P.G. 1990. Expertise Acquisition Through the Refinement of a Conceptual Representation in a Self-Organizing Architecture. Proceedings of the IJCNN-90. Washington DC. 236-240.

Seppälä, T. 1992. Churchlandin neurokomputationaalinen näkökulma: Hermoverkko tieteenfilosofian perustaksi? Proceedings of STeP-92, vol. 2. Helsinki: Finnish Artificial Intelligence Society. 244-253.

Smolensky, P. 1986. Neural and Conceptual Interpretation of PDP Models. In [Rumelhart & McClelland 1986]. 390-431.

Vadén, T. 1991. Adrian Cussinsin ja Paul Smolenskyn konnektionistisista näkemyksistä. Marjomaa, E. ja Vadén, T. (eds.): Ihmisen tiedonkäsittely, symbolien manipulointi ja konnektionismi, Tampere. 220-237.

Vadén, T. 1992. Alisymbolinen konnektionismi. Tampereen yliopisto.

Weiss, S.M. and Kulikowski, C.A. 1991. Computer Systems That Learn; Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning, and Expert Systems. San Mateo, California: Morgan Kaufmann Publishers.

Veronis, J. and Ide, N.M. 1990. Word Sense Disambiguation with Very Large Neural Networks Extracted from Machine Readable Dictionaries. Coling-90. 389-394.

Wildgen, W. and Mottron, L. 1987. Dynamische Sprachtheorie: Sprachbeschreibung und Spracherklärung nach den Prinzipien der Selbstorganisation und der Morphogenese. Bochum: Brockmeyer.

Winston, P.H. 1984. Artificial Intelligence. Reading, Massachusetts: Addison-Wesley Publishing Company.

Zadeh, L.A. 1983. The role of fuzzy logic in the management of uncertainty in expert systems. Gupta et al (eds.) Approximate Reasoning in

Expert Systems, Elsevier Science Publishers. Reprinted from Fuzzy Sets and Systems 11, Elsevier Science Publishers.

Address: Technical Research Centre of Finland
Laboratory for Information Processing
Lehtisaarentie 2 A, 00340 Helsinki
e-mail: tho@tik.vtt.fi