The relationship between connectionist models and a dynamic data-oriented theory of concept formation[1]

Renate Bartsch

ILLC, Faculteit der Wijsbegeerte

Universiteit van Amsterdam

In this paper I shall compare two models of concept formation, both inspired by basic convictions of philosophical empiricism. The first, the connectionist model, will be exemplified by Kohonen maps, and the second will be my own dynamic theory of concept formation. Both can be understood in probabilistic terms, both use a notion of convergence or stabilization in modelling how concepts are built up. Both admit destabilization of concepts and conceptual change. Both do not use a notion of representation in some pre-given language, such as a language of thought or some logical language. Representation in a formal language only plays a role on the meta-level, namely within the theory about concept formation.

I. Two approaches to concept formation: An overview

For a connectionist account of concept formation, I shall concentrate on Kohonen's topological maps and modifications of these (Kohonen 1984, 1991; Ritter et al. 1992, Schyns 1991, Kangas 1990, Scholtes 1993). In a standard Kohonen model, there is a layer of n sensor units, all connected with all units of a two-dimensional hidden layer, which in the course of

---

[1]A short version of this paper is given at the Fachtagung der Gesellschaft für Kognitionswissenschaft. Freiburg, October 12-15th, 1994.

learning becomes the conceptual map. Each hidden unit has a vector of values, which is the vector of the weights of the connections from the input sensor units to the respective unit. To begin with, these connections have randomly assigned weights. The input vectors to the map have n places or dimensions, which is the number of input sensors. Each sensor unit assigns a value for the respective dimension, per instance presented. Thus each input instance gets a sensor-value vector. Every new situation stimulates in principle more or less all units of the map via the connections from the sensor units. The hidden unit with a vector closest to the incoming data vector, calculated in terms of distance within the vector space, will adjust its vector further to the incoming one. In a Kohonen feature map also the hidden units in the close neighbourhood will be adjusted to the incoming vector more or less, depending on their distance to the unit with the best fitting vector. This way we get a vague representation of a pattern, with the best instances activating to the highest degree the central or optimal unit of a neighbourhood of units.

Kohonen feature maps, or modifications of these, are suited to represent concept learning as far as recognizing similarities and differences is concerned. The emerging conceptual map consists of a set of hidden units whereby the distance between these is geometrically measured. They are all connected with input sensor units, but they are not connected among each other, except for some neighbourhood adjustment to units most strongly activated. Each sensor unit can be understood as a dimension or micro-feature that measures the input and assigns a value to it. Concepts that are similar to each other in terms of closeness of micro-feature values. They are represented on the map each by groups of units, whereby the groups belonging to similar concepts are closer to each other than those of concepts that have less in common. The greater the similarity, the greater the overlap is between the respective patterns. Each group is concentrated around a unit best matching the most probable data for a certain concept. The two layers of the Kohonen model provide a transformation of an n-dimensional input into an m-dimensional representation of concepts where similarity and difference between concepts is expressed by geometrical distance in the m-dimensional space. Examples are mostly given for two-dimensional maps; here we have as a measure the geometrical distance on the plane, measured over all units involved in the respective pattern.

The internal similarity degree of a set of instances of a general concept A can be calculated in a way that prototypical instances determine the value to a higher degree than do the marginal cases. The similarity degree has to be calculated over the vectors of the units of the pattern that get activated most by the input instances of A. The similarity degrees of sequences of growing sets of instances for each subcategory of A converge to the internal similarity of the respective subsets of prototypical cases of the category A, as long as the prototypical cases of A are greater in number then other cases, or are very salient. Convergence can also be enforced by multiplying with a strengthening value, which expresses preference for stabilization or preference for keeping up stability. This is typical for the normative attitude with respect to previously established concepts and language use. In this way stabilization can be speeded up, or kept up longer against advers data appearing later on.

In a connectionist model we can implement the similarity measure in the following way:

The perspective **P** under which similarity is judged is analyzed as a cluster of dimensions $<d_1,...,d_n>$, or basic similarity spaces. Each dimension consist of a set of values available in the respective input sensor. In the conceptual map that builds up on the set of hidden units we get neighbourhoods of units with a prototypical value and some similar values around. The vector of values an instance s gets under perspective **P** is $<d_1(s), ..., d_n(s)>$, which per hidden unit is calculated from the input vector and the previous state of the unit in terms of previous connection strength. Then: Under perspective **P**, instances s_1 and s_2 are more similar to each other than s_1 and s_3 if and only if the overal difference between s_1 and s_2, $D_{1,2}$, is smaller than $D_{1,3}$. Hereby, $D_{1,2} =_{\mathbf{Def}} \Sigma_{i=1,...,n;j=1,...,m} (d_{ij}(s_1) - d_{ij}(s_2))$ for all dimensions d_i and all hidden units j, and $D_{1,3}$ is defined accordingly.

We can calculate an average vector from all the instances up to a point. With growing numbers of instances the average stabilizes towards a certain vector. This vector represents the internal similarity of the set of instances. If there is one prototypical type of instances of the category and if instances of that type are frequent, the average vector will be close to the vector that represents the prototype. However, this is different if there are more prototypical centers, like for the example for the use of the German word *Spiel*, which Wittgenstein discussed as an example of family-resemblance between the several typical kinds of games and plays. Here a common average would correspond to a very low internal similarity, which is not at all sufficient to define what 'ein Spiel' is. This untypical and uninformative average would be carried by a unit or a neighbourhood of units which is lowly activated for all instances. On a topological map for the concepts covered by *Spiel*, the average vector would be in the middel of a valley surrounded by a mountain chain. The mountain tops are the units with vectors which are proto-typical each for a different subcategory of games or plays. For each of these subcategories separately we form an average vector. The same holds for other polysemic complexes of concepts.

According to the dynamic theory of concept formation, a concept expressed by an expression e arises as limit towards which sequences of growing similarity sets of data, i.e. instances of use of e with respect to satisfaction situations, converge under a graded notion of set-internal similarity and contrast under a perspective. Similarity not only includes similarity caused by internal (inherent) properties of data, but also similarity by external (relational) properties, especially analogy. The perspective dependent similarity sets are called 'quasi-concepts'. For different expressions e, e', e'', they form oppositions under the guiding perspective. A concept is a limit of a converging sequence of quasi-concepts.

An expression can express several concepts, and in treating concept formation based on linguistic usage, polysemy is built into this process by its context-dependence and perspective-dependence. Polysemy arises by structuring the set of data into subsets according to different perspectives taken for selecting similarity.

Concept formation happens in at least two stages: The first stage is based on the experience of the use of utterances of an expression with respect to satisfation situations. The second is based on knowledge in terms of sets of general sentences held true, and basically consists of a structuralistic ordering of such sets of general sentences into syntagmatic and paradigmatic fields for expressions, whereby we achieve as results taxonomic orderings, definitions of features, and conceptual analyses in terms of features. Here, I shall present the first level of concept formation, the formation of experiential concepts:

We have for an expression e:

1. experiences of utterances: u_e, and 2. experiences of satisfaction situations: s_e, short: experienced satisfaction situations. We form for an expression e associated with utterances u_e sequences of growing similarity sets $S_{e,i}$ of satisfaction situation under a perspective $\mathbf{P}^i$: $S^1{}_{e,i}$, $S^2{}_{e,i}$, $S^3{}_{e,i}$, ..., whereby $S^2{}_{e,i} \supset S^1{}_{e,i}$, $S^3{}_{e,i} \supset S^2{}_{e,i}$, ...

The internal similarity degree of $S_{e,i}$: $D(S_{e,i}) > p$, whereby p is the similarity degree due to the perspective $\mathbf{P}^i$. If $S_{e,i} \subset S'_{e,i}$, then $D(S'_{e,i}) \leq D(S_{e,i})$. We form equivalence classes $|S_{e,i}|$ on the basis of identity of internal similarity degree, which form a sequence under "<". I.e., if an s is added to $S_{e,i}$ and $D(S_{e,i})$ does not decrease, then $|S_{e,i}| = |S'_{e,i}|$.

The biggest member of a **P**-harmonized sequence of equivalence classes of similarity sets is called "the quasi-concept of e with respect to the set of data".

