Hierarchical Categorization

A. Wichert
Department of Neural Information Processing
University of Ulm, D-89069 Ulm, Germany
wichert@neuro.informatik.uni-ulm.de

Abstract

In this paper we introduce a biologically and psychologically plausible neuronal model of hierarchical categorization. The knowledge in our model is represented by a taxonomical arrangement of verbal categories. This categorical representation is psychologically motivated and also offers an explanation of how to deal with uncertain knowledge. It is an alternative to other well known uncertainty calculi. An observer specifies the known features before the hierarchical categorization begins. During the categorization the model learns to favor those categories which often lead to a successful goal. This may help to speed up the search. A computer simulation of a system for the diagnosis of the problem with a car is presented.

Keywords: hierarchical categorization, availability heuristic, neural networks, semantic networks, uncertain knowledge

1 Introduction

The goal of the hierarchical categorization is the accurate determination of a category or some categories out of present facts. The representation of knowledge by categories (Smith, 1995) is the central idea of our introduced model, the abductive memory. This categorically motivated uncertainty calculus is an alternative to the other well known uncertainty calculuses (Duda et al., 1979; Shafer, 1976; Shortliffe and Buchanan, 1975; Zadeh, 1975), for summary see (Lucas and van der Gaag, 1991), or other approaches (Sun, 1995).

2 Categories

2.1 Description

A way to represent the objects by verbal categorization is to describe them by a set of discrete features (Tversky, 1977). The similarity between them can be defined as a function of the features they have in common (Osherson, 1995; Sun, 1995). An object is judged to belong to a category to the extent that its features of it are predicted by the category

(Osherson, 1987).

2.2 Similarity Measure

The similarity of a category Ca and of a feature set J is given by the following formula, which is inspired by the contrast model of Tversky (Tversky, 1977),

$$Simc(Ca, J) = \frac{|Ca \cap J|}{|Ca|} - \frac{|(Ca-(Ca \cap J)|}{|Ca|} = \frac{2}{|Ca|} \cdot |Ca \cap J| - 1 \quad \in [-1, 1]$$

$|Ca|$: number of the prototypical features that define the category Ca.

The value $Simc(Ca, J)$ is called the quality criterion (qc) of the category Ca for the given feature set J. With $qc = 1$ absolute similarity and $qc = -1$ no similarity at all. For example the category **bird** is defined by following features: flies, sings, lays eggs, nests in trees, eats insects. The category **bat** is defined by the following features: flies, gives milk, eat insects. The following features are present: flies and gives milk.

Simc(**bird**, present features)=$\frac{1}{5} - \frac{4}{5} = -0.6$
Simc(**bat**, present features)=$\frac{2}{3} - \frac{1}{3} = \frac{1}{3}$

So the qc of the category bird -0.6 is and of the category bat is 0.3333.

2.3 Salience of a Feature

Features that discriminate among relevant facts should have a higher salience than the one that do not (Smith, 1995). The features of an equal salience have a unary representation, they can only be represented as existent or nonexistent. A category that is described as a set of features can be present with different grades of vagueness corresponding to the cardinal number of the set. A set of features that describes a category can be sometimes divided into subsets that represent some subcategories. Each feature can be also regarded as a kind of subcategory. If this subcategory can not be described by other features, but, nevertheless, should have the properties of variable salience and vagueness, it is described by invisible features. To each feature a number of invisible features is assigned dependent on its salience. This salience determines the range of the invisible features that can be present. The actually present invisible features describe the contribution of a feature to a category. They can be identified as the belief value in the presence of a feature. Through the coupling of the salience of the features with the possible grades of belief values one is forced to think more carefully about the important features than about the less important ones.

2.4 Uncertainty of a Category

An observer determines the presence of features corresponding to each category. But if it is not possible to observe some features, the quality criterion can never reach the maximal value 1. If some features exist for a complete definition of a category, but because of the lack of available knowledge they cannot be named, they cannot be verified. These features are called the unobservable features. Through their number it is expressed, how certain a definition of a category is.

3 Representation

3.1 Units

A category is expressed by a set of prototypic features. This set can be represented by a binary vector, in which the positions represent different features. For each category a binary vector can be defined. A one at a certain position corresponds to a certain feature. This vector can be described by a unit which models a biological neuron (McClelland and Kawamoto, 1986; Palm, 1982) (see Fig. 1). The stored binary weights represent the presence or absence of features that describe a category. Each category corresponds to a unit.

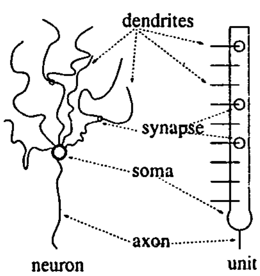

Figure 1: A unit is an abstract model of a biological neuron (Hertz et al., 1991).

