Linguistic Pattern Characterization and Analysis

By

Christian Friedrich-Freksa

B.S. (University of San Francisco) 1974
M.S. (University of California) 1976

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Engineering Science

in the

GRADUATE DIVISION

OF THE

UNIVERSITY OF CALIFORNIA, BERKELEY

Approved:

L. A. Zadeh
....
Chairman Date

Stephen E. Palmer 3/17/81
....

....

....

LINGUISTIC PATTERN CHARACTERIZATION AND ANALYSIS *

by

Christian Friedrich-Freksa
University of California
Berkeley

ABSTRACT

This dissertation presents a model for communicating about objects by means of verbal descriptions. The functions of verbal descriptions are to assimilate knowledge about object features and to convey this information to an interpreter. The interpreter analyzes the object descriptions in terms of its own semantic model and by comparison to reference objects.

A hierarchy of object description languages is developed to demonstrate the advantages of fuzzy languages for communication in complex environments. There is a trade-off between crisp and precise languages suitable for well-defined, small domains, and fuzzy imprecise languages suitable for ill-defined, complex domains. Messages in a language of the former type are determined mainly by

---

* Sponsors: German Academic Exchange Service (DAAD)
National Science Foundation Grant ENG78-23143

definitions of vocabulary and rules which make up the language. In contrast, messages in a language of the latter type are determined by the particular domain and context in which they are used. A fuzzy language does not describe a particular object or point in a feature space but rather it describes a subspace of the feature universe. If this subspace contains only one element in the particular domain of discourse, the fuzzy description has crisp denotation.

In our model, information about objects is represented in terms of linguistic symbols which correspond to possibility distributions over sets of features and by relations between linguistically expressed feature values. Possibility distributions and relations can be manipulated by linguistic operators. It is shown how the chosen object representations capture imprecise and incomplete knowledge and how the intended meaning of the description can be recovered by the interpreter with the aid of the situation context. The target object is identified by elimination of alternative possibilities which are found to be in conflict with the description.

A central theme of this research is that imprecision in descriptions can be captured by fuzzy possibility distributions and can be exploited for effective and efficient communication in complex environments. Possibility is treated as a graded concept and interpreted in terms of "ease of making possible". Linguistic

information processing methods are proposed as an alternative to numerical approaches. Methods for interpreting subjective descriptions are suggested.

For interpretation of linguistic descriptions, adequacy of and agreement between descriptors are distinguished for determining their compatibility. It is shown how possibilistic information can be used to give more informative responses than "yes" and "no" to retrieval requests. The role of linguistic modifiers, determiners, and quantifiers is discussed.

The implementation of L-FUZZY, a dialect of the AI-language FUZZY is described. L-FUZZY incorporates linguistic instead of numerical modifiers and directly represents fuzzy possibility distributions instead of fuzzy set elements. The response to a retrieval request is designed to be more informative in this dialect than in FUZZY. Examples for applications of fuzzy communication are given. The proposed model explains imprecision, ambiguity, vagueness, and misunderstanding in human communication as well as richness, adaptability, double meaning, conciseness, efficiency, robustness, expressness, and flexibility.

Lotfi A. Zadeh
(chairman of thesis committee)

To my teachers

Acknowledgements

I would like to thank the members of my thesis committee, Professors Lotfi Zadeh, Lawrence Stark, and Stephen Palmer for their inspiration, advise, and criticism.

I am grateful to Harry Barrow and Jay Tenenbaum for suggesting an early version of the research problem.

Ramon Lopez de Mantaras, Piero Bonissone, Lucia Vaina, Dennis Allard and Patrizia Violi were valued discussion partners during various stages of the project.

Rick LeFaivre, Wolfgang Wahlster, Susana Berestovoy, Uli Furbach, Klaus Haagen, and Peter Schefe helped me refine my approach by providing alternative viewpoints.

Gerhard Dirlich and Heiner Ellgring read earlier versions of this dissertation and offered detailed criticism.

Donald Glaser's Molecular Biology group generously granted me access to its computing facility. Bob Henry's support and expertise ensured the successful running of my computer programs.

I had the great opportunity of working in various research environments: Lotfi Zadeh's expert systems group; the Berkeley "frames group"; Larry Stark's bioengineering group; the Max-Planck-Institute for Psychiatry in Munich, West Germany; and the IKERLAN research institute in Mondragon, Spain. My office mates in 275 Cory Hall created a rich atmosphere in which I enjoyed working.

A scholarship from the German Academic Exchange Service and a National Science Foundation grant to Professor Zadeh supported my work.

I thank Gabi Zollner for providing me with object descriptions, for helping put my work into perspective, and for her patience. I thank my friends in Europe for their encouragement.

My parents evoked my interest in research. They created the security and freedom which allowed me to pursue my studies the way I did.

TABLE OF CONTENTS

CHAPTER 1

COMMUNICATING ABOUT OBJECTS

"Alice opened the door and found that it led into a small passage, not much larger than a rat-hole: she knelt down and looked along the passage into the loveliest garden you ever saw."

(Lewis Carroll: Alice in Wonderland)

"First find the air cleaner. Through 1972 it is a large round metal black thing with metal clips on it and thin and fat hoses going to it. ...

"Look to a little left of the center of the engine and find a brown or black round plastic thing with five heavy wires sticking out of it. That's the distributor cap."

(John Muir & Tosh Gregg:
How to keep your Volkswagen alive!
A manual of step by step procedures
for the compleat idiot)

"He was a very tall, thin man, with a long nose like a beak, which jutted out between two keen, grey eyes, set closely together and sparkling brightly from behind a pair of gold-rimmed glasses."

(A. Conan Doyle:
The hound of the Baskervilles)

1.1 Introduction

The descriptions above have in common that they create ideas about objects in the reader. These ideas are not complete or crisp enough to enable the reader to reconstruct the described objects in such a way that they would closely resemble their models. However, they may be good enough to enable the reader to identify the objects in a given situation. The environment supplies information missing from the description to make correct identification of the target object possible.

This dissertation develops a family of object description languages and investigates their properties. The goal is to explain how communication about objects in complex environments can take place rather efficiently if context-adaptive fuzzy descriptions are employed. Description matching is examined on a local level: rather than asking how different pieces of descriptions join together to form a model of space [Kuipers (1977, 1978)], we look in detail what individual pieces -- which we will call "descriptors" -- contribute to the denotation of an object. In particular, we are interested in representations of linguistic information and in methods for comparing linguistic descriptors with one another.

We will use the term "linguistic matching" for the process of comparing two linguistic object descriptors or of comparing linguistic object descriptors with sensory information about objects. Linguistic matching is used to interpret a world of objects in order to select that subset of the world which corresponds to the description.

The term "linguistic" is used in this thesis in contrast to the attribute "numerical". Meaning of linguistic descriptions is determined, in part, by the use of the descriptions, whereas meaning of numerical descriptions normally is based on precise context-independent definitions. Linguistic descriptions generally are incomplete and fuzzy. The relevance of an object descriptor for the interpretation of a description depends on the fuzziness of the descriptors, on the other descriptors in the description, and on the set of candidate objects involved in the matching process. The interpretation process is guided by the context of given objects and by the linguistic matcher itself. Suitable object properties typically are color, size, shape, etc., although any other type of characteristics could be used.

A basic assumption of this research is that knowledge can be represented by linguistic descriptions and that linguistic matching can be modeled by computational operations on these symbolic descriptions. At the functional level, linguistic matching can be considered as a process that receives inputs describing an environment and a particular object within that environment, and identifies the object. At a higher level, linguistic matching may involve learning. The processor and its ability to interpret linguistic descriptions, change as a result of the inputs it receives.

The research reported here deals with linguistic matching viewed as a functional interpretation process. Linguistic descriptors are modeled by possibility distributions over object features [Zadeh (1978a,b)]. These distributions preserve imprecision and fuzziness which are inherent in all knowledge about facts or objects [Russell (1923), Black (1963)]. The learning aspects, i.e., the question of how meaning of linguistic descriptors can be acquired, are being investigated in a companion project [Lopez de Mantaras (1980a)]. This work does not address the problem of description generation [compare Breeding & Amoss (1972)].

A result of this research is the PINPOINT model, a computational model of linguistic description matching. The model shows under what conditions incomplete, fuzzy, subjective, linguistic descriptions can be interpreted in such a way that they pinpoint the precise object or set of objects that the describer intended.

1.2 Why Are Object Descriptions Worth Studying?

In complex environments, such as the world of physical objects, conventional computer science approaches for communication break down. It is not feasible, for example, to give each object a distinct name or to describe objects in canonical form. Also, it does not make sense to use all known features of an object to describe this object. Most importantly, however, it is disadvantageous to make distinctions between all possible feature values at all times.

This last point has been neglected in most artificial intelligence (AI) research in which it is generally assumed that distinguishable features have distinguishable labels. One reason for this is that it was not clear how to take advantage of fuzzy representations. Fuzzy representations are capable of variable feature resolution. Another reason is that imprecision and fuzziness in perception, communication, and reasoning have been considered to be deficiencies that eventually could be overcome rather than intrinsic properties of any representation of the "real world".

These considerations raise two questions: 1) how can, do, or should we describe objects in complex environments, and 2) how can, do, or should we interpret these descriptions in order to identify the target objects. Answers to these questions will lead to deeper insight into communication aspects of descriptions: what makes object features relevant for descriptions, how is information about them transmitted, and what does this information mean?

The motivation for this research has been two-fold. First, we have been interested in devising methods for communicating in environments too complex and unpredictable to be tackled by existing approaches. Second, we like to offer explanations for mechanisms underlying the efficiency and effectiveness of human communication as well as its drawbacks by providing a model whose behavior exhibits aspects comparable to human communication.

1.3 The Problem

The problem which this research addresses now can be stated as follows: Suppose we are given an arbitrary scene of objects. We want to have a language to point out any of the objects we perceive to a "communication partner" having comparable perception abilities. This is to be done effectively and efficiently, i.e., the communication partner should be able to identify the target object with a low amount of effort. The efficiency criterion constrains the tolerable complexity of the language: the language must be acquired and stored by each of the communication partners. There is a trade-off between specificity and expliciteness of a language on one-hand and effort for acquisition and maintenance of the language on the other hand.

In addition to conventional computer science and artificial intelligence approaches in which imprecise* and fuzzy* information about environments is represented in precise* and crisp* terms we need ways to capture imprecision and fuzziness as is inherent in natural conceptualizations and descriptions of objects and environments. This is to gain flexibility and efficiency in manipulating descriptions. In particular, the representations should not introduce conceptual discontinuities which are not present in the archetype world [Tribus (1979)]. In other words, it is preferable that the representation of a body of knowledge slightly distorts this knowledge but preserves essential properties, than that it "idealizes" the knowledge and destroys essential properties. "Idealization" of knowledge would also distort this knowledge.

For representation of a given environment, this can be achieved if descriptors capture distinctions between features rather than the features themselves. The descriptions then characterize a feature subspace of the universe of features rather than a precise feature point. This feature subspace could denote an infinite number of different objects, but in a typical context in which such descriptions may be meaningful, only one or a few objects

----

* For definitions of these terms, refer to the glossary in appendix A.

can be found in this feature subspace. The interpretation of a description is done by instantiation of the particular situation context.

Natural language object descriptions fulfill the desiderata outlined in this and the previous section: objects which are frequently referenced have names; others are described in terms of some of their features. Not all detectable features are used in the descriptions, since a few features suffice to uniquely distinguish a given object from other objects in a given environment. Feature labels can have imprecise and fuzzy denotations.

This thesis discusses the semantics of such type of linguistic descriptions and presents an approach to their interpretation. The PINPOINT model for interpretation of object descriptions is presented and a computer implementation of linguistic representation is described.

1.4 PINPOINT Model

A collection of feature descriptors referring to an object or a set of objects will be called "object description". The PINPOINT model illustrates how object descriptions having common characteristics with natural language descriptions can be interpreted in such a way that the target object described can be pinpointed. The model does not deal with the difficult issue of natural language utterance interpretation. Instead, a natural-language object description is translated into a formal meaning representation language for natural languages, called PRUF [Zadeh (1978b)].

Translation from natural-language expressions into PRUF usually is performed by humans, although for sufficiently restricted domains automatic translation is possible [Lopez de Mantaras (1980b)]. The PINPOINT model refers to a data base representing the reference environment to interpret the PRUF representations of object descriptions. The goal is identification of the target object in its environment. The paradigm that we use throughout this thesis is the description of physical objects. The reason is that we can empirically verify if an object has been correctly identified.

The object identification problem can be viewed as a pattern matching problem in which a coarse description roughly characterizes the class of patterns to be accepted and the situation context supplies information to restrict the possible matches. In our system, linguistic object descriptors are represented by fuzzy possibility distributions. In this representation, imprecision of linguistic terms is preserved and can be taken advantage of. Object descriptions are mediated from the describer to the interpreter by means of PRUF expressions. A dialect of the programming language FUZZY [LeFaivre (1974a,b)] serves to interpret PRUF expressions with the goal of identifying the described object. This dialect, L-FUZZY, directly represents fuzzy sets (rather than elements of fuzzy sets) and linguistic modifiers (rather than numerical ones).

1.5 Related Work

Pattern matching and identification plays an essential role in virtually all information processing systems. In this thesis I develop a generalization of the conventional matching paradigm in which information is viewed as crisp and absolute. In the following sections I will relate this work to the fields of Artificial Intelligence, Cognitive Psychology, Linguistics, Fuzzy Set Theory, and Philosophy. References to specific issues will be given together with detailed discussion in the following chapters.

1.5.1 Artificial Intelligence -

The relation of this work to AI is three-fold:
1) several theoretical aspects overlap with issues that arise in scene analysis and language processing research,
2) the research methodology is taken from AI, and
3) we are using tools that are products of AI research and have been developed for AI research.

1.5.1.1 Scene Analysis And Language Processing -

Scene analysis research is concerned with the problem of recognizing objects in visual scenes [e.g. Roberts (1963), Guzman (1968), Winston (1970), Turner (1974), Garvey (1976), Deering (1980)]. The particular appearance of an object must be compared with the description of a class of objects. This requires suitable representation of the objects [Firschein & Fischler (1971, 1972), Minsky (1974), Barrow & Tenenbaum (1975)]. The "feature selection problem" [Duda & Hart (1973), Nagel (1976), Bezdek & Castelaz (1977)] for generating good object descriptions is too little understood to offer general solutions [Bremermann (1972)]. Domain-specific knowledge appears crucial for efficient object recognition [Akin & Reddy (1977), Bajcsy & Tavakoli (1975), Freuder (1976), Garvey (1976), Harmon & Hunt (1977)]. Multisensory information about scenes [Tenenbaum (1973)] and selective search processes [Barrow et al. (1972), Hanson & Riseman (1974), Bajcsy & Rosenthal (1975), Ballard (1978)] simplify identification tasks tremendously by drastic restriction of possible object references.