If with growing n, $| D(S^{n+1}_{e,i}) - D(S^{n}_{e,i})| \rightarrow 0$, we say: the quasi-concept stabilizes.

A sequence of similarity sets is stabilized, i.e. defines a concept, if there is an $S^n_{e,i}$ in $S_{e,i}$ such that for all $S^m_{e,i}$, with $m > n$, $D(S^m_{e,i}) = D(S^n_{e,i})$. All these $S^m_{e,i}$ are members of the equivalence class $|S^n_{e,i}|$. They are called maximal members of the sequence. The equivalence class $|S^n_{e,i}|$ is the constructed experiential concept. It no more depends on the growths of the set of data.

A set $S^n_{e,i}$ in a sequence S **completes a concept** of $e =_{def}$

(1) For all m, with $m > n$: $D(S^m_{e,i}) = D(S^n_{e,i})$. i.e. <u>stabilisation</u>.

(2) For all $\mathbf{P}^j$, with $\mathbf{P}^j \neq \mathbf{P}^i$, and all k such that there is an $s \in S^k_{e,j}$:

$D(S^n_{e,i} \cup \{s\}) < D(S^n_{e,i})$, i.e. $\mathbf{P}^i$-harmony as distinctive in <u>polysemy</u>.

(3) For all e', with $e \neq e'$, and all k such that there is an $s \in S^k_{e',i}$:

$D(S^n_{e,i} \cup \{s\}) < D(S^n_{e,i})$, i.e. <u>opposition</u> under $\mathbf{P}^i$.

Stabilization implies that instances of satisfaction situations of e under $\mathbf{P}^i$ that would change the similarity degree are not anymore incorporated into the concept, but are considered to be marginal cases, or give rise to new concepts. A new concept within the polysemic complex of e is constructed by finding a subset $S_{e,j}$ under some other perspective $\mathbf{P}^j$ the internal similarity degree of which does decrease least when the new satisfaction situation for e is added. The subset $S_{e,j}$ generates a subconcept of an existing one if $\mathbf{P}^j$ is not new but rather a specification of a perspective used so far in the polysemic complex of e. If $\mathbf{P}^i$ is new for it, then $S_{e,i}$ gives rise to a new concept within the polysemic complex of e. We then speak of **Cross-perspectival transfer of expressions in concept formation**, which is achieved by similarity under the new perspective, i.e. by metaphoric operation, or by factual contiguity of part-whole, space, time, cause, effect, result and material relationships, i.e. by metonymic operations.

A concept can in principle be a member of the polysemic complex for an expression if, under a chosen perspective, it stands in a similarity or factual contiguity relationship to some other member of the polysemic complex.

A perspective can be reconstructed as a second order property of a certain kind, i.e. a property of properties, or a set of properties which can be expressed by predicates that are possible answers to a common question or interest, for example: What (kind of Y) is this?, What is its colour? What is its behaviour? What is its function? What about his health? How does he do economically? etc. To look at something under perspective **P**, for example looking for an activity-property, means to attend to the first order property 'activity' and see in what way it is specified in this case.

A polysemic complex is of the same type as second order properties are; it is a set of first order properties. The principles of forming these collections are metonymic and metaphoric relationships, which amount to relationships of factual contiguity and similarity, restricted by the requirement of change of perspective.

I. Recursive definition of a **POLCOMP**(e):

1. $P \in$ **POLCOMP**(e)

2. If for all situations s in which P' is realized, the expression e is taken to be satisfied by s, and there is a P with P $\in$ **POLCOMP**(e) such that metonymic(P',P) or metaphoric(P',P), then P'$\in$ **POLCOMP**(e).

The expression e used under perspective **P** then expresses some property P in the intersection of **P** and **POLCOMP**(e):

P $\cap$ **POLCOMP**(e) = {P}

Polysemy is generated in the following way:

Data:

1. Expression e is used with respect to situation s truthfully, i.e. s is referred to as a satisfaction situation of e.

2. e is used under perspective P.

3. The property that has to be assigned as being expressed by e under P with respect to s has to be eligible as a potential member of the polysemic complex of e.

Goal: Find a property P' with P' $\in$ P and P' being realized in s such that it fulfills the condition for being a member of the polysemic complex of e.

Procedure of concept construction:

I. Take the set of previous satisfaction situations for e.

II. Delineate a (new) similarity set for e under $\mathbf{P}^i$, named: $S_{e,i}$. Choose $S_{e,i}$ such that $s =_i s'$ for all $s' \in S_{e,i}$.

III. Extend that set with the new satisfaction situation s of e such that this extension obeys $\mathbf{P^i}$-harmony and opposition to other $\mathbf{P^i}$-properties, and that we can construct a sequence of growing subsets up to $S_{e,i} \cup \{s\}$ with a converging decline of the internal similarity degree. If that is not possible for $S_{e,i}$, then delineate another similarity set for e under $\mathbf{P^i}$ which satisfies these conditions and name it $S_{e,i}$.

***Result*:** The quasi-concept $S_{e,i} \cup \{s\}$, approximating a concept which is a reconstruction of a property, i.e. a concept, realized in s.

In comparing both models of concept formation, I shall discuss the role of perspectives and especially similarity or feature spaces, basic ones and those generated on the basis of social and cultural interests and on the basis of concept formation that has taken place previously. Furthermore, I shall point out how prototypically organized categories and cluster concepts are treated in both models. The result will be that connectionist types of models can be understood as subcognitive models for a dynamic philosophical theory of concept formation.

II. Concept formation and connectionist models.

The purpose of this section is to show that the dynamic theory of concept formation is germane to models of concept formation as they are developed within connectionist frameworks. Hereby it does not matter whether or not connectionist models are realistic in the sense that they model neuro-physiological processes. It rather suffices that they model some principles of the kind of structuring that occurs in real neural nets. The first principle is that a neurological net adapts to influences from the outside by growth and change of structures on the basis of incoming data. The second is, that degrees of stabilization of certain reaction patterns get explained by relationships between neuro-physiological activation states. At least these two principles are modelled in connectionist frameworks.

1. General properties of both models

There is a common motivation in connectionist approaches to concept formation and the data-oriented dynamic approach:

Firstly, concepts are stages of relative stability within a learning process that tends to converge towards a state of **stabilization**. Massive incoming data can always cause destabilization, if they are advers to the stabilization process at hand. Then a new process of stabilization gets initiated by other common regularities in the old and new data. The activation pattern in the set of hidden units that, at a certain stage in the learning process, is associated on the one hand to an expression type and on the other hand to a type of satisfaction situations may change. In the course of time different patterns may fulfill this connection service; the only constraining condition is that they make this connection between expression and satisfaction situation. There also can be changes in the type of satisfaction situation that belongs to a sentence, for example by different theoretical backgrounds that can be involved. Herewith, of course, the intermediating activation pattern changes, too.

Secondly, concept formation is **context-dependent** and depends on perspectives under which linguistic expressions are used and the accompaning situational data are perceived. Observations of the use of the respective linguistic expressions with respect to satisfaction situations are taken as the basis in forming concepts. (It cannot be answered in general, how far, under which conditions, and in what way generalization is possible across different context-dependent concepts, all expressed by the same word or phrase. This is different per case.) Connectionist systems, up to now, cannot perceive directly satisfaction situations for expressions. But in principle this should be possible. They have to use descriptions of satisfaction situations in some standardized medium, which of course presupposes the use of features in the language of description, as we know them. Or they use linguistic contexts in which the expressions occur. Hereby satisfaction situations will be represented in their conceptual value, as far as this value is linguistically documented. The second road means that they base conceptual ordering on the linguistic distribution of the expressions investigated.

Thirdly, connectionist models, like the theory presented here, take into account not only contextual dependencies, but also the dynamics and **flexibility** of lexical use. These are aspects of the process of formation and generation of concepts. Concepts, as the results of concept formation, are never more than just a more or less stable stage in this process. Dynamics belongs to the theoretical scheme of concept formation. The possibility of stabilization, but also of destablization, and of transfer of expressions onto new concepts is built into the model from the very beginning. Hereby, both approaches can explain easily how, with changing use of linguistic items, the concepts may change that are expressed. They also explain why change in use of expressions always remains a possibility.