The principles in the representation will be demonstrated on an example of two old sayings from country folklore:

(1) If it is April and it snows much then probably the apple harvest will be bad.

(2) If it is April and it rains a lot and it is very cold then the vintage will be good.

The feature "April" corresponds to the first position of the binary vector, it can be only present or absent. The feature "snows" is described by two invisible features because it has a higher salience than the feature "April", this is expressed in the country saying through the adjective "much". The feature "snows" can be either present, maybe present or absent. The two invisible features correspond to the positions two and three. The feature "rains a lot" has the same salience as the preceding feature, it is described by the positions four and five. Because the feature "very cold" has the greatest salience of the other features it is described by three invisible features, namely the positions six, seven, and eight. The uncertainty of the first category which corresponds to the adjective "probably" is expressed by two unobservable features. So the two country sayings are represented by two units which form a module (see Fig. 2). During the categorization the present feature set is presented at the corresponding positions and the quality criterion of each linear unit (Hertz et al., 1991) is calculated corresponding to the weights.

The hypothesis about the presence of a category is made, then it is tested (Elstein et al., 1978) with the aid of the external memory. This procedure corresponds to the abductive inference in which a hypothesis is made and then it is tested (Peng and Reggia, 1990; Josephson and Josephson, 1996). A quality of this kind of rule of inference is that it is unsound, which means that

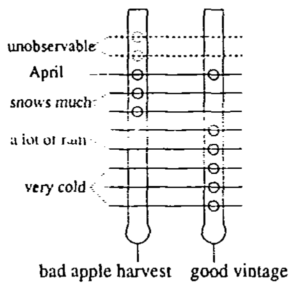

Figure 2: Neural representation of two country lores.

even if the test is positive this does not mean that the hypothesis is necessarily true.

3.2 Taxonomic Representation

With modules only small amounts of categories can be verified, because the represented knowledge is not structured. One of the most effective ways to structure knowledge is the taxonomic arrangement of the information that represents it (Resnikoff, 1989). A well known model of the taxonomic representation is the semantic network (Quillian, 1968). A semantic network represents the knowledge as directed acyclic graph, with nodes corresponding to concepts and the links between them corresponding to relations between them (Otman, 1996). Categories can be split up into subcategories, so that a taxonomy can be constructed and represented by an acyclic graph. The nodes in this graph correspond to categories and the links indicate the "is a subcategory" relation between them. The process of the hierarchical categorization can be performed, in which one moves from more general categories to more accurate categories until the desired category is reached. These ideas correspond to the other models of the modular organization of the memory like (Breitenberg, 1978; Palm, 1982; Rumelhart and McClelland, 1986; Palm, 1990), and the models of the connectionistic realization of the semantic networks (Shastri and Feldman, 1985; Shastri, 1988).

3.3 The Categorization

The categorization task at each node is performed by a module that represents a certain category and which determines another more accurate category. This other more accurate category can be represented either by a module or a unit. The second case represents the goal of a hierarchical categorization. The problem space (Newell, 1990) is defined by the connection between the modules. The known features and the belief values about their presence and their salience are stored in the external memory before the categorization begins. About the other unspecified features nothing is known, their probability of their presence is fifty percent or less depending on their salience,

$$belief\ in\ the\ not\ specified = \lfloor \frac{salience}{2} \rfloor.$$

The feature can be present in different modules and have different salience dependent on the context (Smith, 1995). The external memory calculates the new belief values dependent on the context which is defined by the new salience,

$$new\ belief = \lfloor new\ salience \cdot \frac{old\ belief}{old\ salience} \rfloor.$$

The external memory guides the search in the problem space in the direction of the most plausible category, which is the category with the highest quality criterion value. In the search the first favored category can turn out to be wrong because its subcategories are not present. In this case another category is examined. This search strategy corresponds to hill climbing (Winston, 1992), which is a depth-first search in which the choices are ordered according to the quality criteria values (see Fig. 3). The quality criteria values represent a heuristic measurement of the distance to the remaining goal. The whole model is called the abductive memory because it does the categorization in the abductive way.

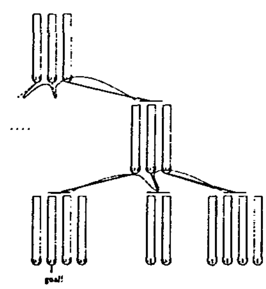

Figure 3: The hill climbing search strategy in the abductive memory during the categorization.

We demonstrate the abductive memory in the knowledge base for diagnosing car problems (see Fig. 4).