Language processing is the largest sub-area of AI. Here, I mention only a few systems which use language for interacting with scenes of objects or object descriptions. The best known of these is SHRDLU [Winograd (1972, 1973)]. SHRDLU accepts object descriptions in English to manipulate objects in a blocks world. All concepts in this artificial world are clear-cut and therefore there is never a difficulty of feature matching. Shaket (1975) also uses a blocks world, but he deals with the fuzziness inherent in natural-language descriptors. HAM-RPM (Hamburg Dialogue Partner Model) [Hahn et al. (1979)] is an ambitious project whose objective is to capture all aspects of natural language interactions to some degree, rather than single aspects to a high degree and others not at all. HAM-RPM's environment is a simulated world allowing for gradable concepts. SWYS (Say What You See) [Hanssmann (1980)] aims at interfacing dialogue systems with 2-dimensional natural scenes. Rhodes & Klinger (1977) describe a system SKETCH which serves to interactively modify graphic displays by means of imprecise natural language instructions. An overview of recent natural language interfaces appears in SIGART (1977). Wahlster (1977) presents an excellent review of approaches for representation of fuzzy information in natural-language AI systems.

1.5.1.2 Methodology -

Emphasis in AI has shifted from developing general intelligent systems [Newell, Shaw, Simon (1960), Newell & Simon (1961)] to the development of expert systems for specific problem domains [Buchanan et al. (1969), Shortliffe et al. (1973), Feigenbaum (1977), Duda et al. (1978)]. Feigenbaum (1980) emphasizes that AI systems should model the best knowledge sources available: human experts. AI is to a large extent an empirical science [Newell & Simon (1975), Nilsson (1974), Raphael (1976)].

The PINPOINT model is designed to exhibit behavior comparable to certain aspects of human communication. No attempt is made, however, to obtain this behavior the same way humans do. This is typical of AI-systems [Boden (1977)]. The objective is to have a model which can be tested and explored to gain better insight into fundamental requirements for flexible communication systems. [Stark & Dickson (1966)]. The model attempts to provide useful representations and interpretations of descriptions and at the same time to adapt to human representations slightly better than previous systems. It should be viewed as a contribution to expert systems for limited domains.

1.5.1.3 Tools -

AI systems are realized on computers [Nilsson (1971), Winston (1977)]. High-level computer languages have been developed [McCarthy (1960), Hewitt (1969, 1972), Sussman & McDermott (1972), McDermott & Sussman (1972), Kling (1974), LeFaivre (1974a, 1977)] which allow for implementation of rather complex systems [Bobrow & Raphael (1974), Hahn et al. (1979)].

We take advantage of experience that has been gained with such systems and investigate how linguistic information can be represented adequately and integrated into existing AI-languages. Linguistic symbols have been treated as atoms in previous AI systems and their detailed denotation has been neglected in programming languages. As a consequence, we can refine existing mechanisms and provide additional tools which may prove useful for the implementation of even higher-level systems.

1.5.2 Cognitive Psychology -

Cognitive psychology investigates how people acquire, organize, and use knowledge [Neisser (1967, 1976), Palmer (1978)]. Specifically, experiments are performed which demonstrate how people characterize [Bartram (1973), Rosch & Mervis (1975), Carroll (1979a,b), Hoffmann (1980)] and interpret [Norman & Rumelhart (1975), Palmer (1975, 1977)] perceived objects.

For our purposes, cognitive psychology research is particularly useful if it is concerned with case studies of cognitive processes [Newell & Simon (1972)] rather than with statistical evaluation of interindividual observations. This is, because we want to model a single hypothetical entity which performs certain cognitive functions well rather than a non-existing "average entity". Kelly's (1955) theory, for example, does not assume that comparable processes in different individuals must have the same structure.

Cognitive psychology also investigates questions of concept discrimination in connection with memory size considerations [Miller (1967)] and the use of linguistic means to denote cognitive concepts [Cliff (1959), Kochen & Badre (1974), Hersh & Caramazza (1976), Zimmer (1980a,b)]. Research in psycholinguistics investigates how language develops in humans, how it is used, and what are its features and problems for communication [Lenneberg & Lenneberg (1975)]. Empirical results from this research help us design a vocabulary and setting up rules for manipulating descriptions.

1.5.3 Linguistics -

In the 1950's information processing by computers was almost exclusively numerical. Problems of any type were cast into arithmetic shells and then solved by standard mathematical methods. In the 60's, symbolic information processing was advanced when it became clear that many problems could not be fit in a natural way into mathematical structures for which adequate computational tools existed [Newell & Simon (1972)]. When AI moved away from "toy" problems to "real-world" problems in the 70's it became apparent that many situations are not well enough defined to be captured by symbols, unless the symbols are given an interpretation reflecting the nature of the phenomena they describe [Kickert (1979)]. Research in linguistics investigates how people use natural language (NL) for communication. NL has characteristics substantially different from mathematical languages. In particular, NL is acquired not by definitions but by use. Therefore the meaning of NL expressions cannot be derived from definitions but from context. As a consequence, meaning of NL expressions may not be determined to the same extent as we expect from mathematical language [e.g. Kempson (1977)].

Traditionally, this indeterminacy has been viewed as a disadvantage of natural languages compared to mathematical languages. This view is certainly justified if we want to use language to describe clear-cut and well understood phenomena. However, if we want to use it to describe phenomena which are not as clear-cut, natural language may be better, since it reflects the knowledge about the phenomena more adequately [Popper (1965)]. Models of indeterminate NL expressions have been developed [Zadeh (1972), Lakoff (1973), Kay (1979)] and applied to artificial systems [Mamdani (1976)]. The "linguistic approach" to computer modeling as described by Zadeh was enabled by the calculus of fuzzy sets [Zadeh (1976b), Bonissone (1980)].

1.5.4 Fuzzy Set Theory -

Zadeh (1965) proposed to generalize Boolean set theory in such a way that elements do not have to be either perfect members or perfect non-members of a set, but can be members of a set to a certain degree. The full and partial members form a "fuzzy set" [Zadeh (1965-1979), Bellman et al. (1966), Gaines (1976)]. Fuzzy set theory and fuzzy logics have been used to model non-clear-cut situations encountered in "real-world" problems [Bezdek & Castelaz (1977), Bonissone (1979a,b), Buckles (1979), Freksa (1980b), Goguen (1974), Haar (1977), Imaoka & Sugeno (1979), Jain & Nagel (1977), Kickert & Koppelaar (1976), MacVicar-Whelan (1976), Mamdani (1976), Rieger (1980), Wahlster (1979), Zadeh (1976a)]. Extensive bibliographies can be found in Gaines & Kohout (1977), Kandel & Byatt (1978), and Gupta et al. (1979a).

Computer systems have been developed to represent and utilize fuzzy sets and to support fuzzy reasoning [Adamo (1980), Bonissone (1979a), Freksa (1980a), LeFaivre (1974a, 1977), Mandic et al. (1980), Shaket (1975), Umano et al. (1975), Wenstop (1976)]. The approach presented in this thesis is based on the fuzzy set theoretic interpretation of linguistic labels and derives general rules which serve to simplify the "linguistic reasoning" process.

1.5.5 Philosophy -

A pragmatic approach to representation of and communication about objects leads to issues which have concerned philosophers for a long time. Dreyfus (1979) doubts that thought processes can be formalized to the extent that artificial intelligence becomes possible. His conclusions are mainly based on empirical "evidence" from existing implementations and on (partly invalid) assumptions about neurobiology rather than on logical inconsistencies of AI approaches. Therefore, his skepticism should not keep us from looking for more adequate ways of assimilating intelligent behavior.

Berkeley (1709) distinguished between objects and their representations, Frege (1892) between "Sinn" (sense) and "Bedeutung" (denotation) of representations. Russell (1905) said that denoting phrases never have meaning in themselves, but the propositions in which they occur have meaning. Wittgenstein (1922, 1951) argued that the meaning of a word is the way of its employment such that if we talk about "different meanings" of a word, we think of different functions.

With regard to precision of language, Frege (1879) judged word language to be inadequate and demanded a system of symbols which is free from every ambiguity. Wittgenstein (1922) attempted to develop such an ideal language without vagueness but he conceded later [Wittgenstein (1953)] that imprecision in language can be useful. Russell (1923) criticized traditional logics for assuming that there could be a precise language for representing something that is real. Popper (1976) advocates that we never should use more precision than the particular situation asks for, but this view is not universally accepted [Haack (1979)].

1.6 Thesis Overview

Chapters 2 - 4 deal with theoretical aspects of object description and interpretation. Chapter 2 develops a system of seven object description languages starting with a conceptually simple approach and stepwise revising it to account for aspects which the previous one could not handle. The result is a fuzzy object description language L7 on which the following chapters are based.

Chapter 3 discusses the semantics of L7 in detail. It explains the relationship between linguistic descriptors and their reference objects. It is shown how fuzzy object descriptors may obtain crisp denotation when they are instantiated in a particular context and how fuzziness may be utilized for representation of subjective concepts. Different types of referential meaning for descriptors are discussed and a hybrid representation scheme for the descriptors is proposed.

Chapter 4 deals with the interpretation of fuzzy object descriptions. Different aspects of compatibility between descriptors are presented. The object identification process is discussed as a series of possibility restrictions. Qualitative feature matching is introduced as an efficient method for obtaining informative matching results. A discussion of the interpretation of descriptions referencing sets of objects follows.

Chapters 5 - 7 deal with applications of our approach. Chapter 5 describes how linguistic labels are implemented in the computer language L-FUZZY, how L-FUZZY queries are interpreted, and how the system responds to the queries. Chapter 6 discusses sample applications for linguistic modeling in the "soft sciences", in person - machine interaction, and as a general communication tool. Chapter 7 concludes the dissertation with a summary, a comparison with other computer systems, and with a comparison to properties of human communication.

In the appendix, a glossary of terms and an introduction to the AI-language FUZZY -- on which the dialect L-FUZZY is based -- are presented.

CHAPTER 2

A HIERARCHY OF OBJECT DESCRIPTION LANGUAGES

> "All traditional logic habitually assumes that precise symbols are being employed. It is therefore not applicable to this terrestial life, but only to an imagined celestial existence."
>
> Russell (1923, p.88f.)

In this chapter, we will construct a language to communicate about objects in complex environments. As we will see, complexity considerations make it desirable to abandon very detailed and precise descriptions at the risk of inducing ambiguity or inaccuracy. We will develop this language from a hierarchy of languages whose basic level requires accurate and crisp descriptions. On the higher levels we gradually relax these requirements and allow for approximate and fuzzy descriptions. On the top level, detailed agreement of vocabulary and context between communication partners using this language is not required. Thus, the resulting object description language "conceptually decouples" issuer and interpreter of a description. At the same time, a description is coupled to the specific context of the described object and is meaningful especially in this context. This language is suitable for much more complex situations than the lower-level languages.

In essence, the criterion for superordination of a language in the hierarchy is the ability to describe objects in more complex domains while the efficiency of the communication system can be roughly maintained. Some of the intermediate levels may appear rather arbitrary and artificial. They have been introduced mainly as stepping stones to facilitate introduction of the higher levels and to facilitate comparison between the top and bottom levels, by modifying single language features individually. A different structure could have been used. The higher-level languages are developed by taking advantage of structure that is inherent in the world of discourse.

2.1 Notation

We will employ a notation similar to PRUF notation [Zadeh (1978b)] to describe the object description language hierarchy. In this notation, an object can be characterized by means of a property list [Weissman (1967)], i.e.,

(OBJECT [<property 1> = <value 1>]
[<property 2> = <value 2>]
:
[<property n> = <value n>])

The interpretation of this characterization is that it denotes an object which maintains properties <property 1>, <property 2>, ... <property n>. This can be expressed differently: the possibilities, to which object the description refers, are restricted by the particularizations [Zadeh (1978b)] <property x> = <value x>, $1 \leq x \leq n$.

This representation presupposes that the class OBJECT is given. Accordingly, we can restrict subclasses of objects, e.g.,

(HOUSE [<property x> = <value x>])

Subclasses may be obtained by possibility restriction of superclasses or may be introduced independently. The set of objects belonging to the reference context will be characterized by the same notation, i.e.,

(CONTEXT [<property 1> = <value 1>]
[<property 2> = <value 2>]
:
[<property n> = <value n>])

The types of properties and values that can be used for particularization will be introduced in the following discussion of the various levels of the hierarchy.

2.2 Seven Hierarchy Levels

> " ... both precision and certainty are false ideals. They are impossible to attain, and therefore dangerously misleading if they are uncritically accepted as guides. The quest for precision is analogous to the quest of certainty, and both should be abandoned.
>
> "I do not suggest, of course, that an increase in the precision of, say, a prediction, or even a formulation, may not sometimes be highly desirable. What I do suggest is that it is always undesirable to make an effort to increase precision for its own sake -- especially linguistic precision -- since this usually leads to loss of clarity, and to a waste of time and effort on preliminaries which often turn out to be useless, because they are bypassed by the real advance of the subject: one should never try to be more precise than the problem situation demands."
>
> Popper (1976, p.24)

The communication paradigm which we will use for all seven levels L1 - L7 of this hierarchy is the following: in a given world W, a describer X constructs a verbal description of a particular object $O \in W$; an interpreter Y analyzes this description to identify target object O.

In this hierarchy, we will describe first the basic, most primitive level. On each of the higher levels we will introduce some features which allow the language to be used for more complex situations. The objective is to keep communication efficient with respect to memory requirements and computational effort. We will show how we can compensate for ambiguity introduced by more efficient methods.

2.2.1 L1: "zero-dimensional" Symbolic Object Labeling -

Suppose, X and Y live in a world W of n distinct objects. X wants to be able to point out to Y any object O in W. Both, X and Y can look at the objects in W, but they only can communicate about the objects by means of words.

The conceptually simplest way [compare Kempson (1977, pp.12f)] of communicating under these constraints is to have distinct labels for all objects in the world. We will assume here that these labels are random, i.e., there is no known systematic relationship between the labels and their corresponding objects. Then, X simply utters the label and Y identifies the object by comparing the label with the labels corresponding to the objects in W. X's message will match exactly one object label.