Fourthly, both approaches do **not presuppose a set of basic concepts** or conceptual features as an innate basis of concept formation, such that all concepts would be syntactically composed of these basic ones. Rather conceptual features, i.e. concepts that function in linguistic conceptual analysis, are constructs that become available later in concept formation as certain structures that emerge in the conceptual network, and are made explicit by operations of stating hyponymies and oppositions. These are based on syntagmatic relationships (expressing general predicative and relational associations within recurring satisfaction situations) and paradigmatic ordering by appearance within the same types of contexts or the same types of satisfaction situations. The operation of substitution by analogy is an extension of the ordering that is manifest within the data: If A appears within the context B and within the context C, and D appears also within the context B, then there is a good chance that D likewise can appear within context C. The more contexts A and D have in common, the more likely it is that they also have some further contexts in common. This is equivalent to having a lot of common semantic features. Common semantic features of A and D can be defined by subsets of common contexts, i.e. by subsets of their common distribution.

Fifthly, though both approaches do not presuppose innate conceptual features, they presuppose **some innate pre-conceptual prerequisites**. Firstly, on the basis of our innate perceptual and motor abilities there are basic perspectives or dimensions of measuring incoming data. Secondly, there are perspectives formed by goals, interests, desires, orientation that come with how we are equipped for being situated in our surroundings and how we get adapted to these. These perspectives, under which perception of situations takes place, are not concepts to begin with. Rather later they might be conceptualized as higher order parts of conceptual orderings. These perspectives, then, get reconstructed as such higher order concepts, and as such they may become expressed linguistically. In conceptual analysis, which is an activity of formal education and is part of scientific and philosophical methodology, they furtheron can function as explicit feature-dimensions.

Generally, perspectives determine, or are themselves a selection of similarity spaces under which an input is analyzed. Basic perspectives, now, are the basic dimensions ('micro-features') of these similarity spaces. That a certain perspective is in use means that certain similarity spaces are activated. Its dimensions or micro-features are then the means of analysis or interpretation of the incoming data by measuring: In terms of values of these micro-features we get the relevant parts of the causal effects or impacts a situation has on our sensor and motor system. Micro-features are not concepts for the concept learner, because he cannot interprete them in a world, which still has to be constituted. Only after partial worlds have been constituted by forming objectivizing relationships between his experiences, in these worlds higher order classifications can be traced which are interpreted in our theory of concept formation as projections on the world of certain body-internal micro-features or combinations of these. For example the perspective Colour, which is not an innate concept, but an inborn measuring device for light frequencies, has as a corresponding higher order concept, the second order concept 'Colour', which is built on sets of first order colour-concepts, and is, like these, concious and understood intentionally as something in the world: In our world there is colour. However, the perspective Colour by itself is nothing more than our physiological endowment for perceiving colours. Concepts are concious and used in intentional acts of predication and reference, perspectival prerequisits need not function on that level. Only on the level of second order concepts we are aware of what colours are. This awareness is part of our knowledge about colours. Only via such a correspondence between micro-features (basic perspectives, dimensions) and second order classifications in the world can micro-features, or combinations of these, indirectly get an interpretation in the world, and only hereby they can be assigned the status of concepts. Mostly, perspectives are reconstructed at the level of concepts by theoretical generalizations, or by meta-theoretical kinds of linguistic analysis.

2. An outline of connectionist concept formation

A sensor and motor system consists of the sensory and motoric organs and their connections with certain parts of the brain, including these parts themselves, and interrelations between these parts. The effects on the sensor system, which are partly conditional on activities of the motor system, can be of different kind and strength. This strength can be coded by numerical values for the degrees of impact of the causal effects for each micro-feature involved in the similarity space. Thus, a similarity space defines a specific type of vector consisting of an ordered series of micro-features or dimensions. Each feature is to some degree activated by the incoming stimulus. These degrees are the numerical values of the vector that then represents the input. There also may be vector types which consist for some part of registered features of motor activity. These can form the conditional context for other sensory features which form the other part of the vector. Anyhow, the data are n-place vectors in n-dimensional spaces.

In the world, there are causal chains between states and processes which are realizations of relational as well as non-relational properties. They have an impact on our sensory organs with their perspective-dependent similarity spaces. By such impact, experiential predecessors of formal concepts are effected. They are built up by a cumulative effect of recurrent similar input. These evolving experiential predecessors of formal concepts are patterns, which are, as just described, causally connected by perception and its motoric conditioning to daily situations, routines, actions and things. But they have no similarity with these, other than by parallels in internal and external relationships. The causal relationship between recurrent properties and relationships in the outside world and properties of the internal patterns gives rise to structural parallels: Noticeable structures in the world have parallel structures in the pattern. In a theory of concept formation, these patterns can be understood as concepts only if they are interpreted by the theory as representations of the situations and things with whom they were causally connected in the process of being formed, and with which they are causally connected again in new situations of action and perception. For the concious concept forming individual these patterns themselves are nothing, anyway not concepts. They are sub-cognitive, underlying our intuition of cognitive structures. I also would not want to call our experience connected to such an activated neural pattern a concept. This experience is an experience one has in activating a certain concept. The concept is a class of these experiences, corresponding to the pattern as a class of partly similar, i.e. overlapping, activations. The experiences, and also our concepts as orderings of these experiences, are understood as being of or about things in the world, not of the pattern activations. Nevertheless, intersection, inclusion, and exclusion relationships between activation patterns parallel such relationships between (extensions of) properties in the world.

The two layers of the Kohonen model provide a transformation of an n-dimensional input into an m-dimensional representation of concepts where similarity and difference between concepts is expressed by geometrical distance in the m-dimensional space. This second map can be connetd with a third layer of units, which then each get an m-dimensional vector as input from the second layer. Input that goes through a system of two layers onto a third layer is already selected, or filtered, not only by the input sensors, but also by the concepts established on the second layer. Thus the second layer 'interprets' incoming data and thus influences the resulting activations on the third layer. Examples are mostly given for two-dimensional maps of m units; here we have as a measure of closeness of two concepts in terms of similarity or factual contiguity in syntagmatic relationships the geometrical distance on the plane, measured over all units involved in the respective pattern.

The representations on the two- or more-dimensional map by themselves are not very informative, if we just look at them. They merely show very rough structural relationships between concepts. What the concepts are qua content only becomes apparent in the connections of the units of the map to the sensor units and their weights after the learning process. What these weights mean in terms of conceptual content and not just in terms of vague relationships of similarity between concepts can only be known by studying them in relation to the input data. Since the input data for us are always experienced input data, we can make sense of the weights only in terms of what is part of our mind. This, on the other hand, is according to the present theory itself already expression of these constellations of sensor units and their connection to the conceptual map, including the weights and the causal connection via the sensors to something outside. Though this all is causally primary to our concious experiences, these are conceptually primary in our identification and understanding of the causal factors involved.

There is no content in connectionist representations by themselves. They are related to content only by being causally related to input data. Depending on the (random) state of the hidden map to begin with, or on its state after having learned some concepts, learning a new concept results in patterns located quite differently on the map and being of different shape, depending on input experiences in relationship to other concepts. There is no fixed place and shape for a pattern corresponding to a concept. We acquire knowledge of the input by placing our experiences of the input data into classifying and objectivizing relationships with other experiences. Our experiences, which all together make up our mind, depend on incoming stimuli as well as on the analysis provided of the stimuli by the input layer, i.e. the sensors, and by the hidden layers, in which they are placed into orders of classification and contiguity. Since the hidden layers get modified in the course of making experiences, also our experiences of incoming stimuli will change. Thus one and the same type of stimulus can trigger different types of experiences, due to the state of the hidden layers involved in analysis. Nevertheless, there evolve more or less stable relationships between patterns, corresponding to such relationships between concepts. Our feeling that we use or activate a certain concept might best be understood as an awareness of the respective pattern's activation. But, strictly speaking, the pattern cannot be the concept. The concept is part of cognition, and on that level we can think of it as an equivalence class of conceptually complete sets of experiences of satisfaction situations. Each such set is an equally fine representation of the concept. For many concepts we can imagine instances. But an instance is not the concept; rather it is an image which exemplifies the concept because it fits into the concept. In a derived sense we can say that we imagine the concept if we image one or more instances with the additional thought that more instances can be imagined. In this way we reconstruct a representation of the concept in the sense explained above. The invariance we feel in all the instances is due to the internal similarity measure of the set of instances reached in the stabilization of the concept. The connectionist pattern, on the other hand, can be thought of as modelling the neuro-physiological basis of the respective concept. By the neuro-physiological pattern, built up in the learning process, it is measured whether an instance fits into the concept. It is the realization of the similarity measure towards which the process of stabilization has converged.