The known features and the belief values about their presence are stored in the external memory before the categorization begins. About the other unspecified features nothing is known, their probability about their presence is fifty percent or less

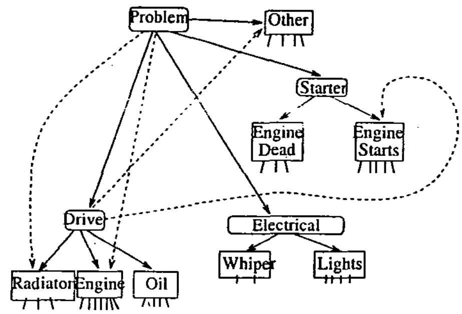

Figure 4: The taxonomic representation of disorders of a car. The dotted arrows represent the shortcuts to categories which are less certain. The rectangular boxes represent the possible goals of the hierarchical categorization. Their number is drawn in addition.

dependent on their salience. The only feature known by the observer is POOR_ENGINE_PERFORMANCE with the maximal belief value (salience=6, belief=6).

1) ! MODULE ELECTRICAL with qc=0 !

2) ! MODULE LIGHTS with qc=0 !

3) ! MODULE WIPER with qc=0 !

4) ! MODULE STARTER with qc=0 !

5) ! MODULE ENGINE_STARTS with qc=0 !

6) ! MODULE ENGINE_IS_DEAD with qc=-0.166667 !

7) ! MODULE OTHER with qc=-0.333333 !

8) ! MODULE DRIVING with qc=-0.333333 !

9) ! MODULE ENGINE with qc=-0.166667 !

R E S U L T :
DIRTY_AIR_FILTER_OR_BAD_SPARK_PLUGS with qc=0.222222
FUEL_PUMP_BAD_OR_FUEL_PIPE_BROKEN with qc=0.0555556

The information which was stored in

the external memory was sufficient for the determination of the disorder but insufficient to guide the search, which took nine steps.

3.4 Heuristics

A heuristic could help to speed up the search in the problem space that is defined by the categories and the links between them. During the hierarchical categorization some categories come to mind more easily, because they were determined to be more frequent than the other. Some psychologists (Tversky and Kahneman, 1973) assume that the individuals estimate the frequency of an event, in our case the determination of a category. This kind of heuristic is called the availability heuristic (Tversky and Kahneman, 1973; Wickelgren, 1977; Schwartz, 1995). The idea how to implement this kind of heuristic was given by the psychologist William James already around 1870 (James, 1985): *Habits are due to pathways by the nerve centers, in getting out they leave their traces in the paths which they take.* In our model this corresponds to the strengthening of the links between the categories which describe a successful search by a small factor. The successful search corresponds to the path of links between the categories from the category where the search began to the results of the hierarchical categorization. Paths that corresponds to wrong search directions are weakened by a small factor (see Fig. 5). During the repeated search the categories that receive strong links are favored. The search direction corresponds to the highest value that is the sum of the quality criterion of a category and the strength of the link to it.

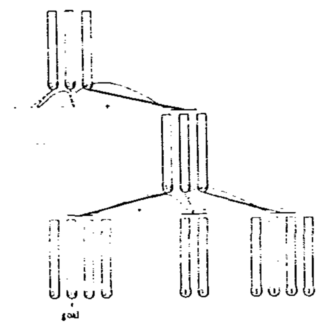

Figure 5: The learning of the availability heuristic. Links between the categories which describe a successful search are strengthened by a small factor.

Suppose that we are trying to diagnose the problems with cars that are used near a dusty desert where sandstorms are very. Very often the problems of such cars are caused by dirty air filters. After the strengthening of the corresponding links between the categories during the previous categorizations which describe this fact, the search takes with the availability heuristic only two steps.

```
1) ! MODULE DRIVING with qc=-0.333333 !
2) ! MODULE ENGINE with qc=-0.166667 !
----------------------
R E S U L T :
DIRTY_AIR_FILTER_OR_BAD_SPARK_PLUGS
with qc=0.222222
FUEL_PUMP_BAD_OR_FUEL_PIPE_BROKEN with
qc=0.0555556
```

This can explain the fact that memory contents are accessed faster in the human memory if they were accessed often before (Anderson, 1995).

4 Summary

The abductive memory is an efficient biologically and psychologically motivated engineering tool for representation and an easy access of uncertain knowledge. The categorical representation offers an alternative to the traditional uncertainty calculus (Duda et al., 1979; Luger and Stubblefield, 1993; Shafer, 1976; Shortliffe and Buchanan, 1975; Zadeh, 1975). This work was supported by the SFB 527.

References

Anderson, J. R. (1995). *Cognitive Psyhology and its Implications.* W. H. Freeman and Company, fourth edition.