Using the notation from section 2.1, this object description has the form

(OBJECT [name = <label>])

The only property that can be used as descriptor in L1 is a name. An example for an object description in L1 is [Valentin (1929)]:

(OBJECT [name = Wrdlbrmpfd])

We must assume that the links between labels and their corresponding objects are not elements of W -- otherwise there was a systematic relationship between the objects and their labels. For this reason, X and Y each must have a memory for the labels and must be able to associate with each object the corresponding label. The size of the vocabulary (i.e., the number of object labels) is proportional to the number of objects in the world. The memory required to store the vocabulary is proportional to the size of the vocabulary. Also, the label selection time for X and the object identification time for Y are proportional to the number of objects in the world (since we assume non-systematic assignment of labels).

L1 allows for perfect communication between X and Y. Each description is accurate and unambiguous by definition. However, L1 is of practical use only for small worlds because of memory requirements, vocabulary acquisition effort of X and Y, and label search time. Size of memory, label acquisition time, and search time for label and object could be reduced if there was a systematic approach to the object labeling process.

2.2.2 L2: One-dimensional Absolute Object Labeling -

In this model we introduce a systematic approach to the label assignment process. Instead of agreeing on a label for each individual object, X and Y will agree on a system (i.e., a rule or a set of rules) by which labels are assigned to objects. A single feature dimension which is common to all objects in W will be chosen to label the objects. Then, X can replace his declarative data base (the random vocabulary) by a procedure which generates a label by measuring the feature value on this feature dimension. The interpreter, Y, also can replace his vocabulary by a procedure, provided that there exists an inverse to X's feature measuring function. This means that he needs a way of matching the feature value reported by X with the actual object.

Using our notation, an object description in L2 has the form

(OBJECT [<feature> = <value>])

where <feature> designates a dimension which is defined for all objects in W. An example for an object description in

L2 is

(OBJECT [height = 168 cm]).

The label generation time has become independent of the number of objects. To set up a common vocabulary, X and Y only must agree on a rule rather than on each individual label -- this also is independent of the number of objects. The object identification time is still proportional to the number of objects in the world, if we assume that the objects in W are not arranged in a systematic fashion by their feature value.

L2 has the advantage over L1 that memory requirement and label selection time have become independent of the cardinality of W. This means that it should be more useful for large worlds. A severe disadvantage is that this labeling method does not guarantee unique labels for distinct objects, if we assume finite feature value resolution. Therefore, Y cannot uniquely identify the object which X described, in general. He can identify a class of objects which must contain the object in question. L5, L6, and L7 will deal with this problem.

The object identification effort could be reduced if there was more structure in the object arrangement to allow for a more directed label matching process.

2.2.3 L3: 1-D Relative Object Labeling -

In this model we introduce the concept of "context" to increase efficiency of communication. In realistic situations, X and Y do not want to communicate about all the objects which they are able to communicate about; rather, in any given conversation they focus on a specific subset of objects.

Let us assume, therefore, that X and Y have agreed on a subset $W_x \subset W$ which they refer to. Typically, $|W_x| \ll |W|$. This assumption reduces the complexity of the problem if we describe the object feature in relative terms instead of absolute. For this purpose, X sorts the objects in W_x according to their one-dimensional feature value. Now, he can describe the object simply by referring to its ordinal number. To decode this message, Y must sort the objects as well, before he will be able to interpret X's description.

An object description in L3 has the form

(OBJECT [<feature> = rank <ordinal number>])

Implicitly or explicitly, a context must be specified:

(CONTEXT [<subset> = <set descriptor>])

The feature rank refers to this context. Note that the context feature brings back the advantages of a small world while maintaining the capability of communicating about all objects in a large world.

An example for an object description in L3 is:

(OBJECT [height = rank 3])

with

(CONTEXT [location = room 275])

which denotes the set of objects in room 275 which rank third with respect to their height.

The sorting process requires O (m * log m) comparisons for $m = |W_x|$ and a memory size proportional to m. The sorting effort pays off if many references have to be made to objects in the same subset W_x, because the sorting must be done only once. The information which has to be transmitted for each object reference from X to Y is reduced, because less information is required to distinguish m objects than to distinguish n objects if $m < n$, $n = |W|$. A disadvantage is, however, that exhaustive comparison of the objects in the subset is necessary to obtain the object's ordinal number.

2.2.4 L4: 1-D Context-adaptive Absolute Object Labeling -

To combine advantages of absolute and relative object labeling we will introduce "context-adaptive object labeling". Rather than sorting all objects in W_x, we select from W_x the object with highest feature value and the object with lowest feature value. These two extreme values define the end points of an absolute scale on which the object feature is measured.

An object description in L4 has the form

```
(OBJECT [<feature> = <class> <scaled value>])
```

with

```
(CONTEXT [<subset> = <set descriptor>]
         [feature values = <natural number>])
```

where <class> refers to a feature dimension with a discrete set of ordered values and "feature values" specifies the number of values that <feature> may assume on the corresponding scale. It is the inverse of feature granularity.

An example for an object description in L4 is:

```
(OBJECT [height = size 7])
```

with

```
(CONTEXT [location = shelf 4]
         [feature values = 10])
```

Here, the feature value is given in absolute terms as in L2. However, the scale for this value is given by the context and the feature scale partitions. Thus, it is interpreted as a relative value, in effect.

The effort of setting up the scale results in the advantage over the previous model that the objects do not have to be sorted completely; m comparisons are necessary to determine the end points of the scale. In addition, we combine the advantages of measuring in absolute terms with the small amount of information to be communicated from the relative measuring method.

All levels of the hierarchy discussed so far are based on assumptions which are not very realistic for practical applications:

1. X and Y must completely agree about the situation context. In applications in which the context is "the objects in room 275" this may be tolerable; but if X and Y want to communicate about the objects in a less well-defined context, e.g., the set of visible objects, there may be a slight disagreement about the members of the set between X and Y, because X and Y may have different points of view. If X and Y want to communicate about all recognizable objects, the disagreement will be even greater, unless they have identical perception.

2. the vocabulary that could be used by X and Y must refer to precisely defined feature values; this can make it difficult to describe objects which were not anticipated during the design of the vocabulary.

3. the meaning of X's and Y's vocabulary must be identical; this means that before being able to communicate, X and Y must agree on the meaning of the terms they use.

The next model will introduce some flexibility for the use of the vocabulary and accommodate disagreements in context and vocabulary of X and Y.

2.2.5 L5: 1-D Linguistic Object Labeling -

In this model, not the two most extreme feature values of the objects in the context W_X must be determined, but instead an object with "comparatively high" feature value and an object with "comparatively low" feature value with respect to the other objects in W_X. These values serve to define a "fuzzy scale" on which the feature of the object to be described will be measured. The scale is partitioned into overlapping fuzzy subsets* which correspond to linguistic labels. This implies that several labels become applicable for the same feature value. Correspondingly, a given label may include more possible feature values than before.

----

* For an introduction to the theory of fuzzy sets see, for example, Ragade & Gupta (1977).

To compensate for lost precision in the intended meaning of a label due to this relaxation of requirements, we introduce the concept of modifiers of fuzzy sets [Zadeh (1972), Lakoff (1973)]. Modifiers act as operators on fuzzy sets and yield a larger variety of fuzzy concepts. Modification is done by shifting, sharpening, or fuzzification of fuzzy sets. Such modification cannot be done in a natural way for crisp sets or values. A similar effect could be obtained by enlarging the number of labels but with modifiers we maintain the advantages of a small label set.

An object description in L5 has the form

(OBJECT [<feature> = <linguistic value>])

where <linguistic value> refers to the label of a possibility distribution. A possibility distribution denotes an arbitrary element of a fuzzy set. Thus, the linguistic value "tall" has a possibility of referring to any particular height value that belongs to the fuzzy set of tall heights.* For example, an object description in L5 is:

(HOUSE [height = tall])

----

* The notation used here deviates from the PRUF notation used by Zadeh (1978b). Possibility distributions and their representation are discussed in more detail in section 3.4.

with

(CONTEXT [location = vicinity])

Here, "HOUSE" denotes a subclass of the class "OBJECT" in the previous examples, "vicinity" specifies a fuzzy location.

This is the first model in our language hierarchy in which we do not maintain the requirement of total agreement between the communication partners with respect to their reference sets of objects and their vocabularies [compare Bellman & Zadeh (1977)]. If this disagreement gets too strong, misinterpretations of the object descriptions become possible. In the next model we will show how we can compensate for this deficiency.

2.2.6 L6: Multi-dimensional Object Labeling -

Up to now, objects were described in terms of a single feature, in all models. This was done mainly for simplicity of presentation -- it is not a very realistic approach. In practical situations it is not easy or possible to find a single feature that discriminates well across a large number of objects, as we know from pattern recognition research. In our paradigm, the feature would have to discriminate well across all n objects of W. There is another reason why a single feature is not attractive for object descriptions: it requires a label set with cardinality well above m, the number of objects in W_x, to obtain small classes of referenced objects. Thus, we will introduce multi-feature object descriptions in the present model.

An object description in L6 has the form

```
(OBJECT [<feature 1> = <linguistic value 1>]
        [<feature 2> = <linguistic value 2>]
        :
        [<feature k> = <linguistic value k>])
```

For example, an object description in L6 is:

```
(BOOK   [color = red]
        [size = quite small])
```

The objects are described in terms of a variety of features. In general, not all objects of a set share all properties by which they can be characterized. Therefore it does not seem appropriate to replace the feature value of the previous models by a fixed-format feature vector. Instead, descriptions will be constructed from a set of characteristic and well-discriminating features and represented by a free-format property list.

Multi-dimensional object labeling drastically enhances the specificity of descriptions. The number of objects that can be discriminated increases exponentially with the number of feature dimensions. This allows us to maintain a small label set and still deal with complex worlds.

A drawback may be that several features have to be measured and described; this enhances the task complexity, but not necessarily the description generation and interpretation times, because all dimensions could be dealt with in parallel.

2.2.7 L7: Subjective Object Labeling -

In L5 we accounted for the fact that many feature descriptors are applicable to different objects to a higher or lower degree. We did not require a rigorous method for defining the fuzzy scale for the linguistic labels, and in fact, the fuzzy sets corresponding to the same linguistic labels could be slightly different for X and Y. The reference concepts for X and Y have become subjective, or in other words, X and Y have become "conceptually decoupled".

Because of this and because of the fact that objects are not distributed homogeneously in their multi-dimensional feature space, we need a method allowing for correct identification of the described object. The object description must provide information for a correcting mechanism which allows the interpreter to identify the target object even if some of the object descriptors do not completely agree with the interpreter's concepts.

In L7, this is done by redundant object descriptions [compare Critchley (1975), Reball (1978)]. Similarly as in coding of messages, where we can correct for noise in the transmission of a message [Shannon & Weaver (1949)] by increasing the Hamming distance of a code [Hamming (1950)], we can compensate "conceptual distance" between describer and interpreter by enriching the descriptions. The interpreter then can be less rigid in the interpretation of single descriptors but relies on the coincidence of several more or less applicable descriptors instead.

A description in L7 has the form

```
(OBJECT [<feature 1> = <fuzzy linguistic value 1>]
        [<feature 2> = <fuzzy linguistic value 2>]
        :
        [<feature k> = <fuzzy linguistic value k>])
```

An example for such a description would be

```
(BOOK   [author = Mao Tse Tung]
        [title = bible]
        [color = red]
        [size = quite small]
        [cover = plastic])
```

Thus, we trade conciseness and crispness of description for flexibility in interaction and variability in background of the communication partners.

2.3 Example Comparing L1 - L7

This example demonstrates the use of L1 - L7 for the description of single objects. The same scene of objects, a fish scene, is used in all cases for comparison purposes. This makes the example somewhat artificial, since we have to use a simple scene to demonstrate L1. The advantages of L7, on the other hand, come to bear only in more complex situations. The basic effects should become clear, however. Occasionally, L1 will serve as meta-language for explanation of the examples.

Fig. 2.1: Fish scene W

L1: Let W be a set of a dozen fishes of different length. They are labeled (randomly) with unique labels from the set of letters {a, b, ... , z}.

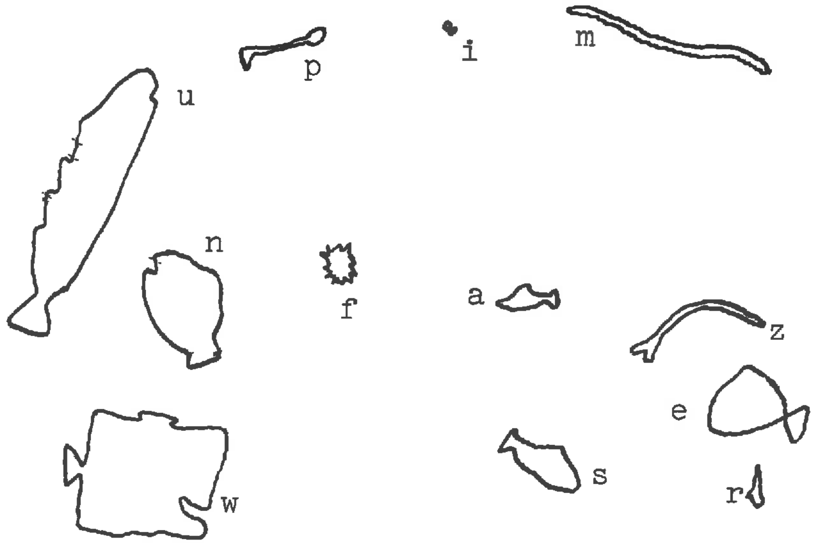


Fig. 2.2: Labels corresponding to fishes in Fig. 2.1

X wants to point out the fish labeled "e" to Y. He "looks up" this label and utters "e", or, more formally,

(FISH [name = e])

Y looks up the fish corresponding to the label and thus identifies the fish pointed out by X.

L2: X measures the length of the fish and reads the measurement "55 mm" to Y, or

(FISH [length = 55 mm])

Y generates a template of length 55 mm and compares this template with each fish. One of the fishes which match the template's length is the target fish.