3. Different types of conceptual maps

A modification of the simple Kohonen feature map is Kohonen's (1991) Hypermap. It can be used to distinguish concepts and their contexts. In Scholtes 1993 we find an example for classifying terms by their contexts. This indirectly provides a conceptual map representation of closeness of concepts as far as conceptual closeness is expressed by the fact that the terms occur in certain contexts, and not in others. To the map, vectors are fed which consist of two parts, one part for the concept expressing terms, and one for the linguistic context-type in which these terms occur. Each vector has the bipartite form $X_p \oplus Y_c$. The map can first learn the context-types separately, by feeding vectors with constant X_p and varying input Y_c. Thus the map gets pre-organized according to context-types. Then the whole vectors are fed into the map. Within the regions of certain context-types different neighbourhoods are built up that represent the concepts that co-occur with the respective context-types. In this way, context-types can be learned and depending on these, terms can be learned as being more or less compatible with certain context-types. The map forms paradigms of terms with respect to context-types. Since common context-types are indicators for common features in concepts expressed by the terms, the order that is formed on the map will represent hyponymy and opposition between concepts in a more natural way than can be done in a connectionist model designed on the basis of features, which are already known from linguistic analysis and which presuppose much of what should be learned. In fact, by letting certain known features function as micro-features in the design of our model, we impute on the system already much of what we know about the classifications, while our aim is to understand how the system can learn the classifications and conceptual features to begin with. Scholtes 1993 provides a method of concept representation by means of contexts. Based on his procedure, we can see a representation of conceptual features by orderings on the sets of contexts or context-types. This parallels the second level of the dynamic theory of concept formation, on which linguistically explicated concepts are formed.

In Scholtes 1993 we also find a two-layered recurrent model, building on Kangas two-layered Kohonen maps with recurrent fibres, where units of the first can fire to the ones of the second layer. By the recurrent fibres a vector representing the activation of the first map due to a left context of a word can be fed back into the first layer in concatenation with the vector for the word itself. This is a way to deal with sequences, here of two words in a row. At the same time activation of the first layer is fed into the second layer. In the first layer map we get, to some degree, a kind of syntactic categorization of the words, as far as it is expressed by sameness or difference of their left contexts in sentences of a certain data set. After the first layer has organized itself, we get on the second layer an ordering of the phrases consisting of a word together with its left context. In principle, the first map forms a paradigmatic ordering on the basis of similarity defined by sameness of context, while the second map forms a syntagmatic ordering, i.e. an order by contiguity together with a context.

A different kind of map is one with hidden units that are able to fire to other units of the same or of another layer. A special type of hidden layer is one in which its units are interconnected. Although every new situation, in principle, stimulates all units of the hidden layer more or less via the connections from the sensor units, only those that have a vector of values that does not deviate much from the incoming vector of the stimulation will be triggered strongly enough to exceed a certain threshhold of activation. Thereby they will send an impuls to other (possibly output) units. Depending on whether intra-connections of the layer are restricted to neighbours on the plane or exist also directly between geometrically distant units, a triggered unit can fire and activate other units in the neighbourhood or far away. Interconnections can be strengthened by being activated more often or more intensively. If two units are repeatedly activated at the same time, their interconnection can become stronger such that also in the event that only one of them gets activated directly from an input, the other unit can get some of the spreading activation via the interconnecting fibre. Thus, if a system often has experienced that a dog barks, it can, when seeing a dog that does not bark, expect that it might bark. This is the typical way of associating one event with another via contiguity. Also paradigmatic association works like this: one easily gets from dogs to cats if they are close together on a map with neighbourhood interconnections.

In hidden layers where the units are also connected among each other, closeness of concepts may not only be represented by geometrical distance, which is based on neighbourhood adjustment, but also by degree of strength of the interconnections between the units. Those concepts that are activated together more often than others, form stronger connections. Thus typical contextual associations, so-called 'lexical solidarities' and other syntagmatic relationships, can be represented additionally to concept-similarity. We can associate, for example, 'dog' with 'bark', 'dog' with 'run', 'frog' with 'leap' and with 'swim', 'horse' with 'ride', and so on. But also cases like 'men' and 'women', 'women' and 'children' will be associated, which appear in phrases of recurrent use. Further, an individual concept can be represented by an activation pattern on a net of units which stands for some typical properties of the individual. If some of the units get activated strongly enough by some input one associates the other typical properties of the individual, or the whole individual as far as it is represented by an activation pattern on the whole net.

Another modification of a conceptual map would be, what I here call a 'compound' map. Here the relationship between the conceptual map and sensor units has to be designed differently, such that we get a compound map consisting of several Kohonen maps on a plane. There may be a similarity space of sensors for colour and form features, and there may be another for movement and behaviour features. Each of these spaces may be connected to different parts of the compound map of hidden units. In this way, we get a division on the map between concepts expressed by verbs (of movement and behaviour) on one part, which is more strongly connected with the movement sensors, and concepts expressed by nouns for things on another part of the map. Animals would be represented inbetween. We can imitate this process of constructing a conceptual map on the linguistic level by presenting short actor-action sentences (up to a certain length) to the map in form of a vector with a numerical code for the words of the sentences. To induce such a division faster, the map can first be trained to recognize the verb contexts of the nominal terms, and than it can be trained to the whole sentences. The connections among the units of the map are trained by feeding these sentences.

A conceptual map can be connected to other layers of units. Since also the connections of the hidden units to possible output units (for a classification task) or to other related units can be strengthened in the process of learning, these patterns expressed by vectors in a group of hidden units, can be associated with other units containing vectors for patterns related to the first by recurrently occuring factual relationsships, or by similarity. Likewise the connections between a layer of conceptual patterns and a layer of units which contain vectors for expressions associated with the conceptual pattern can be strengthened. Thus linguistic patterns can be associated with conceptual patterns. -- Of course, linguistic patterns are also special kinds of conceptual patterns, namely patterns of linguistic forms realized within linguistic utterances. Each linguistic expression or linguistic form is a concept of a class of utterances.

A two-layered model of concept formation has been worked out by Schyns (1991). He first provides for non-linguistic concept formation on a first conceptual map, taking into account that children have already formed many concepts before they have learned lexical items. Secondly, a lexical map is formed in the same way, containing only representations of linguistic expressions. Thirdly, by association a connection is formed between these two layers, whereby pre-linguistic concepts are associated to lexical items, and cases of overgeneralization get corrected in this supervised module. The first two modules are not supervised.

By starting out with language-independent concept acquisition as an unsupervised learning process, Schyns (1991) aims at a psychologically real model of concept formation. This is not in contradiction with the model of concept formation proposed here, where concept formation is reconstructed not as a psychological process; rather a philosophical and logical reconstruction is intended. But if we, for the moment, think in terms of psychological implementation, we can state: If pre-linguistic concept formation is already available, it would just become integrated into the phase of linguistically guided concept formation and would speed it up. On the other hand, much correction would then be necessary because of over- or under-extension of pre-linguistic concepts, compared with the use of the respective linguistic expressions and the concepts defined on the basis of this use. It does not make any difference for the principles and operations of concept formation, whether we start with pre-linguistic concepts and adjust them later to linguistic usage, or take linguistic usage as a starting point right away. The same principles that are used in linguistically guided concept formation are also used in pre-linguistic concept formation, though there with fewer restrictions. But the omission of the restrictions at this point in the acquisition process requires that they are applied later on, anyway. In a formal theory of concept formation as lexicon acquisition, they can better be integrated into the theory right from the beginning. Because here the issue is not just the formation of concepts, but of linguistically expressed concepts, i.e. meanings; and meaning of an expression is based on its use.