Breitenberg, V. (1978). Cell assemblies in the cerebral cortex. In Heim, R. and G.Palm, editors, *Theoretical Approaches to Complex Systems*, pages 171–188. Springer-Verlag.

Duda, R. O., Gaschnig, J., and Hart, P. E. (1979). Model design in the prospector consultant program for mineral exploration. In Michie, D., editor, *Expert Systems in the Microelectronic Age.* Edinburgh University Press.

Elstein, A., Shulman, L., and Sprafka, S. (1978). *Medical Problem Solving - An Analysis of Clinical Reasoning.* Havard University Press.

Hertz, J., Krogh, A., and Palmer, R. G. (1991). *Introduction to the Theory of Neural Computation.* Addison Wesley.

James, W. (1985). *Psychology, the Briefer Course.* University of Notre Dame Press, Notre Dame, Indiana.

Josephson, J. R. and Josephson, S. G. (1996). *Abductive Inference, Computation, Philosophy, Technology.* Cambridge University Press.

Lucas, P. and van der Gaag, L. (1991). *Principles of Expert Systems.* Addison-Wesley.

Luger, G. F. and Stubblefield, W. A. (1993). *Artificial Intelligence, Structures and Strategies for Complex Problem Solving.* The Benjamin / Cummings Publishing Company, Inc., second edition.

McClelland, J. and Kawamoto, A. (1986). Mechanisms of sentence processing: Assigning roles to constituents of sentences. In McClelland, J. and Rumelhart, D., editors, *Parallel Distributed Processing*, pages 272–325. MIT Press.

Newell, A. (1990). *Unified Theories of Cognition.* Harvard University Press.

Osherson, D. N. (1987). New axioms for the contrast model of similarity. *Journal of Mathematical Psychology*, (31):93–103.

Osherson, D. N. (1995). Probability judgment. In Smith, E. E. and Osherson, D. N., editors, *Thinking*, volume 3, chapter two, pages 35–75. MIT Press, second edition.

Otman, G. (1996). *Les Représentations Sémantiques en Terminologie.* Masson.

Palm, G. (1982). *Neural Assemblies, an Alternative Approach to Artificial Intelligence.* Springer-Verlag.

Palm, G. (1990). Assoziatives gedächtnis und gehirntheorie. In *Gehirn und*

Kognition, pages 164–174. Spektrum der Wissenschaft.

Peng, Y. and Reggia, J. A. (1990). *Abductive Inference Models for Diagnostic Problem-Solving*. Symbolic Computation. Springer-Verlag.

Quillian, R. (1968). Semantic memoiry. In Minsky, M., editor, *Semantic Information Processing*, pages 227–270. MIT Press.

Resnikoff, H. L. (1989). *The Illusion of Reality*. Springer-Verlag.

Rumelhart, D. and McClelland (1986). A distributed model of human learning and memory. In McClelland, J. and Rumelhart, D., editors, *Parallel Distributed Processing*, pages 170–215. MIT Press.

Schwartz, N. (1995). Social cognition: Information accessibility and use in social judgment. In Smith, E. E. and Osherson, D. N., editors, *Thinking*, volume 3, chapter ten, pages 345–376. MIT Press, second edition edition.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ.

Shastri, L. (1988). *Semantic Networks: An Evidential Formulation and its Connectionistic Realization*. Morgan Kaufmann, London.

Shastri, L. and Feldman, J. (1985). Evidential reasoning in semantic networks: A formal theory. In *Proc. of IJCAI'85*, pages 465–474. IJCAI.

Shortliffe, E. H. and Buchanan, B. G. (1975). A model of inexact reasoning in medicine. *Mathematical Biosciences*, (23):351–379.

Smith, E. E. (1995). Concepts and categorization. In Smith, E. E. and Osherson, D. N., editors, *Thinking*, volume 3, chapter one, pages 3–33. MIT Press, second edition edition.

Sun, R. (1995). A two-level hybrid architecture for structuring knowledge for commonsense reasoning. In Sun, R. and Bookman, L. A., editors, *Computational Architectures Integrating Neural and Symbolic Processing*, chapter 8, pages 247–182. Kluwer Academic Publishers.

Tversky, A. (1977). Feature of similarity. *Psychological Review*, (84):327–352.

Tversky, A. and Kahneman (1973). Availability: A heuristic for judginig frequency and probability. *Cognitive Psychology*, (5):207–232.

Wickelgren, W. A. (1977). *Cognitive Psychology*. Prentice Hall.

Winston, P. H. (1992). *Artificial Intelligence*. Addison Wesley, third edition.

Zadeh, L. A. (1975). Fuzzy logic and approximate reasoning. *Synthese*, (30):407–428.