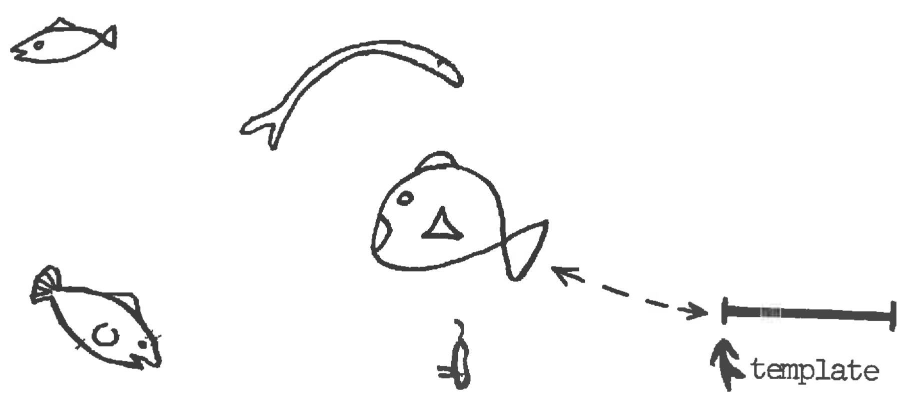


Fig. 2.3: Subset W_X of scene and reference scale.

L3: Let $W_x \subset W$ be the subset of fishes in the right lower part of the scene W (Fig. 2.1) and let's assume, X and Y have agreed to converse only in the context of W_x. X sorts these fishes in his representation in increasing order of length and refers to a particular fish in terms of its ordinal number in this sequence, "the 4th",

(FISH [length = rank 4])

Fishes of the same length obtain the same ordinal number. Y sorts the fishes in his representation accordingly and matches X's description against the ordinal numbers in his system. One of the matching fishes is the fish described by X.

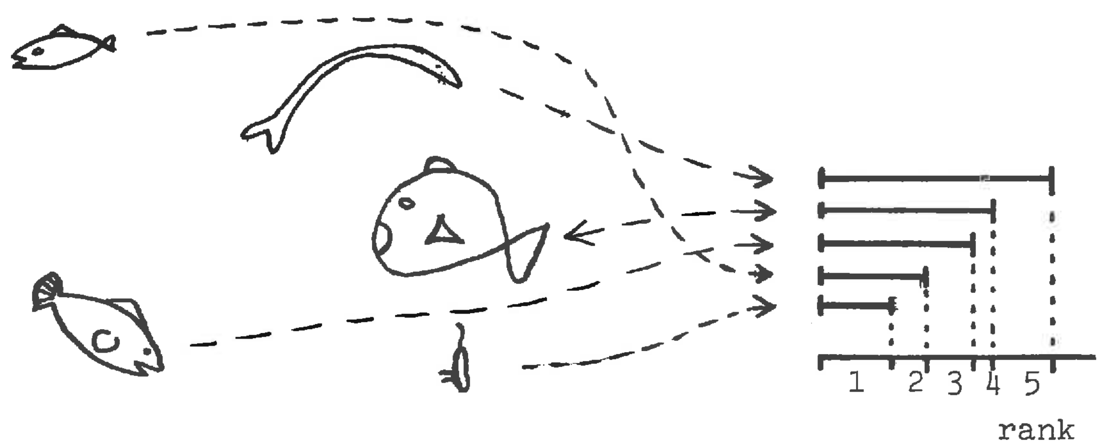


Fig. 2.4: W_x is sorted, its elements are ranked.

L4: X selects the shortest fish (fish "r") and the longest fish (fish "z"). Their end points become pivot points for a linear scale (say from 1 - 10), such that fish "r" corresponds to size "1" and fish "e" corresponds to size "10". Now he compares the target fish "e" to this scale and obtains "size 7".

(FISH [length = size 7])

Y creates a scale accordingly and obtains a range of lengths for which "size 7" applies. The target fish is one of the fishes whose length is in that range.

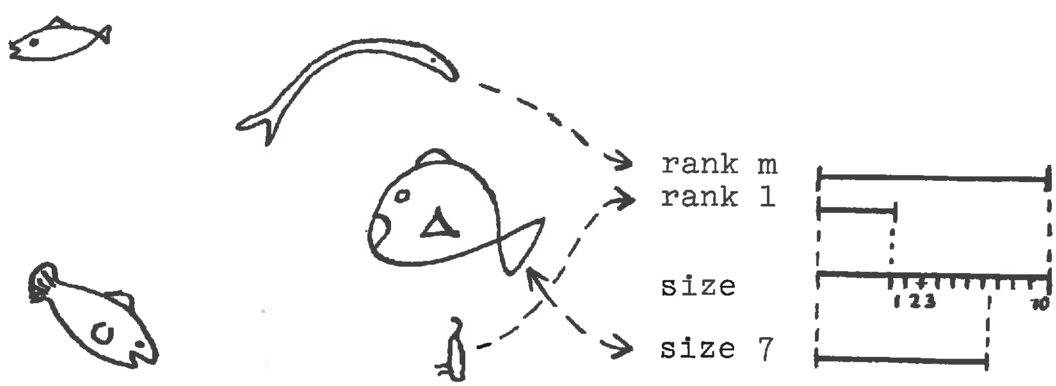


Fig. 2.5: Context-adaptive scale for objects.

L5: X selects a "relatively short" fish ("a") and a "relatively long" fish ("z"). Their lengths become reference values for a fuzzy scale of 3 overlapping fuzzy sets, labeled "short", "medium", "long". Two modifiers, "slightly less than" and "slightly more than" serve to shift these fuzzy sets to the left or to the right, respectively.

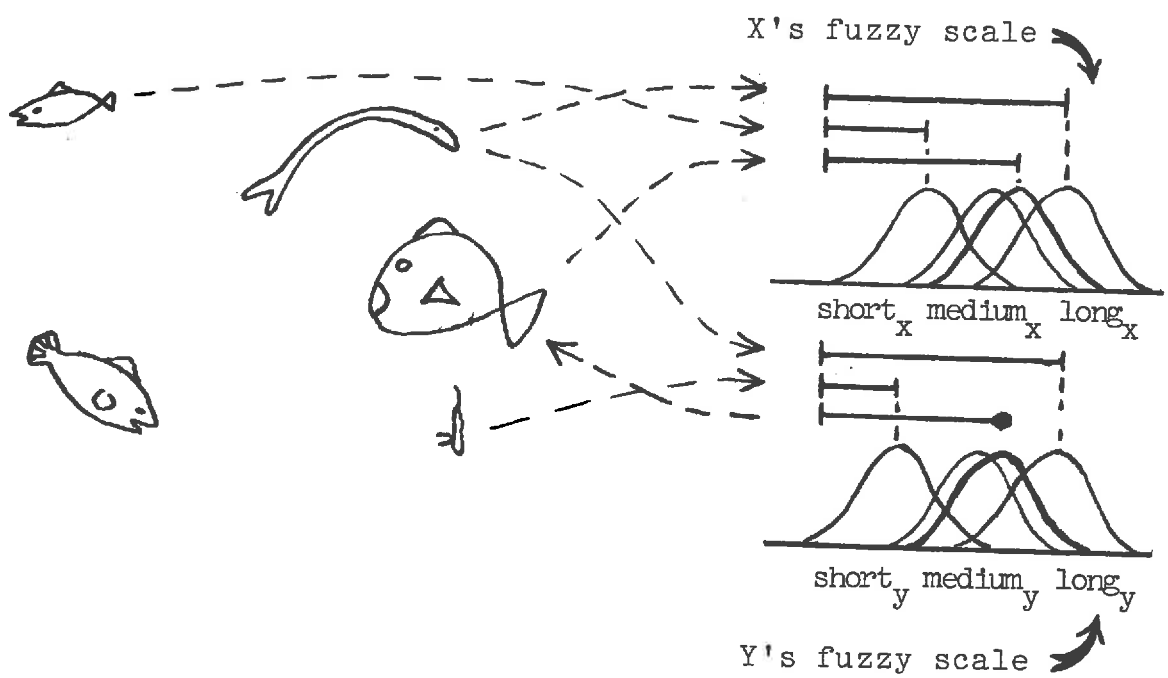


Fig. 2.6: Fuzzy feature scale for W_x.

X selects the label of the fuzzy set which best fits the target size, in our case "slightly more than medium", i.e., fuzzy set "medium" modified by a right shift:

(FISH [length = slightly more than medium])

Y also selects a "relatively short" (fish "r") and a "relatively long" fish ("z") -- which are different from the ones which X selected -- and constructs a fuzzy scale with the same labels, but with different possibility distributions. Y then selects the possibility distribution denoted by the label "slightly more than medium" and obtains a fuzzy set of possibly referenced fishes which will include the target fish, unless Y's selection of a "relatively short" or a "relatively long" fish differ substantially from X's selection.

L6: Now let W be the set of all fishes on earth and $W_X \subset W$ the subset of the dozen colored fishes which we called W up to now (Fig. 2.1). Suppose that besides measuring the length of a fish, X and Y are able to determine color and diameter. Thus, X can use up to three dimensions to describe the target fish. Figures 2.7 and 2.8 show how the fishes are arranged in a subspace of in this 3-dimensional feature space.

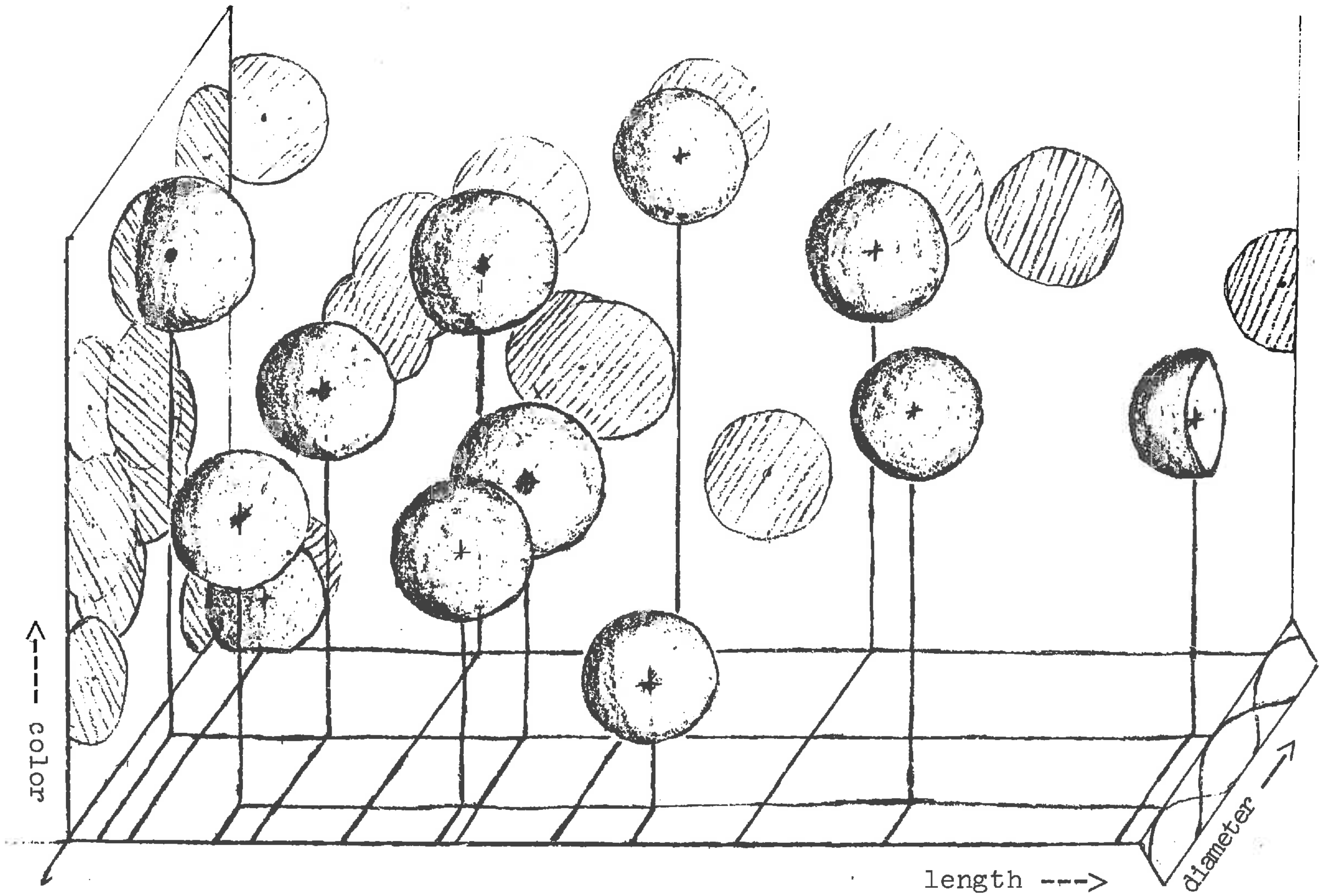


Fig. 2.7 Fuzzy feature descriptors in a 3-dimensional feature space.
The spheres represent 3-dimensional fuzzy sets;
the shaded circles represent their projections onto 2-D feature planes.

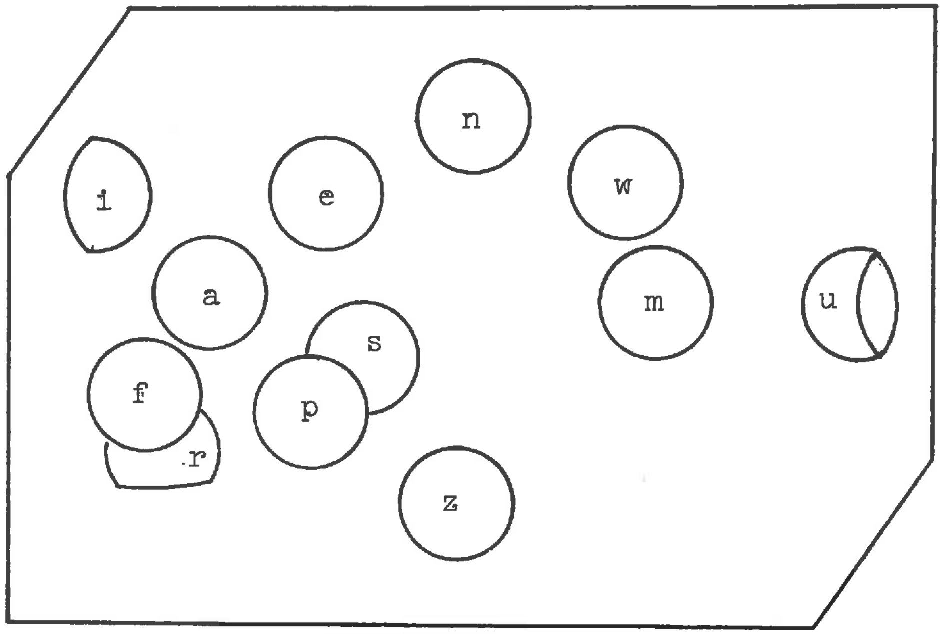


Fig. 2.8 Object names corresponding to fuzzy feature descriptors in Fig. 2.7

All feature dimensions are treated independently in the same way as in the previous example for L5. X utters

(FISH [length = medium]
[color = orange])

and Y constructs a 2-dimensional feature space for his fuzzy scales of length and color and obtains a fuzzy subspace which contains the objects of W_x to which X's characterization can possibly refer.