I disagree with Schyns (1991) on that he treats learning linguistic expressions in a module totally separated from the use of the expressions. This is not favourable for a data-oriented theory of linguistic parsing. It also is of no help for a general theory of syntax, which we do not want to get uninteresting or even circular because of very heavy theoretical assumptions, namely assumptions which are justified by nothing else than by making a theory of language work which contains an autonomous syntax. The semantics of sentences, as it appears in the parallel orderings on sets of their satisfaction situations and constellations, helps greatly in developing something like syntactic structures. Syntactic structure for a large part corresponds with the ways of synthesis between concepts in sentences by which constraints on basic situations are constructed such that they addable salva stabilitate to the intersections of positively involved lexically and syntactic-morphologically expressed concepts. The elimination of semantic information as input into syntax acquisition has given rise to unnecessarily involved theories of innate linguistic universals or principles that have to guide learning of a natural language. These unfortunate developments are the necessary result, if the most important source of integrated input data, namely semantic input together with sentence input, is cut off in linguistic theory formation by creating separate modules of syntax, of semantics, and of cognition. We better can think of models in which maps on which some pre-linguistically established conceptual orderings and some previously established phonological and expression orderings function such that by them data of syntactic-semantic language acquisition get pre-organized in terms of strengthening and weakening connections.

In a distributive representation of a concept as an activation pattern at some stage, the role of a sign that expresses the concept is especially important. This is so because the activation pattern has no fixed boundaries. It rather stands per instance of contextual use in a connection with other patterns by overlap or simple extension. A boundary, which is required for every discrete representation, can only be drawn by repetition across several contexts and by, at the same time, connecting a single sign or expression with that part of the pattern which shows up repeatedly. For each instance, the association with the sign, previously established in learning, can activate the part of the pattern that, at a certain stage, represents the identity of the concept: Starting with the activation of the sign, the activation goes via the associative connection to the part on the conceptual map that represents the concept. The discreteness introduced by the sign is necessary for delineating logical implications which are not primarily syntactic inferences by operations on signs, but are relationships within the contents, if in these parts can be delineated. A (syntactic) rule of implication between two signs, for example a meaning-postulate in the syntax of realistic formal semantics, is reflected in the cognitive semantic relationship of inclusion between delineated parts of patterns. The rule can thus be induced from out the connectionist semantic background, rather than being learned or programmed into the syntax independently.

4. Perspectives as similarity spaces for similarity measures

Input data can be analysed under different perspectives, which means that subspaces of the overall available similarity space can be activated. Thus, spaces filter out information which is relevant under a pre-established perspective. The perspective then consists in a certain frame for measuring input activation, whereby something is perceived as far as it is represented by certain activation values for the micro-features, or dimensions, which constitute the frame, or similarity space. The vectors, consisting of sequences of numerical values, are interpreted as induced by the causal connection of external stimuli to the similarity space, by which the incoming data are canalized into micro-feature values. Measurement by micro-features provides a numerical value in the respective dimension of the vector, which as a whole represents the instance.

We can imagine that similarity spaces are pre-ordered according to the capability of our sensory organs, such that we have basic subspaces and others that are formed by combining a selection of these under certain perspectives. We can imagine that vectors per basic similarity space are locally distributed within sets of hidden units -- i.e. groups of units are accessible for a certain kind of vectors --, while combinations of different kinds of vectors formed by different kinds of sensory and motoric input are globally distributed over groups of groups of units. In such a model, the colour sensors, for example, would be connected with the units of a certain group on a map, the form-sensors with the units of another group, etc. Thus we would have short vectors of basic micro-feature values in different units, for example one type vector for colours in a group of units, one type of vector for form in another group, one type of vector for kinds of movement in a third group, etc. In such a model, groups of vectors of features belonging to a single basic space are locally group-distributed such that an incoming vector hits every unit in the respective area. Per unit, its value together with the previous state of the unit (i.e. it previous connection weights) results in a new state of the unit represented by the vector of the now adjusted weights of its connections to the input units. Combinations of these vectors, which will be relevant for building up the patterns of things and situations, are fully distributed across one or more sub-maps, i.e. areas or groups of hidden units.

In such a compound model there are resultant basic vectors of values of micro-features, one vector in each unit. The units get their input by each being connected to every sensor unit of a basic similarity space which is constituted by certain input organs. Such a resultant vector itself may function further as a value of a 'higher order' micro-feature in the combination of basic similarity spaces. Thus we have basic vector types defined by basic similarity spaces, and we have matrix types defined by combinations of vector types formed in these basic spaces. A matrix consists of columns, whereby each column is a basic space vector. The micro-features, or dimensions, of the matrix type are then the types of basic space vectors. The matrix is in fact a higher order vector, i.e. a vector the dimensions (micro-features) of which are the basic space vectors which make up the columns of the matrix. The basic space vectors are distributed within a local net of units of a sub-area, and the matrixes, or vectors of vectors, are distributed within the whole net, consisting of the sub-areas together.

If we additionally assume that the hidden units in the net are also connected with each other, then we get greater connection strength between certain units by feeding in similar matrices. For example, we can have connections between units of different sub-areas of the map, which were connected with different basic similarity spaces. Then resultant vectors in one area can influence the ones in another area by association with growing strength. It can then happen that certain form and movement features can trigger the expectation of certain colour features, if the association has grown strong enough. The units with basic space vectors which co-occur as columns in certain recurrent matrices will form a stronger connection than others. All the units involved together in such an interconnection are the distributed representation of the concept that is a probability based generalization of the matrixes formed by input into spaces of spaces. Certain subsets of columns, i.e. subsets of units with basic space vectors, are then representations of parts of the concepts. These can be common features of several related concepts. This means, that within the distributed representation of concepts we get an overlap, inclusion, and complement structure of features on this level, where each complex feature is represented by a Kohonen neighbourhood of units of similar basic vectors. In this way, taxonomic relationships appear in such a connectionist model. Such concept representations are something like Hofstadter's (1982) 'living symbols', which are the substratum or the contents on which our intuitions about logical and especially taxonomic relationships are based. On this level, implication is based on inclusion, contrariness on exclusion, similarity in terms of common properties and relationships on intersection. Like in the set-theoretic relationships in Formal Semantics, these logical relationships are extensionally realized and become available by directly addressing concepts in their extensional realization as activation regions in connectionist models. No rules of inferences are necessary on the 'semantic' levels of extensional representation, not on the level of activation patterns, not on the level of concepts, and not on the world level of realistic formal semantics. On the syntactic level of expressions, on the other hand, we need rules of inference and meaning postulates, which have to be proven correct (and preferably complete) with respect to semantic models, in which they have to be grounded.

Imagine now that we have such higher order, or complex, micro-features. These are micro-feature values that are themselves basic vectors. These vectors consist of dimensions, for example the lower order micro-features 'form', 'weight', 'colour'. Then in the conceptual net, the representation of an iron bar, for example, will be distributed over the net. The micro-feature values for 'long form' and 'heavy in my hand' are situated distributively within different areas of the net. When a net learns the pattern of an iron bar, the connections between units of these and some other micro-feature values typical for iron bars will be strengthened. If then an iron bar is presented they will be strongly activated together, even if at an instance one of the features is not directly perceivable, for example the metallic glance of the material, which becomes evident when the rust or paint coat of the iron bar is scrabbed off. Likewise, the distributed representation of the expression *iron bar* will be associated with all these higher type micro-feature values, which are distributed over the conceptual net in the following way: A group of units are activated which contain vectors for the micro feature value 'long form'. They are in the neighbourhood of the optimal unit for this feature. Also a neighbourhood of units around the best for the micro-feature value 'heavy in my hand' is activated, and the same for further micro-feature values involved in 'iron bar'. In this way the conceptual pattern for 'iron bar' is activated and since it has a strong connection with the expression *iron bar*, also the visual and auditory pattern for this expression on the expression net is activated. The other way around, when one of these patterns for the expression is perceptually activated, then also the pattern of an iron bar gets activated via the strong associative connection between the layer of expression representation and the layer of conceptual representation.

Summarizing, we have in a connectionist model with compound maps three levels of representation: Firstly, we have basic space micro features. The sensor units, which accept the input, represent the micro features. From the sensor units, where each unit is responsible for measuring a micro-feature, there are sent to the hidden layer units, per instance of presented data, the values for the respective micro-features. Thus we have at least as many sensor units as we have basic micro-features. Secondly, in the learning process, units of the so-called hidden layer get the weights of their connections to the sensor units adjusted according to the connectionist learning process described above. This result can be expressed by a vector associated to each unit of the hidden layer, at a certain point of time in the learning process. Third, in a compound map these vectors are the second level micro-features that figure in the distributed representation of concepts on the compound hidden layer. The concepts can be expressed by vectors of basic vectors, either by an n-fold concatenation of vectors into a single vector, or by matrices of n-vectors as colums. The basic space vectors are represented distributively only over a sub-map, the compund vectors or matrices themselves are represented distributively over the whole compound map. On the compound hidden layer we have thus two levels of representation, a level of distributed sub-net representation, each restricted to a sub-map, and a level of overall distributed representation which is structured by these sub-nets.

Connectionist models of this kind are a suitable sub-cognitive background for the dynamic theory of concept formation. They provide ways of representing dynamic concepts, or 'living symbols' in the sense of Hofstadter (1982/85). They explain how the association between expressions and their satisfaction situations can be implemented. They also explain why we can assume similarity measures such that we can speak of similarity degrees and of sequences of converging similarity degrees in the formation of experiential concepts. Under the basic similarity spaces there is, for recurrent input, convergence in terms of weight of connections of hidden units with the sensory units. The convergence goes towards the units with the best vector in a neigbourhood of more or less similar vectors that are connected to a certain similarity space. The best (compound) vector represents a prototypical case of input. It is central in the neighbourhood of units with vectors which represent less typical items of the same category. Under combined similarity spaces we have convergence towards patterns of units in the hidden layer, distributed over the areas of the compound map, which are each connected with a different basic similarity space that is part of the combination. Coherence of a concept can be calculated in terms of strength or weights of the interconnections between its best hidden layer units in the compound map, whereby these weights are increased by co-occurent activation of the units that form the pattern. In a connectionist network we can account for prototypicality within a category and for a measure of internal similarity in a category such that frequent, prototypical instances have more influence on the convergence of the similarity degree than seldom or marginal instances.

A similarity space, which structures the input, consists of several dimensions (micro-features), for example for objects these might be the dimensions of colour, measuring the frequencies of lightwaves, the space of quantity, measuring in the dimensions of weight and spatial extension, and the space of forms, measuring in different form dimensions. They get combined in the conceptual map. For colour alone, the higher the density of a colour is, the more its value will be strengthened. This means that instances with the highest density, i.e. the most salient ones, have a greater influence on forming the representation of a certain colour, for example of Red. A red with low density, which contains a lot of white, has mixed frequences in it. This would cause units around the unit representing the most saturated red to be activated more often together with instances of other colours. In a certain natural surrounding of concept formation a not very intensive colour red can be the prototypical case according to frequency of occurrence, while the value which would represent the strongest red will not be represented in full strength in the conceptual map. We then would have a prototype with respect to salience and one or more prototypes with respect to frequency of occurrence. Ideally, the representations are head-shaped around the units with the prototypical vectors, i.e. the vectors best fitting the sensory impact of the typical examples. But if there are several prototypes to a category the representation of the category is more shaped like a group of mountain tops. Thus the colour Red can be represented by a group of different reds which are about equally frequently manifested, and for which designers and other specialists in colours might have different codes or names.

In the above model, a general concept A that has to be analyzed by means of several basic similarity spaces at once, will be represented distributively by a pattern on a compound map, which is divided into different areas or groups of units, which are each connected with a different basic similarity space constituted by a different sort of sensors. In such a pattern for A there is formed a prototypical representation by re-inforcing most the connections between those units that are activated most frequently together. These connections represent the internal coherence of the concept, and induce by triggering one part, strongly activated by one or more typical features, also the activation of the typical other features that are bound to it by high probability of co-occurence. This associative feature activation is strongest for the prototypes and thus for them predictions of other features on the basis of one or more directly activated ones is most probable. Because of this we predict for (typical) birds that they can fly, while we do not for untypical ones. We can also get several proto-typical patterns that deviate from each other in the degree to which one or the other feature is actualized, if the frequency of occurence is almost evenly distributed between subsets of instances. Then we would have a family of prototypes for one category instead of a single prototype. These define subcategories within the larger category, whereby an instance belongs to the subcategory to which it is closest. The subcategories for

German *Spiel* (Engl. *play, game*) would be an example. Vagueness is built in, because something might belong to one (sub)category to some degree, and to some other (sub)category to another degree, or even to the same degree.

Summarizing, we conclude that connectionist models provide a subcognitive modelling of the notion of perspective. The simplest perspectives are the input sensors. Each input neuron can measure a certain aspect of the input and give a value to it. In this way it selects or filters the input. Such a sensor neuron, or group of neurons of the same kind, constitutes a basic dimension (micro-feature), e.g. Colour, by measuring the frequences of incoming light. Each basic dimension is a perspective. And each space built up from perspectives as co-ordinates or dimensions, again is a perspective. The simplest kind of complex perspective is a set of input sensors. Further we can build sets of such sets. These then give rise to sets of input vectors in distributed representations on the conceptual net, or to long vectors consisting of combinations of vectors in local area representations on a non-compound net, which is part of the whole compound net. In this way, perspectives can be built up recursively. Each complex perspective is thus a net of neurons that, by measuring against its own state, analyzes the input into a number of aspects, whereby the number of basic aspects of analysis is equal to the number of neurons. Thereby each basic aspect is a dimension of values. Though the basic dimensions are preset, and may be constructed such that they get adapted to input in usage, the combinations can be preset or selected, as certain selections are learned as valuable by previous experience. What is valuable or relevant depends on interests and desires. The perspectives select from the input. The thus filtered input is transmitted as a set of vectors into the hidden layer, the conceptual map. On that map we find, after a converging process of analysis of growing sets of input, similarity and opposition of concepts such as they appear under the perspective in question. These internal map may further function as a secondary, though soft, selector for input that is already pre-filtered by the sensory units. The output of such a map may be a hereby pre-organized input to another layer of maps, or to motoric facilities.

5. Polysemic complexes in Kohonen-models

Also polysemic complexes can appear in conceptual maps. They are held together on a compound conceptual map by some similarity or by contextual association, coded in terms of connection strength between units on the map, and by being associated by connection strength with a single expression on the layer of expression representation. This latter association delineates a pattern on the conceptual map. If we see, for example, a tree, we have an activation of a whole pattern of the visualized situation and many specific features of this tree. The sign *tree*, if associated with the respective perceptional activation on the conceptual map delineates in this whole activation pattern the tree-specific features by way of high-lighting these by the connection strength with the sign perception. The context in which a polysemic expression appears activates a certain part of the conceptual map, directly or, if linguistically explicated, via its expression. This context activation pattern is associatively stronger connected to some part of the polysemic complex than to its other parts. In this way, the context selects a part of the polysemic complex and hereby the context dependent meaning of the sign is activated.

There are at least two ways to represent polysemic complexes in Kohonen-models. The first is to use a hypermap (Kohonen 1991) per polysemic complex. The input vectors have two main parts, one for the concept and one for the perspective. To represent the concept we may use short sentences in which the word that expresses the concept is the grammatical subject part. Each such sentence consists of the concept expression and its context within this sentence. This means that the concept part itself consists of two subparts. The perspective, then, is the third part of the input vector. The perspectives can be expressed by a short question. The map is first trained to the different perspectives, by keeping the other parts constant, then to the sentential contexts by keeping only the concept expression constant, and last the whole texts each consisting of a perspective-introducing question and a sentence are fed into the map. If one finds suitable values for the different constants in the mathematical formulas that organize the learning process, one should end up with a map that shows per perspective the different concepts under it.