L7: W_X is not any more the well-defined, crisp set of fishes; instead, it is the fuzzy set of fishes "in the vicinity of X and Y". Now, X uses more features for his object characterization

(FISH [length = slightly more than medium]
[color = red]
[diameter = thick])

Y may have a different understanding of "less than medium" length, "red" color, and "thick" diameter in this fuzzy context, than X, but he agrees with X that "slightly more than medium" length denotes a shorter length than "long" and that "red" is closer to "orange" than to "green", etc. Thus, he will be able to find the target object (or a superset) if there are no objects to which this description would fit better according to his own conception.

2.4 Summary And Conclusions

The following table summarizes comparative performance values for L1 - L7 and states the main trade-offs involved:

	vocabulary acquisition	memory requirement	label search	matching	advantages	disadvantages
L1:	O(n)	O(n)	O(n)	O(n)	accurate, unambiguous	prohibitive for large worlds
L2:	const.	const.	const.	O(n)	only small memory required	single property for all objects, ambiguity possible
L3:	O(m*log m)	O(m)	O(m)	O(m)	able to deal with very large world	objects must be sorted
L4:	O(m)	const.	O(log m)	O(m)	objects don't have to be sorted completely	more ambiguity possible
L5:	O(m)	O(m)	O(log m)	O(m)	no common reference of vocabulary necessary	misinterpretations possible
L6:	$O(k* \sqrt[k]{m})$	$O(k* \sqrt[k]{m})$	$O(k* \sqrt[k]{m})$	$O(k* \sqrt[k]{m})$	small vocabulary complex world	several dimensions to deal with
L7:	$O(r*k* \sqrt[k]{m})$	$O(r*k* \sqrt[k]{m})$	$O(r*k* \sqrt[k]{m})$	$O(r*k* \sqrt[k]{m})$	avoids ambiguity, enhances flexibility	longer message required

Fig. 2.9 Comparison of performance criteria as for L1 - L7 as functions of size of domain n, size of context m, number of feature dimensions k, and redundancy factor r.

The hierarchy of object description languages L1 to L7 contains various levels of communicating about objects: from a very rigid language which is unambiguous independent of the context in which it is used and suitable only for small domains and well-coordinated communication partners to a non-rigid language for ill-defined domains which can be disambiguated only by the context and which allows for some discrepancy in interpretation between the communication partners. Proceeding from L1 to L7 can be viewed as "conceptually decoupling" the receiver from the transmitter of a communication message. At the same time, descriptions on higher levels become less meaningful when considered out of context. Thus, descriptions become "conceptually coupled" to a particular context, in the higher-level languages.

The few constraints on feature labels that must be accommodated in L7 allow that one of the communication partners is a person, even if we cannot determine the exact denotation of his or her linguistic labels. Only the relative order of feature labels must be known. Thus, the use of linguistic labels from human language in our model does not presuppose that any person has the same interpretation of that label.

The approach is based upon the contention that no two individuals have identical representation (and therefore interpretation) of a linguistic label and that one individual may have different interpretations of a linguistic label in different situations [Critchley (1975)]. Nevertheless, communication between individuals by means of linguistic labels is possible.

CHAPTER 3

SEMANTICS OF DESCRIPTORS IN L7

In chapter 2, we developed a family of description languages L1 - L7 without stating in detail what the descriptors mean. In the present chapter we will explain what the descriptors represent and how they are to be interpreted. From now on we will mainly consider L7, since L7 includes all interesting aspects of the lower-level languages L1 - L6. Our discussion will be limited to the relationship of reference between an object description and the object (or set of objects) which it denotes [Russell (1905)]. This is important for identification of the described object (the "target object"). The reference relationship is introduced as a gradable, relative concept. We discuss how graded reference is defined, represented, and manipulated in our approach.

3.1 Descriptors Indicate Possibility

Suppose you are asked to pick up a person at the airport whom you have never seen before. You are given a verbal description of that person which you will use for his or her identification. If you are lucky, the description suffices to discriminate your "target person" from the other people you see at the terminal. Perhaps it is even redundant, i.e., it contains more features than you would need to identify the person. These extra features are useful to increase your confidence in the result of the inference process that lead to the identification of the person.

What is the nature of descriptive information? As we have seen in chapter 2, feature characterizations have different effects from object labels. The information they carry does not code the associated objects but it selects a subset of the universe of discourse.*

A universe of discourse can be segmented according to a variety of criteria. In particular, we may have <u>probabilistic</u> information about object features and we may have <u>possibilistic</u> information [Gaines & Kohout (1975)] about object features. In the following discussion we will argue that in the absence of probabilistic information, an

----

* Some codes are designed in such a way that they exhibit properties of feature descriptions.

object description conveys information about possible feature values that the target object may exhibit. We will first investigate characteristics of crisp descriptions and subsequently generalize to account for the fuzzy case.

3.1.1 Restriction Of Reference Set -

Consider the following person description[*]:

> "A man who is between 180 and 190 cm tall
> and weighs between 60 and 80 kg."

We can precisiate[$] this description in terms of the intentional meaning representation

```
(PERSON [sex = male]
        [height = >180,190< cm]
        [weight = >60,80< kg])
```

Here, $>x_1,x_2<$ denotes an unspecified crisp value from the interval $[x_1,x_2]$.

----

[*] We shall use the term "description" to mean more accurately "denoting phrase". We shall use the term "characterization" if we want to emphasize the fuzzy or incomplete nature of a denoting phrase.

[$] Compare Gaines (1976), Zadeh (1979d).

What does this description denote? The three descriptors restrict the possibilities of persons which the description may refer to. They have the effect of subset selection from the set of all persons in the given context as indicated in this Venn diagram:

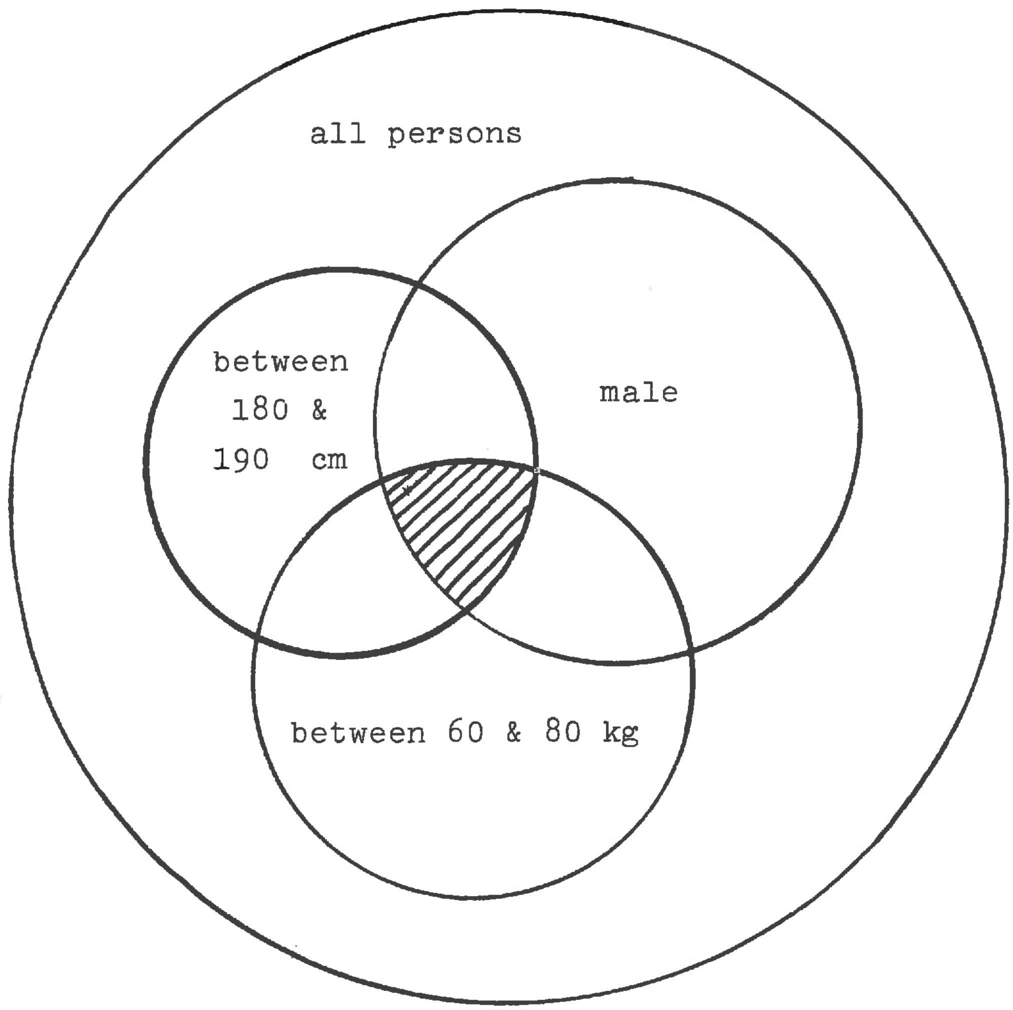


Fig. 3.1 crisp restriction of possible object references

The features of the actual object -- i.e. person, in this example -- will correspond to exactly one point in the shaded area of the Venn-diagram. The description does not specify to which point it refers, but in the absence of additional information (e.g. about the objects in the domain) each point in the shaded area is equally possible.

3.1.2 Possibilistic Vs. Probabilistic Information -

The difference between possibilistic and probabilistic data stems from the different origin of their underlying observations [Zadeh (1980a)]. First of all, if a descriptor may correspond to several possible instances, this means that it is underspecified or imprecise. If a probability value is associated with each of the possibilities, this means that additional knowledge about the possible instances is available or that some law governing the assignment of particular instances is assumed to hold. Second, a prerequisite for using probabilistic (or statistical) information is that the different possibilities are known, whereas conversely no probabilistic information is required for analyzing various possibilities.

We can obtain <u>possibilistic</u> information even if we do not know anything about the process that decides about various alternatives or if we are unable or unwilling to capture some event with high precision. On the other hand, we obtain <u>probabilistic</u> information if we capture events precisely but are unable or unwilling to describe the different observations analytically.

As an example, consider that we are interested in predicting the outcome of throwing a die. A possibilistic analysis would result in the finding that the elements of the set {1, 2, 3, 4, 5, 6} are possible outcomes. A probabilistic analysis, on the other hand, would assume that this are all the possibilities and would require a non-deterministic model which would predict the die's behavior in a statistical sense. Clearly, the two approaches answer different questions: the possibilistic analysis describes what <u>can</u> happen if the die is thrown, or in the case of object descriptors, which object <u>may</u> be referenced; the probabilistic analysis describes how <u>likely</u> an event is going to occur, or in the case of multiple events, how <u>frequently</u> we can expect the different types of events to occur [compare Gupta et al. (1979b)].

As a consequence of these differences, possibility and probability have different properties. In particular, possibilities are superadditive, i.e., the existence of one possibility does not automatically influence the existence of another. In contrast, probabilities are additive, i.e., if one event becomes more likely, one or more other events automatically become less likely. Superadditivity is a desirable property for systems dealing with incomplete knowledge, since addition of knowledge does not necessarily involve revision of all related knowledge.

In summary, possibilistic information requires a minimum of assumptions and is suitable for analyzing unique situations involving incomplete or imprecise information while probabilistic information requires a statistical model and is suitable for analyzing repetitive situations involving unreliable information.

In this dissertation, we focus on possibilistic aspects of descriptions and PRUF descriptions are assumed to convey only possibilistic information. Thus, we are concerned with the question which interpretations of descriptions are possible, although for a good understanding of reasoning it will be important to ask which interpretations are probable, plausible, reasonable, applicable, relevant, consistent, and adequate, as well.

It should be pointed out here that possibilistic information usually represents general rules or knowledge and therefore is more easily available than probabilistic information [compare Zadeh (1978b, p.402)]. A possibilistic statement may not appear as strong as a comparable probabilistic statement, because it may be quite fuzzy; on the other hand, it may be more useful in each particular situation in which it is employed, because it allows for predictions that can be validated by individual instances [compare "incompatibility principle", Zadeh (1973, p.28)].

3.2 Possibility Comes In Degrees

In the last section, we dealt exclusively with imprecise but crisp descriptors and a descriptor - feature reference either was possible or not. In this section, we will investigate what happens if descriptors are fuzzy rather than crisp. We will use the concept of a "possibility distribution" [Zadeh (1978a, 1979c)] to extend the more conventional view of a dichotomy between possible and impossible reference. Specifically, if we are dealing with incomplete information, the dichotomy between all or none possibility becomes useless, since the missing information could make the difference between "entirely possible" and "entirely impossible". This would be incompatible with the "principle of graceful degradation" [Norman & Bobrow (1975), Goguen (1976)] which calls for a gradual degradation of inferences if knowledge degrades gradually.

3.2.1 Fuzzy Descriptors Vs. Crisp Descriptors -

"... all language is more or less vague."

Russell (1923, p.90)

In this section, we replace the crisp person description of section 3.1.1 by a fuzzy description, for example, "A tall man"

or

(PERSON [sex = male]
[height = tall])

Now, "tall" denotes an unspecified value from the fuzzy set FS (height) = tall* rather than from a crisp interval. The most important difference between the fuzzy descriptor

"[height = tall]"

and the crisp descriptor

"[height = >180,190< cm]"

is that there is no sharp boundary between height values which conform with the description and those which do not. Some height values conform better with the description than others.

3.2.2 Types Of Uncertainty -

We should note at this point that, if one value conforms better with a descriptor than another, this does not mean that we are less certain whether or not this value conforms with the descriptor than we are about the other, as Schefe (1980) suggests. The reason is that "tall" is an

----

* The notation used here deviates from the notation in Zadeh (1978b). Details are given in section 3.4.

intrinsically fuzzy concept suitable to characterize the height of an object in each situation in which it is used, sometimes better, sometimes worse [Black (1963)]. For this reason it is inadequate to ask "whether or not" the value in question conforms with the descriptor, but we can ask to what degree this value conforms. Asking "yes - no" questions about fuzzy data cannot elicit the full content of this data, it only can help obtain a crude approximation by reducing the size of the answer label set to 2. Note however, that in many cases, the resulting "yes" or "no" answer still must be considered a fuzzy answer. This implies that in some cases, in which a "yes" answer is given, a "no" answer may have been possible as well.