The other way is to have two layers of Kohonen maps, one for the linguistic expressions and another layer which is a compound of Kohonen-maps, for each perspective one. Then on each perspective specific map we find the concepts under this perspective. The linguistic association has to be learned between the representations on the linguistic map and the compound conceptual map. In this way the representation of an expression *e* will be associated by connection strength with all the representations of concepts expressed by *e* under the different perspectives. The part of the polysemic complex which has strongest connection values to the expression *e* on the expression map is the meaning of *e* that is chosen by default, i.e. as long as the context of use of *e* does not require otherwise. This is the case if the context induces the default perspective, i.e. the most used one. Per polysemic expression and perspective, under which the expression is used, we get the concept expressed by intersecting the polysemic complex pattern with the perspective pattern (cf. the parallel in the logical construction of properties under perspectives in Bartsch 1987). Intersecting these two in connectionist terms just means that the part is selected that is mostly activated by the polysemic complex pattern and by the perspective pattern together.

We can, to some extent, answer the question, which other perspectives are possible, given that *e* is already used in perspective **P**. The new perspective must be different from any perspective that would correspond to a hypernym in the taxonomy under **P**. This is so because otherwise we would get an overgeneralization of the concept expressed by *e* under **P**. As far as the discussion in this chapter goes, we see perspectives established in the input system, for basic perspectives in the set of input-sensors, whereby groups of sensor-units are connected to the hidden units in a certain area of the compound map.

6. Generalization

Generalization by abstracting a common set of features from all the different context-dependent uses is mostly not possible, and certainly cannot be taken for granted. Nevertheless, in the history of Linguistics and of Philosophy the so-called Humboldt Principle 'one sign - one meaning' has been assumed as being true or as being at least a guiding methodological principle in searching for a single 'underlying' meaning also in cases of polysemy. Furthermore, all traditional logic is based on an unambiguous vocabulary, necessary to secure correct inferences. By integrating context-dependence of meaning into formal languages these can be a better tool for modelling the semantic properties of natural languages, whereby inferences can be handled in a context-dependent way (cf. Sag 1981, Bartsch 1986). Depending on how far the context-types are distant from each other, there are often no common features which would constitute a common, context-independent concept. Or the common part is so weak that it makes up a concept which does not stand in the required oppositions to other related and lexically expressed concepts. For example, the generalization over all context-types of the verb *run*, we gives an empty set of common features; if we just take those in which some movement is involved, we get not more than just 'movement' as a generalization over these concepts, which is much to wide because it does not form the oppositions to 'go', 'walk', 'leap', etc. Therefore it is not specifically connected to the expression *run* but also to its oppositions, and thus it is not one of the concepts expressed by *run*. The same is true of the verb *open*. This verb expresses several context-dependent concepts. This is evident from the oddity of gapping across these contexts, as in *John opened the can and Fred the door,* or *John opend the book and Fred the can.* These cases are not quite as bad as *John opened the door and Mary the nut* or *John opened the door and Fred* *the meeting*. Still, sufficient conceptual identity is missing for making these constructions semantically acceptable. Of course, opening a can and opening a book have more in common that opening a can and opening a door, and certainly more than opening a meeting and opening a door. But what they do have in common seems to be difficult to express linguistically by the use of other expressions than *open* itself. Nevertheless we feel that there is something in common which makes it natural to use *open* in both cases. However, what is common is not a generalized, context-independent concept 'to open'. Such a concept, the lexical meaning, is just a myth. Rather the whole polysemic complex of *open*-concepts is the lexical meaning, whereby some of these concepts are closer to each other than others. But the features by which they are close to each other are not sufficient to determine something like the lexical meaning of the word.

We can think in terms of a connectionist representation of these uses of *open* in a kind of Kohonen map, especially a hypermap (Kohonen 1991), where the different context-dependent concepts expressed by *open* can be distinguished by their appearance in the different areas of different kinds of contexts. There the associative bond between the different uses will appear as closeness on a conceptual map, based on common micro-feature values, though these values might be of such a kind and make up such a combination that we have no words for these. And also an adequate elaborate description of what seems to be in common, abstracting from the contexts, seems to be impossible. Since we cannot think of what 'open' amounts to without referring to something which exemplifies the property, we would need some thing which has those properties of a door, a can, a nut, and a book all at the same time and which permits to apply the concept 'open' in all these cases. To open such a thing would be the generalization for these three cases. But we cannot imagine any such thing. This is different for a door and a window; here we can imagine a kind of thing that can be a door, when made high enough and placed on the ground level, and it can be a window when placed above the ground. Both we open in approximately the same way. In the concept 'open

X' we can replace X by a door or a window. X represents what door and window have in common. This property X is something which makes that everything of kind X can be opened in the same way. For book and can, what is common is just that it is a material thing. Here a generalized 'open' would not be more than 'to make two parts distinguishable by a spatial cleavage'. But that is much to broad because it extends across oppositions: it would also hold for cut (partly) apart, for tear apart, etc. Though such a generalization is automatically included in concept formation it is not distinctive for the concepts expressed by *open*. Data based concept formation and connectionist models of concept formation can both account for context-dependent concepts, and the relatedness between concepts expressed by the same word, i.e. they can account for polysemy; and they also can account for the few cases where a distinctive generalization is possible. As far as for an expression a conceptual generalization is possible which preserves the oppositions to other related concepts, it is automatically included by these models of concept formation.

I now shall illustrate generalization in a case in which it is clearly possible with respect to an expression, and I shall indicate what kind of operations are involved:

Starting with satisfaction situations of sentences which contain, for example, the word *horse* we get different situational concepts of a horse:

a horse as running, a horse as eating, a horse as drinking, a horse as standing, a horse as lying, etc.

In order to get the concept of horses as individuals we have to unite all the situational 'horse'-concepts into possible life-histories of horses, i.e. into space-time-and causally connected sequences of horse-situations. These sequences are thought of as being parts of continuants. They approximate an individual concept of a horse. Similarly, in order to construct the action concept of, for example, beating a space-time and causal connection has to be made across the situations of beating, of being beaten and of observing someone beating someone. It is evident that a generalization across different situational, i.e. context-dependent, horse-concepts or beat-concepts is not a matter of abstraction, but rather a matter of firstly establishing factual and possible historical connections between situations in which the same expression is used. When these connections have been established and therewith individuals, c.q. actions, have been constructed, an abstraction over these constructs may lead to a generalization by which the general concept of a horse, c.q. the concept of beating, is achieved. What this abstraction means in data oriented concept formation is that we can scan through a variety of 'horse'-situations bound together in possible life-histories. There is no abstract horse constructed. It is more that we are able to construct possible partial horse-individuals from the set of 'horse'-situations we have come across and will come across. The data might be concentrated around something like a proto-typical horse, which would be no more than a typical collection of typical 'horse'-situations in a partial possible 'horse'-life-history. The basic perspectives guiding this mode of construction are, when they are reconstructed on the conceptual level, what Kant in his "Critique of Pure Reason" saw as general schemata which play a role in making up specific, i.e. empirically grounded, schemata: General categories of quantity, quality, relation and modality are applied in synthesis to space-time parts not yet filled with specific empirical material in order to construct our general notions or schemata of objects, events, and states. These serve in integrating, for example, the different empirical 'horse'-impressions or 'horse'-situations into an individual concept of a horse. The generalization over different individual horse concepts, then, results into the general concept, or empirical schema, of a horse.

Also when we learn quality concepts, for example 'blue' we start with context-dependent, situational 'blue'-concepts:

blue of the sky, blue of the sea, blue of the corn-flower, blue of sample X, blue of sample Y, etc.

Here the generalization is simpler. We do not need to construct individuals, or actions, first. Generalization consists in forming a concept of a proto-typical blue, i.e. a blue without other colours in it, except for white as the neutral colour mix. By this we get a scale from light blue to a saturated blue. This saturated blue could be the proto-typical blue.

The following requirements are necessary for the possibility of generalization: It must be possible to abstract from features of the context-dependent situational concepts of, for example, the use of *horse* in satisfaction situations. The features that are abstracted from, in the sense that they do not occur in all these situation, are the ones that are not central or essential for something being called *a horse*. The central features that are kept are 1. the ones of the hypernymic concept introduced by the perspective of classification, and 2. the specific features that play a role in defining the oppositions under the perspective. The first ones are, for example, the animal features under the natural kind perspective 'kind of animal'; the second are the ones that distinguish horses from the other kinds of animals. The abstraction is not a separate operation or step one has to do somehow. It is automatically included in the very process of ordering our data such that concept-formation takes place. It is not possible in a distinctive way if hereby central features get cancelled out, i.e. if hypernomy induced by the perspective is violated, or if necessary specificity with respect to oppositions is violated. The first violation would lead to use of the expression under another perspective, i.e. to a metaphoric or metonymic use, and therewith would give rise to polysemy; the second violation would lead to overextension.