As an example, consider two detailed descriptions of a person in which there is pairwise agreement for each descriptor, except for one: in one description it may be "height = tall", in the other "height = not tall". Both descriptions may refer to the same person, if the person is neither particularly tall nor particularly small. The law of contradiction does not hold for fuzzy concepts in a rigid sense [Black (1937), Goguen (1969), Lenneberg (1975), Gaines (1976)].

On the other hand, fuzzy data leave uncertainty about precise values [Goguen (1967)]. This uncertainty could be analyzed statistically if a model about statistical distribution of values in the corresponding fuzzy sets was available. However, we must not forget that fuzzy sets and corresponding probability distributions are distinct. In general, there is only a weak relationship between the two which indicates that something is improbable if it is impossible. The inverse relation does not hold. This connection between possibility and probability has been expressed in the "possibility - probability consistency principle" [Zadeh (1978a)].

3.2.3 Fuzzy Laws Of Excluded Middle And Of Contradiction -

The view that "something either is possible or impossible", where "either - or" and "possible - impossible" are viewed as absolutely exclusive, appears not to be useful for reasoning in a world in which we are bound to rely on incomplete and imprecise knowledge. A piece of additional information or slightly improved precision could alter conclusions about the possibility of an event drastically. We are looking for a model that acknowledges that incompleteness and imprecision are unavoidable and which reflects this fact in its knowledge representation and in its reasoning processes. Then, slightly improved or degraded information only should result in slightly refined conclusions as stated in the "principle of graceful degradation" [Norman & Bobrow (1975), Bobrow & Norman (1975), Goguen (1976)].

The law of excluded middle of classical logics can be generalized for the fuzzy case if we interpret "either x or (not x)" in the following way:

> "if x is entirely possible then (not x) is entirely impossible; if x is less than entirely possible, then (not x) is less than entirely impossible."

This can be expressed more formally by

$$\text{Poss}\ (x \mid v) + \text{Poss}\ (\text{not}\ x \mid v) = 1 \qquad \forall\ v \in D,$$

where v is a free variable in the domain D. This means that the principle of contradiction generalizes to a "principle of trade-off" between the consistencies of a label and its negation with a given feature value. Note that the law of contradiction is a special case of the principle of trade-off as stated above, whereas in Black (1937, p.55) the crisp case is not defined [Hempel (1939)].

As an example, consider that we have "x is tall" and "x is not tall" in the same context. According to the "trade-off principle", this means that x definitely has some height, but the height value cannot be expressed precisely with either the label "tall" or the label "not tall". But since both labels are applicable to some degree, x's height must be between "tall" and "not tall".

3.2.4 Concepts Can Be Stretched -

A fuzzy concept capable of representing a collection of values can be likened to a piece of rubber capable of assuming a variety of shapes, some by forceless deformation, some by stretching. The more the concept must be "stretched" to assume a certain value, the less adequate is this concept to characterize the value. The less "force" is required, the more easily it is possible that the concept refers to a particular value.

Following Zadeh (1978a), we deviate from the modal logic model [Hughes & Cresswell (1968)] in which events either are possible or impossible. Instead, we view possibility as a gradable dimension. This view is in harmony with natural language usage of the concept of possibility. For example, we say an event is easily, quite, entirely, marginally, absolutely, hardly possible, or there is a slight, great, small, or very real possibility for a given event to occur.

For example, a detective or a judge uses graded possibilistic information to determine to what extent a suspect's alibi is in conflict with an action of which he is accused. This can be done by examining physical constraints for getting from the location of the alibi in the time between action and alibi. Incidentally, the art of solving criminal cases or other problems exhibiting exceptional circumstances may be rooted in the ability of manipulating possibilistic information while neglecting probabilistic information.

3.3 Descriptors Are Subjective

Descriptor definitions are not assumed to be universal with respect to different describers and interpreters. This means that the same descriptor may be defined in different ways (as discussed in section 3.5) and even if defined by the same method, the feature values of a descriptor do not have to agree for two individuals. In the case of purely artificial communication systems, "subjective descriptor" means that different parts of the communication system may use different denotations for the same descriptors.

For the purposes of this thesis, we will assume that the feature dimensions on which the feature values are based (the left hand sides of the PRUF descriptors) are universal to the communication partners. This means that the same criteria are taken into account to determine a particular feature value in a given communication situation. In order for two individuals to communicate successfully (i.e., "to get an idea across") even if the communication medium (the object description) denotes different things for the two individuals, there must be "circumstances" which make the describer's idea appear on the interpreter's end. How this is done will be elaborated in chapter 4.

3.4 The Concept Of A Possibility Distribution

Zadeh (1978a) defines the concept of a possibility distribution as a "fuzzy restriction which acts as an elastic constraint on the values that may be assigned to a variable." He elaborates, "if F is a fuzzy subset of a universe of discourse U = {u} which is characterized by its membership function μ_F, then a proposition of the form "X is F," where X is a variable taking values in U, induces a possibility distribution Π_X which equates the possibility of X taking the value u to $\mu_F(u)$ -- the compatibility of u with F. In this way, X becomes a fuzzy variable which is associated with the possibility distribution Π_X in much the same way as a random variable is associated with a probability distribution."

Russell (1923, p.87) states: "The fact is that all words are attributable without doubt over a certain area, but become questionable within a penumbra, outside which they are certainly not attributable. Someone might seek to obtain precision in the use of words by saying that no word is to be applied in the penumbra, but fortunately the penumbra itself is not accurately definable, and all vaguenesses which apply to the primary use of words apply also when we try to fix a limit to their indubitable applicability."

Recognizing that the penumbra is not accurately definable, we must be aware that a possibility distribution as defined above, can only be a first approximation to modeling fuzzy descriptors [Watanabe (1978)]. A more adequate model would be ultra-fuzzy sets [Zadeh (1980b)]. However, since we admit interindividual differences in the definition of possibility distribution functions, we treat the definition of the penumbra as fuzzy, in effect.

Having this method of representing imprecise information, we can enrich descriptions with imprecise knowledge, since "a vague belief has a much better chance of being true than a precise one, because there are more possible facts that would verify it", as Russell (1923, p.91) points out. This does not imply, of course, that we should fuzzify precise knowledge, since precise knowledge may have more power in restricting possibilities than imprecise knowledge. The informativeness of a given descriptor depends on the related features in the context.

3.4.1 Possibility Distributions Vs. Fuzzy Sets -

To clarify the concept of a possibility distribution it may be helpful to relate it to the concept of a fuzzy set. A fuzzy set is a <u>collection</u> of pairs of set elements and associated fuzzy set membership values, whereas a possibility distribution represents a fuzzy <u>restriction</u> on the elements (or values) that can be assigned to a variable.

In other words, a fuzzy set represents a conjunction of elements while a possibility distribution represents a disjunction.

To distinguish the two concepts, Zadeh (1978b) uses the notation

$$X = F$$

to indicate that the variable X is assigned the fuzzy set F, and

$$\Pi_X = F$$

to indicate that the possibilities of the values that the variable X may assume are restricted by the fuzzy set F. Zadeh proposes to write "γ-F" to indicate conjunction and to write "δ-F" to indicate disjunction. For example, "γ-warm" would indicate the fuzzy set of temperatures that can be considered warm, while "δ-warm" would indicate a particular temperature whose value is constrained by the restriction that it must be from the fuzzy set "γ-warm". Since we are mostly concerned with possibility distributions, in this dissertation, we will interpret "warm" by default as disjunctive value. This has the advantage that we can treat crisp values and fuzzy values in a uniform manner. For example, we will write

temperature = 30'C

and

temperature = warm

to indicate that something has a (particular) temperature value which is compatible with the set of temperatures denoted by the descriptor on the right hand side of the equation. Actually, as we have discussed in chapter 2, we do not want to interpret "30'C" as absolutely crisp, but merely as more precise than "warm".

In the present study, we use a restricted set of possibility distributions. Values that are fully compatible with a given linguistic label will constitute the "core" of a possibility distribution and have a possibility value of 1; values that are fully incompatible with the label have a possibility value of zero; in between, we will require a monotonic transition from 0 to 1 which constitutes the "penumbra". The fundamental distributions are assumed to be unimodal, and can be combined to form more complex distributions, in principle. Here, we will deal only with fundamental distributions and their complements, however. Details of possibility distribution representation will be discussed in chapters 4 and 5.

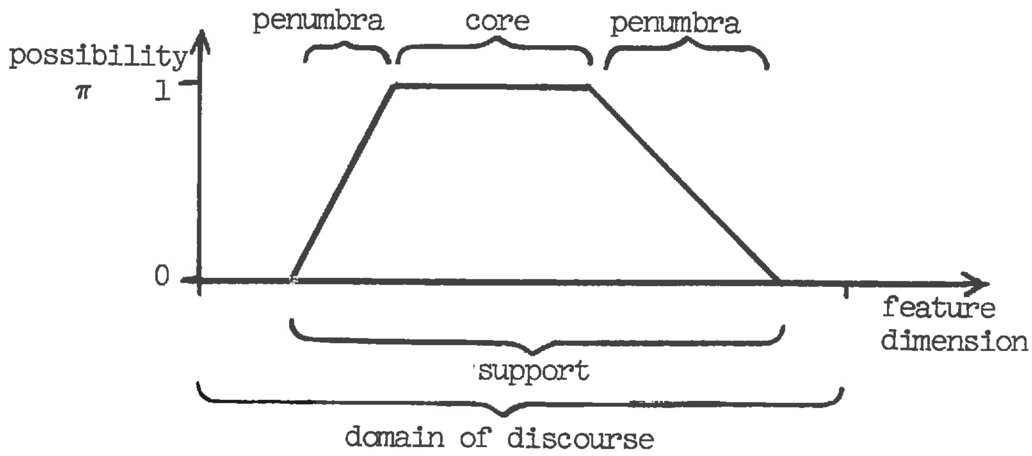


Fig. 3.2 fundamental possibility distribution

3.4.2 Acquisition And Refinement Of Knowledge -

We can use the concept of a possibility distribution to distinguish between two types of learning. "Incremental learning" can be modeled by the process of defining concepts in terms of possibility distributions and by the process of gradually adapting possibility distributions according to constraints which the "student" successively experiences. This process is being studied by Lopez de Mantaras (1980a).

The other type of learning can be viewed as having "gestalt" character: it corresponds to the process of precisiation (or de-fuzzification) of knowledge. No knowledge means: everything is possible and additional knowledge suggests that some values are more easily possible than others. Thus, the denotation of descriptors refined in this way may successively become clearer.

3.5 How Are Descriptors Defined?

A descriptor may denote several different levels of meaning. These levels correspond to a "depth of understanding" of the underlying concept. Not all of the levels must be defined in order for the descriptor to be used in a meaningful manner. It is conceivable that a descriptor adopts increasingly deeper levels of meaning through use.

For the purpose of object descriptions we will limit the discussion to five semantic levels. These are sufficient for object denotation, but they cannot capture deeper forms of understanding of the objects they denote.* The levels correspond roughly to the languages L1 - L5, although they have not been motivated by complexity considerations.

3.5.1 Descriptors As Tags -

The most fundamental and primitive use of a descriptor is to use it merely as a name tag for an object, without giving it any interpretation. In this case, the descriptor

----

* "Understanding" sometimes is assumed to be an all-or-nothing event. The viewpoint taken in this thesis is that understanding must be determined relative to the goal for which the understanding process is undertaken. A statement can be called "completely understood" when all intended semantic levels have been grasped by the interpreter according to the intentions of the statement issuer.

stands for itself and the only meaningful question an interpreter can ask is whether or not this tag is identical to a given object label.

3.5.2 Ostensively Defined Descriptors -

A descriptor obtains some meaning if we specify a feature dimension which it elucidates and a way to measure values along this dimension. A feature dimension can be specified in terms of a procedure which yields one-dimensional feature values or by enumeration of partially ordered objects. A particular descriptor can be defined ostensively by associating it with the instance of a feature to which the descriptor applies [Russell (1948)]. At this point, the only meaningful inquiry about the descriptor is an identity test. However, by defining a descriptor in terms of a set of values and defining a set of descriptors which characterize the same feature dimension, we open the way for relating various descriptors and feature values to one another. This corresponds to a deeper insight into the meaning of a descriptor.

3.5.3 Descriptors Defined In Terms Of Relations -

If we have a way of measuring feature values along a given dimension, we can define a descriptor in terms of this feature dimension plus an indication of a feature value relative to other descriptors which are defined on the same feature domain. This allows us to order feature descriptors partially and to compare them with one another. The ability of comparing meaning is a prerequisite for a calculus for descriptors.

3.5.4 Descriptors Defined By Instances And Relations -

If both, absolute instances of and relations between a set of descriptors are known, a descriptor becomes more meaningful yet: not only can a feature be compared with a descriptor in relative terms, but it also can be tested in absolute terms.

3.5.5 Descriptors Defined By Possibility Distributions -

Finally, if a descriptor can be defined in terms of possibility distributions over the entire domain of discourse, both absolute and relative information is available for any feature value. This is the deepest level of direct referential information about an object that we can expect.

3.5.6 Example -

As illustration for gradual acquisition of several levels of meaning consider this model of a person who learns a new adjective in a foreign language by experience:

1. Our subject hears the adjective being used in an assertion in connection with a noun. He does not know what the adjective describes but he would be able to repeat the assertion.

2. From the context, our subject is able to infer which quality is characterized by the adjective and he notices the feature value of the corresponding object but he cannot infer the scope of this value. Now he would be able to use the adjective in an identical situation but he could not apply it to similar situations.

3. The adjective is contrasted to other adjectives describing different values of the same quality. Now he learns a multitude of situations in which the value may be applied but he would be unable to exceed the range of the examples that were presented to him.

4. The subject learns a partial definition of the adjective which allows him to use it in situations which are beyond the scope of the examples which he has learned.

5. The subject learns a complete definition and gains full control over the use of the word.

3.6 Linguistic Labels

We describe objects in terms of linguistic labels and linguistic operators. "Linguistic" is contrasted here to "numerical". Linguistic symbols are not axiomatically defined as numerical symbols (numbers) are. For this reason, we can not develop a calculus for linguistic symbols in the same way as for numerical symbols. We must manipulate linguistic symbols through approximations as long as we do not have learning systems which are able to grasp the meaning of linguistic labels by active experience.

In order to use linguistic symbols in such a way that we can relate them to one another in a systematic fashion we must give them relatable interpretations or we must associate them with an axiomatic system for which rules already have been developed.

Both options have been provided for in the PINPOINT model: 1) two linguistic descriptors can be related to one another by indicating their relative position along the feature axis, and 2) a descriptor can be given an absolute interpretation in terms of a possibility distribution.

3.7 Linguistic Operators

The reference set of a linguistic label can be modified by linguistic operators [Bolinger (1972)]. Cliff (1959) proposed to represent each adjective and each adverb by a number. If an adjective is modified by an adverb, their respective numbers are multiplied to obtain a new value on the feature dimension of the adjective. For example, if "bad" were represented by a negative number, "pleasant" by a positive number, and "very" by a number greater than unity, "very pleasant" would obtain a higher value on the "bad - pleasant" dimension than both "bad" and "pleasant".

Zadeh (1972, 1975a) and Lakoff (1973) represent adjectives by fuzzy sets and adverbs serve to modify fuzzy sets. Fuzzy set modifiers may precisiate or fuzzify their associated descriptor, they may shift the emphasis of the descriptor, or they may cause a combination of these effects. Some controversy arose about which modification should be associated with a given adverb in English. The standard paradigm for this question is the linguistic modifier "very". Does "very" shift a fuzzy set [Zimmermann (1979)] or does it modify the shape of the membership function, as Zadeh (1972) suggested?

To find an answer to this question, we will look in detail at a particular example. Consider the statements "the water is warm" (1)

and

"the water is very warm" (2)

What is the difference between (1) and (2) ? Assuming, the statements refer to the same context, say to water in a lake and the description's purpose is to help the interpreter determine whether he or she will gain pleasure from taking a swim.

CASE 1: Suppose, "warm" restricts the possible water temperatures as indicated in this possibility distribution:

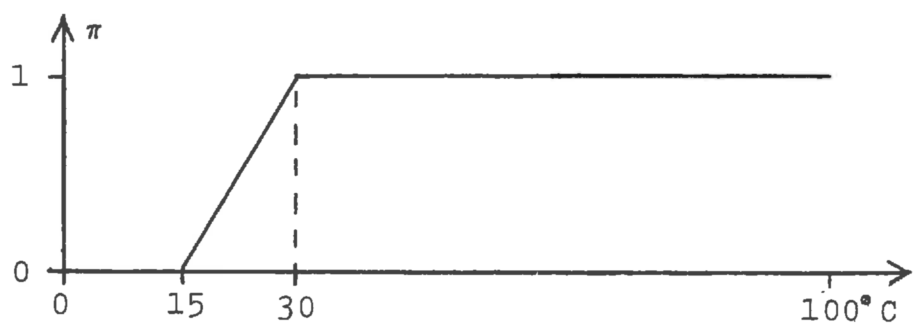


Fig. 3.3 possibility distribution defining "warm"

This distribution signifies that, in the given context, water that is described as "warm" cannot have a temperature below 15'C, without difficulty it can have a temperature above 30'C, and between 15'C and 30'C for any temperature $t_2 > t_1$ the reference of "warm" to t_2 would be more easily possible than to t_1. The absolute possibility value is not of significance here; only relative possibility values are considered.

The label "very warm" denotes higher temperatures than the label "warm". Or, in terms of possibilities, for a low temperature, there is a lower possibility to assign the label "very warm" than to assign the label "warm". The resulting possibility distribution for "very warm" would be a subset of the distribution for "warm" as shown in the figure

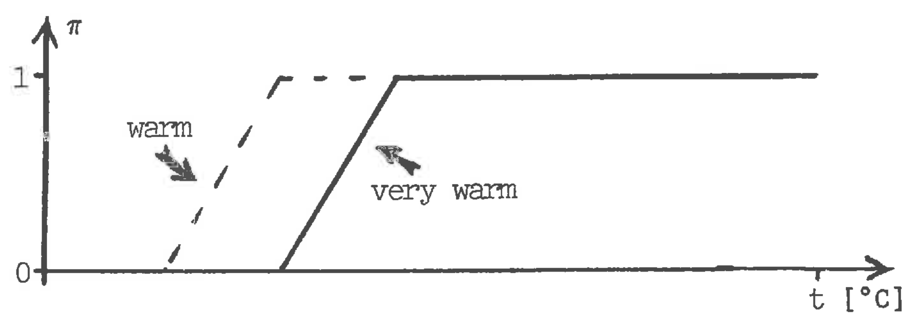


Fig. 3.4 possibility distribution defining a subset "very warm"

This means that all temperatures that can be labeled "very warm" can be labeled "warm" as easily, but the label "very warm" is less fuzzy since it indicates more precisely which temperature the water actually has.

CASE 2: There are situations in which "very warm" temperatures are not considered a subset of "warm" temperatures, as in the question "is the water cold, warm, or very warm ?" Possibility distributions corresponding to these labels are shown in this figure:

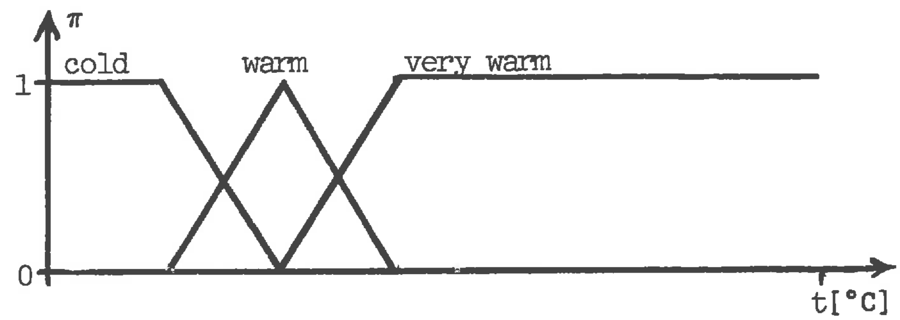


Fig. 3.5 "cold", "warm", and "very warm" as contrasting labels

Here, the modifier "very" does not have the effect of subset selection; it acts as a shift operator on fuzzy sets and does not make the description more precise.

Does this mean that the same linguistic modifier can have very different effects in very similar contexts? It does not seem so. The label "warm" in the second case can be viewed as an abbreviation for the label "warm but not very warm", with "warm" and "very warm" having the meanings of the first case as indicated in this figure:

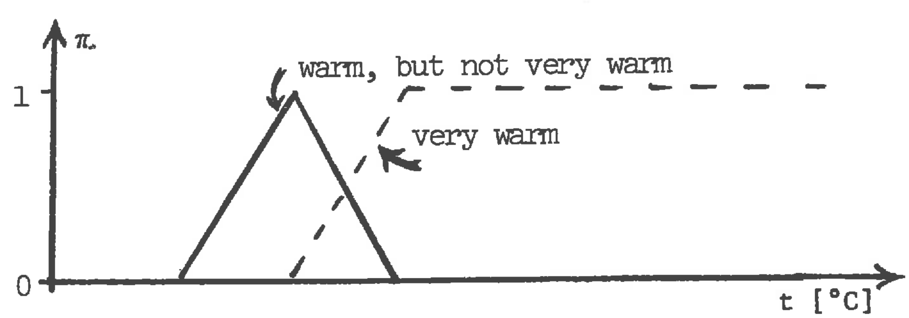


Fig. 3.6 "warm but not very warm" in case 1

Similarly, modifiers of linguistic labels, the support of whose possibility distributions is not located at the beginning or the end of their domain interval, can have multiple interpretations. For example, we can fuzzify or precisiate the description "standard-sized pencil" with the descriptors "approximately standard-sized" and "exactly standard-sized", respectively. The interpretation of these descriptors according to case 1 above is depicted here:

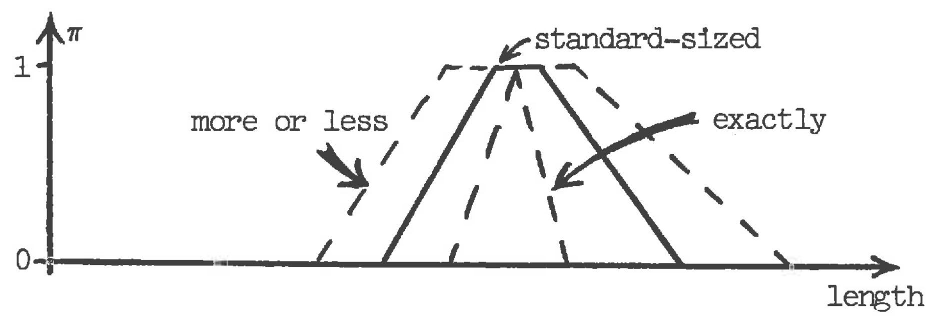


Fig. 3.7 fuzzification and precisiation

3.8 Hybrid Representation Of Linguistic Descriptors

We provide three ways of representing linguistic descriptors to reflect the different meaning levels outlined in section 3.5:

1. representation by label

2. representation by relations

3. representation by possibility distribution

This hybrid representation scheme corresponds to the different ways a descriptor may be defined. Not all three representations must be present simultaneously, but they can coexist.

Representation of descriptors merely by their labels is trivial, but sufficient for trivial matching tasks. Relational representation is suitable to express an ordering of linguistic labels along a feature dimension; to express subset and superset relations between linguistic labels to account for precisiation and fuzzification; to express right shift and left shift of linguistic values to account for emphasizing and deemphasizing the associated concept. Representation of a possibility distribution, finally, corresponds to detailed "bottom-up" concept definition.

The next chapter discusses how these representations are used for description interpretation.

CHAPTER 4

INTERPRETATION OF OBJECT DESCRIPTIONS

This chapter discusses how PRUF descriptions are compared against a data base on objects in order to yield meaningful responses to object identification requests. We will discuss under what circumstances two object descriptions can be considered consistent with one another. We distinguish between object descriptors, i.e., descriptors which refer to features of individual objects, and set descriptors, i.e., descriptors which characterize sets containing those objects.

The important aspects which make the interpretation task non-trivial are imprecision and fuzziness of the descriptions. Imprecision and fuzziness are relative properties: we call a descriptor *imprecise* if we can resolve the corresponding object feature to a higher degree than the descriptor itself. A descriptor is considered *precise* if feature resolution for the two representations agree. A descriptor can be called *overprecise* if the object features cannot be verified to the same degree of precision

[Popper (1976)]. Correspondingly, a descriptor can be more, equally, or less fuzzy than a given reference feature [e.g. Kochen (1979)].

4.1 Compatibility = Adequacy + Agreement

To determine compatibility between object descriptors and actual object features, we consider two aspects: is the descriptor adequate in the given context and if so, to what extent do the features in question agree? The adequacy of a descriptor influences the weight that a particular descriptor should be attributed and possibly the effort that should be taken to determine the feature agreement.

Consider, for example, an object world containing several sticks of differing length. It appears adequate to describe one of these sticks as "the longest". In addition, this feature will agree with one of the sticks, so we can say the description is compatible with one of the objects.

Now suppose, we look for a "long" stick, but all the sticks around are short. The description still appears adequate, since length is a feature which applies well to sticks. However the feature value "long" does not agree with any of the candidate objects. The description is not compatible with any of the objects.

Let us come back to the first example, but let the sticks have approximately the same length. Now the description of a particular stick in terms of length ("the longest") has become much less adequate, even though "length" is a feature which applies to sticks. Again, the description is not well compatible with any object. I am suggesting that a descriptor of this kind in such a context should be given less significance (weight) even though one of the sticks <u>is</u> "the longest". Similarly, if we look for "the long" stick and there are several sticks which qualify as "long".

Of course, we can have descriptors which are neither adequate nor agree with respect to their feature values, for example, if there are only short sticks and we look for "the long" stick.

Adequacy and agreement are related to sense and denotation of descriptions, respectively. Frege (1892) discusses their significance in detail. As Frege, we will not consider connotations that might be associated with particular words in a description. For example, we will not distinguish between name and nickname for a given object, but we will use them interchangeably as labels to reference the object. We will use "adequacy" of a description in a given situation to determine a measure of confidence of the interpreter in its interpretation.

4.2 Object Descriptors

The object descriptions we deal with are human created. They can be represented either in PRUF notation, elucidating the significant aspects of the description, or they can be in English if the intended meaning can be extracted automatically and be converted into PRUF notation. Lopez de Mantaras (1980b) developed an ATN-based parser capable of translating simple object descriptions from English into PRUF representation.

In interpreting object descriptors, we first look for adequate feature references, then we compare feature values. As result of the comparison we obtain two measures describing the quality of the match: a measure of agreement between description and object and a confidence measure qualifying this interpretation. The confidence measure indicates to what extent the overall description is adequate to describe the object in question.

4.2.1 Candidate Set Selection -

A PRUF descriptor has the form

[<property> = <value>]

<property> serves as key for comparing corresponding object features in the description and the data base. We can distinguish three cases:

1. the data base contains a feature slot identical to <property> or can be transformed into a slot which is identical. Then a comparison of the corresponding feature values will be adequate.

2. the data base contains a feature slot which is related to <property>, but a precise correspondence cannot be established (e.g. <property> = tall, <feature slot> = big). A comparison of the corresponding values becomes less adequate (but not meaningless). The confidence in the result of the comparison is reduced.

3. no related feature slot can be found in the data base. It is unknown whether or not the feature in question is meaningful for the object, i.e., the descriptor is not adequate, in the given situation. There is no confidence in information from this descriptor.

Feature values referring to <property> of case 1 can be compared directly. To compare feature values of case 2, we generate a new slot in the data base which describes the requested feature and the confidence in this description. Then the feature values can be matched. If we admit incomplete data bases, case 3 objects cannot be excluded from the candidate set, but feature values in question cannot be compared. This results in a lowering of confidence that those objects are in fact good candidates.