If across the instances of use of an expression in satisfaction situations distinctive generalization is not possible, one has not established a single concept, but a polysemic complex of concepts for the expression. This means that so much context-dependence remains as there are concepts in the complex. This context-dependence cannot be eliminated, except by chosing different names, such that each concept from the complex would have its own name. But then the connectedness of the different concepts of the polysemic complex would not be expressed in the lexicon. There is no difference in principle between context-dependence and context-independence, if there are finitely many contexts of use. One simply can use different names. But if the set of possible contexts is not limited to a fixed set, and if the possible contexts are not predictable otherwise in the lexicon, then there is a difference in principle between context-independent and context-dependent use of lexical items. The lexicon of a natural language is context-dependent. The few cases of context-independent use, i.e. the few cases in which distinctive full generalization across all contexts is possible, are borderline cases.

7. Objectivization, objectivity

For a learning individual, a pattern becomes a concept-'implementation' only when he or she is able to make a difference between a situation outside and the experience of the situation, as far as it is based on agreement with the pattern. Making this difference implies that the perceiver has to be aware of the fact that a series of perspective dependent perceptions of the same situation has been experienced or can be experienced, and that local-temporal connections of these experiences of the situation with experiences of surrounding situations have been established. Thus, a notion of reality has to be developed in the Husserlian fashion within the mind by placing experiences within space-time and causal connections that establish the identity of objects. In this way one forms partial objects, and then, by extending the space-time regions, places them into bigger partial objects and situations, as well as into courses of situations with make up partial worlds, being extendable to larger partial worlds, etc. This means that every experience is placed into wider horizons of possible extensions, and by this embedding, awareness of reality is formed as distinct from a single experience and as distinct from the awareness of a pattern of similar experiences. Knowing that some experience is real and not just imagined means knowing that it is embedded in these processes of transcending the experience by further experiences that relate to it in certain ways in space and time.

Concepts are not just stabilizing and slowly changing patterns in the brain, which appear in activation as loose experiences and feelings in the whole of experiences and feelings which constitute the mind. They are, whenever activated, instantiated within these broader (historical) horizons which establish in the mind the awareness of objectivity. Only when our experiential concepts are organized within concepts of (historical) objects, can they be used for characterizing objects and thus be used in representing knowledge. Without this integration into concepts of objects, i.e. into individual concepts, we just would have situational impressions and patterns of these. We have horse-concepts by ordering the context-dependent experiences of horse-situations, where we see, feel, and hear horses. These experiential concepts are first merely concepts of impressions of horse-situations. However, after grasping the objectivity of situations by uniting many real and possible situational impressions into the space-time-causal identity of a situation we form concepts of objective horse-situations, the situational horse-concepts. In a third step, the concept of horse individuals can be formed by uniting certain horse situations into a space-time-causal continuous unity, i.e. a (possible) history of an individual object. By having taken the second step we have knowledge about the objective presence of certain horse situations. By having taken the third step we have knowledge about individuals and can devise possible situations in which an individual plays a role. We can now design situations and start out to see whether they are real. We can also make our designs real in performing actions. We understand our situational experience in perception as being real situations, by locating them within further, objectivizing, (possible) surroundings or Husserlean 'horizons'. On the first step no intentional language use is possible: There are merely associations between re-occuring impressions and utterances, established as reaction patterns. Only on the second and third step we have concepts in the full sense available, which can be used about something. They have ankers, in the form of instances, in an objective, real or possible world. An objective world is one that contains objects, i.e. situations, material, and individuals. A world consisting merely of unordered impressions would not be an objective world. Intentional use of language is only possible with respect to an objective world.

8. Conclusion:

Connectionists' models provide a way to understand the basic relationship between recurrent situations and their stabilizing and changing neural patterns. Thus they are already now in principle able to account for the first step in the formation of experiential concepts. Since space-time and causal contiguity, which are essential for the second and third step, can be represented by association, connectionism should in principle also be able to represent situational concepts, at least of those situations that take short time spans. Here in a single long vector, three vectors can be connected, one that typically represents the temporal beginning, one that represents the middle, and one that represents the end of a situation. This would be a partial, but often sufficient representation of a situation that can be fed into a net. However this still would merely provide a subjective picture of situations. In the objective view one has to connect the subjective, egocentric, impression with a representation of the view of other participants in an action. This presents a serious difficulty in a connectionist model. An additional difficulty is the representation of individual concepts. Here we have the problem of representing long stretches of time, because several partial objects have to be united into a single space-time path which is a (partial) history of the individual. The individual is a possible completion of each of its partial histories. Without this concept of a historical object, several situational occurences of an individual could not be understood as being occurences of a single individual. In cognitive models which would be able to capture individual concepts, representations of episodes or situations must be represented within courses of history, and it is difficult to imagine how that will be possible in the kind of connectionist models we know today. Right now, I cannot imagine how situational representations assumed under the perspectives of other participants, and how history could be represented without a linguistic medium and a mental model of time, in which episodes can be represented in temporal order. This, of course, is also necessary for understanding texts. For these tasks a traditional symbolic model of cognition with a semantics including space and time still seems necessary.

The dynamic data-oriented theory of concept formation and understanding advocated here (cf. Bartsch 1993), proceded further than connectionist models so far can do. In our model, syntactic categories can be understood as stabilizing growing sets of expressions with an internal relationally determined similarity based on semantic correspondences and contextual, syntagmatic, relationsships. In this way, syntactic categories are constructed as meta-linguistic concepts, i.e. conceptual structures on growing sets of expressions. They have certain cognitive and semantic values. The cognitive values are operations performed on conceptual structures in understanding an expression, the semantic values are operations in interpreting an expression with respect to a world.

References:

Bartsch, R. 1986. Context-dependent Interpretation of Lexical Items. In: J. Groenendijk, D. de Jongh, and M. Stokhof (eds.), *Pragmatics and Lexical Semantics. Selected Papers of the fifth Amsterdam Colloquium.* GRASS Series 7. pp. 1-26. Dordrecht: Foris Publications.

Bartsch, R. 1987. The Construction of Properties under Perspectives. In: *Journal of Semantics 5*: pp. 293-320.

Bartsch, R. 1993. Concept Formation and Understanding: An Overview. In: D.Nauta, A. Nijholt, J. Schaake (eds.), *Pragmatics in Language Technology. Procedings of the fourth Twente Workshop on Language Technology*: pp. 71-86. University of Twente. Department of Computer Science. P.O. Box 217. NL-7500 AE Enschede.

Hofstadter, D. 1982/1985. Waking Up from a Boolean Dream, or subcognition as Computation. In: *Metamagical Themas: Questing for The Essence of Mind and Pattern*. Penguin Book.

Kangas, J. 1990. Time-Delayed Self-organizing Maps. In: I. Alexander and J. Taylor (eds.), *Proceedings of the International Joint Conference on Neural Networks 2. Vol. 1*. pp. 331-336. San Diego.

Kohonen, T. 1984. *Self-organization and Associative Memory*. Berlin, New York: Springer Verlag.

Kohonen, T. 1991. The Hypermap Architecture. In: T. Kohonen et al. (eds.), *Artificial Nural Networks*. Vol. 2. pp. 1357-1362.

Ritter, H., T. Martinez, and K. Schulten 1992. *Neural Computation and Self-organizing Maps*. Addison Wesley Publishing Company.

Sag, I. 1981. Formal Semantics and Extra-Linguistic context. In: P. Cole (ed.), *Radical Pragmatics*. New York: Academic Press. pp. 295-318.

Schyns, Ph. 1991. A Modular Neural Network Model of Concept Acquisition. In: *Cognitive Science 15*. pp. 461-508.

Scholtes, J. 1993. *Neural Networks in Natural Language Processing and Information Retrieval*. Ph.-Dissertation. Department of Computational Linguistics. University of Amsterdam.