4.2.2 Feature Matching -

The candidate set selection process provides us with a set of objects which contains the target object. Now the feature values have to be screened for compatible candidates. In the following discussion, we will analyze the possibilistic content of descriptors with respect to reference objects. Initially, we will assume that the descriptors are represented by possibility distributions; subsequently we will show how this prerequisite can be relaxed. We will define some labels to refer to possibility distributions:

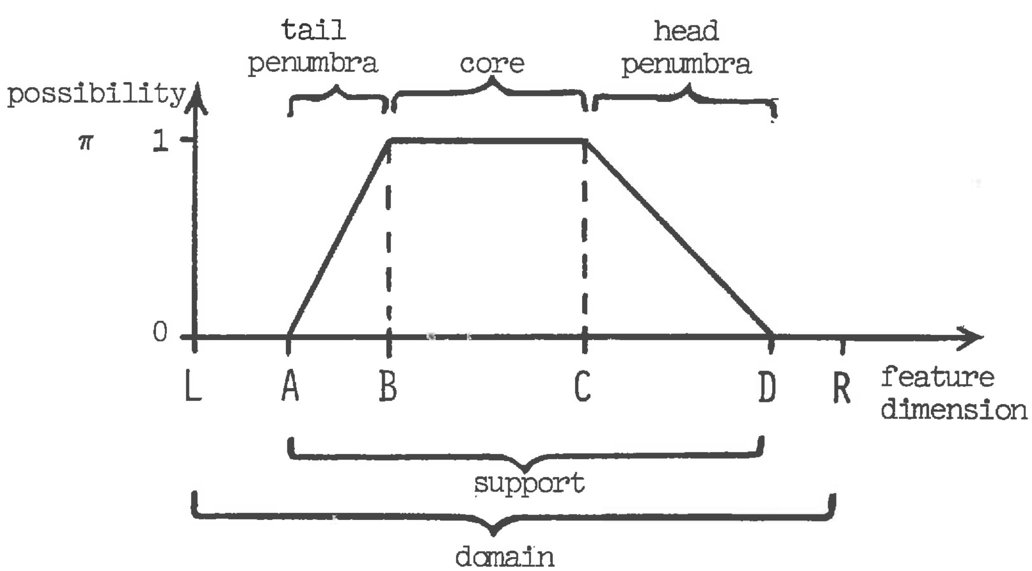


Fig. 4.1 terminology for possibility distributions

The domain of a possibility distribution is an interval or an ordered set of discrete points. In our discussion, we represent both cases by continuous intervals with the understanding that in the discrete case, only the domain points are defined. We denote the left and right extreme points with "L" and "R", respectively.

The interval on which the possibility distribution assumes non-zero values, is called the <u>support</u> of the distribution and is referred to by "[A,D]".

The interval on which the distribution assumes unity, is called <u>core</u> and labeled "[B,C]".

[A,B] and [C,D] constitute the <u>penumbra</u> of the distribution. If one of the penumbral intervals coincides with the starting point or with the end point of the domain (L or R), it is called <u>head</u>, the other penumbral interval is called <u>tail</u>. If none coincides, the right hand interval will be referred to as "head", the left hand interval will be referred to as "tail". Descriptors associated with the object description will be abbreviated <u>ROD</u> ("request object distribution"), those associated with the target object, will be abbreviated <u>TOD</u> ("target object distribution"). In the description of relationships between TOD and ROD, we will investigate equality relationships between intervals (denoted by "="), subinterval relationships (denoted by "c"), and overlap relationships, (denoted by the percent symbol "%").

We will now consider the situation in which an adequate feature slot has been found in the candidate selection process and investigate how we can determine agreement between descriptor and target object. We will start at a point of no knowledge and stepwise try to find some evidence to refine the judgement.

4.2.2.1 Initial Probing Of Matching Hypothesis -

The core of a possibility distribution indicates the range of feature values that can be achieved without stretching the associated concept. If the core of the TOD is a subset of the core of the ROD, we have positive evidence that the descriptor may refer to the target feature.

An example for this situation is a query containing the descriptor for some water temperature

ROD = lukewarm water

if the target object is described by

TOD = water of 30'C.

If we assume that the cores of these descriptors relate to one another as shown in Figure 4.2, the matching hypothesis is supported.

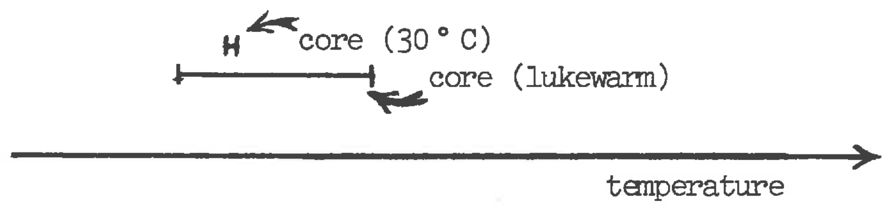


Fig. 4.2 core (TOD) contained in core (ROD)

On the other hand, if the cores of the two descriptors do not overlap at all, we have a first indication that the object descriptor may not refer to the target object. An example for this situation is:

ROD = warm water

TOD = cold water

if their respective cores are related like this:

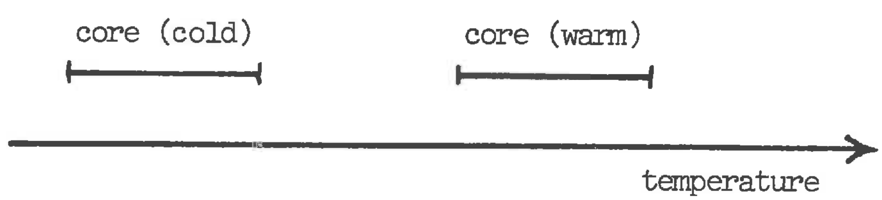


Fig. 4.3 core (TOD) and core (ROD) do not overlap

If the cores of TOD and ROD overlap partially, the match is inconclusive: parts of the descriptors agree fully, other parts do not. These parts all describe the same feature; thus, partial incompatibility between the descriptors indicates that the descriptors per se denote a different range of feature values. It does not indicate that they may not have an identical reference. An example for this situation is

ROD = very warm water

TOD = hot water

with the following cores:

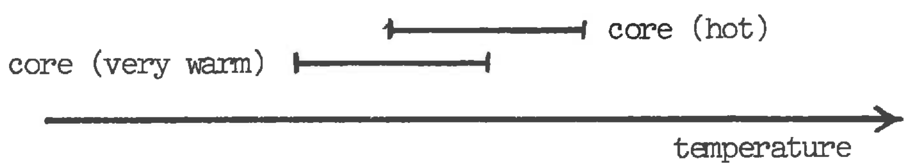


Fig. 4.4 core (TOD) and core (ROD) overlap partially

In summary, we can make a preliminary decision on the matching hypothesis by looking at the cores of their distributions only and distinguish three conclusions, "yes", "no", "perhaps".

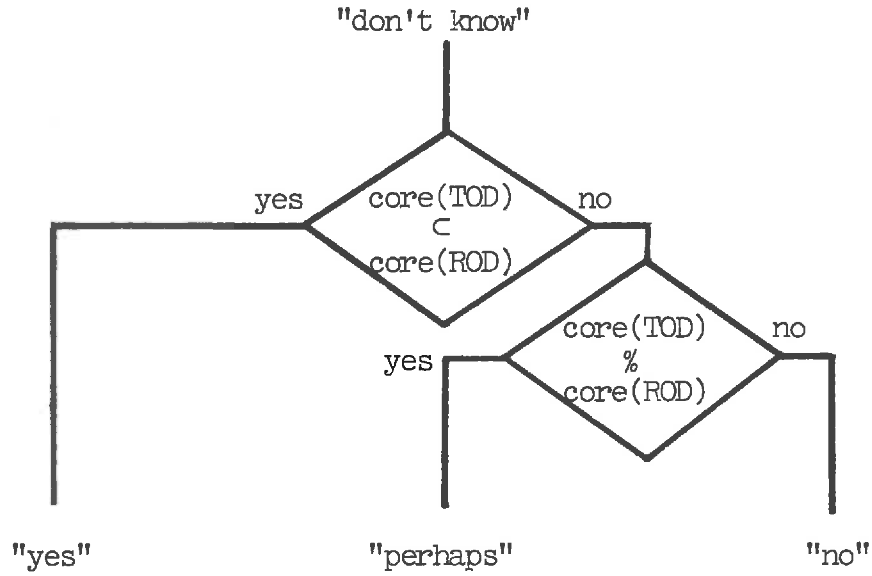


Fig. 4.5 decision tree for initial hypothesis

4.2.2.2 Qualification Of Hypothesis Support -

We consider the case in which the initial probing of the hypothesis resulted in a "yes" answer. We compare the heads of ROD and TOD. The heads are those parts of the possibility distributions which may coincide with the beginning or the end of the feature domain and thus may correspond to emphasized feature values. Thus, if the heads of the two distributions agree, we have further evidence that the same feature value may be referred to by both descriptors, and we can emphasize the response to the matching request by "yes, indeed". For example,

ROD = warm water

TOD = very warm water

have coinciding heads, if they are represented by the following distributions:

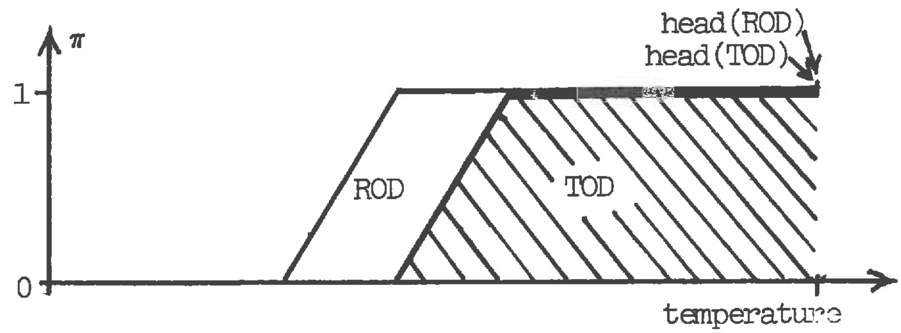


Fig. 4.6 head (TOD) agrees with head (ROD)

In this particular example, the head of TOD and the head of ROD coincide with the end point of the feature domain.

If, on the other hand, the heads do not agree, we may de-emphasize the matching result by "yes, but", indicating that it still may be possible that both descriptors refer to the same feature, but that they emphasize different prototype values. An example for this situation is:

ROD = hot water

TOD = water of 84'C

and is depicted by these distributions:

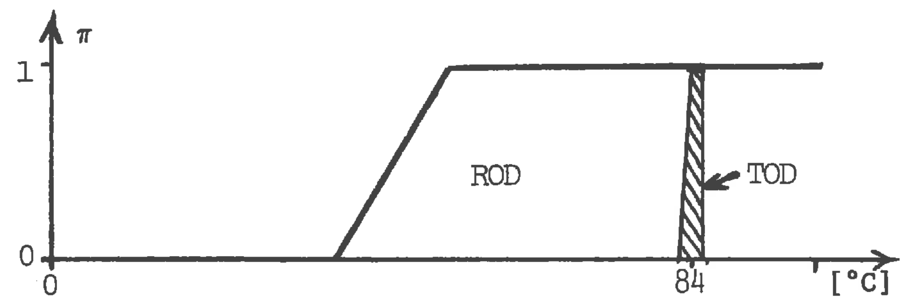


Fig. 4.7 head (TOD) does not agree with head (ROD)

4.2.2.3 Qualification Of Hypothesis Rejection -

Now, we consider the case in which the initial probing of the hypothesis resulted in a "no" answer. We compare the support intervals of ROD and TOD. The support intervals designate the portions of the possibility distributions to which a corresponding reference value possibly may refer. Thus, if the support intervals of ROD and TOD do not overlap, we can confirm the rejection of the hypothesis by "no, indeed not". An example for this situation would be

ROD = hot water

TOD = cold water

with the corresponding support relationship:

Fig. 4.8 support (TOD) and support (ROD) do not overlap

If, on the other hand, the supports of ROD and TOD do overlap, then there is a remote possibility that the corresponding descriptors refer to the same feature value. We can qualify the "no" answer by "no, but". At least one of the associated concepts has to be stretched to achieve this agreement. For example,

ROD = hand warm water

TOD = very warm water

if the corresponding possibility distributions look like this:

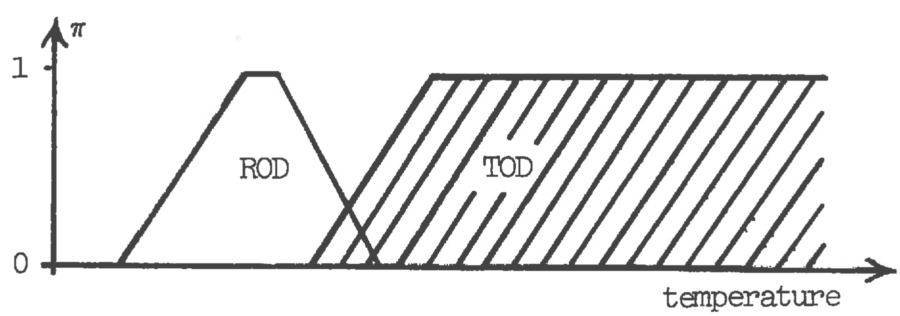


Fig. 4.9 support (TOD) and support (ROD) do overlap

4.2.2.4 Conclusion For Emphasized Hypothesis Support -

We return to the result of section 4.2.2.2 in which the matching hypothesis was emphasized by "yes, indeed". In that case, the heads of the distributions were found to be equal. At last we can check whether the tails do agree as well. If they do, the two distributions agree completely. The matching hypothesis is supported "absolutely". For example,

ROD = warm water

TOD = warm water

If, on the other hand, only the heads agree, but the tails do not, the response will be "indeed".

4.2.2.5 Conclusion For De-emphasized Hypothesis Support -

The "yes, but" case of section 4.2.2.2 can be further qualified, as well. If the support of TOD is completely included in the support of ROD, then the entire distribution of TOD is included in ROD (since the core of TOD was included in the core of ROD). Thus, all objects referenced by TOD are also referenced by ROD. The question regarding the agreement of the two descriptors is answered by "yes".

If, on the other hand, the support of TOD is not completely included in the support of ROD, there is a marginal possibility that TOD refers to an object which is not covered by ROD. For this reason, the affirmative response will be weakened to "quite possibly".

4.2.2.6 Conclusion For De-emphasized Hypothesis Rejection -

We now revisit the result of section 4.2.2.3 in which the matching hypothesis was de-emphasized by "no, but". In this case the support intervals of the two distributions were found to overlap. At last we will check whether the core of TOD overlaps the support of ROD. If so, the two descriptors come rather close in their reference values, the matching result will be "not quite".

If the core of TOD and the support of ROD do not overlap, the response will be "no".

4.2.2.7 Conclusion For Emphasized Hypothesis Rejection -

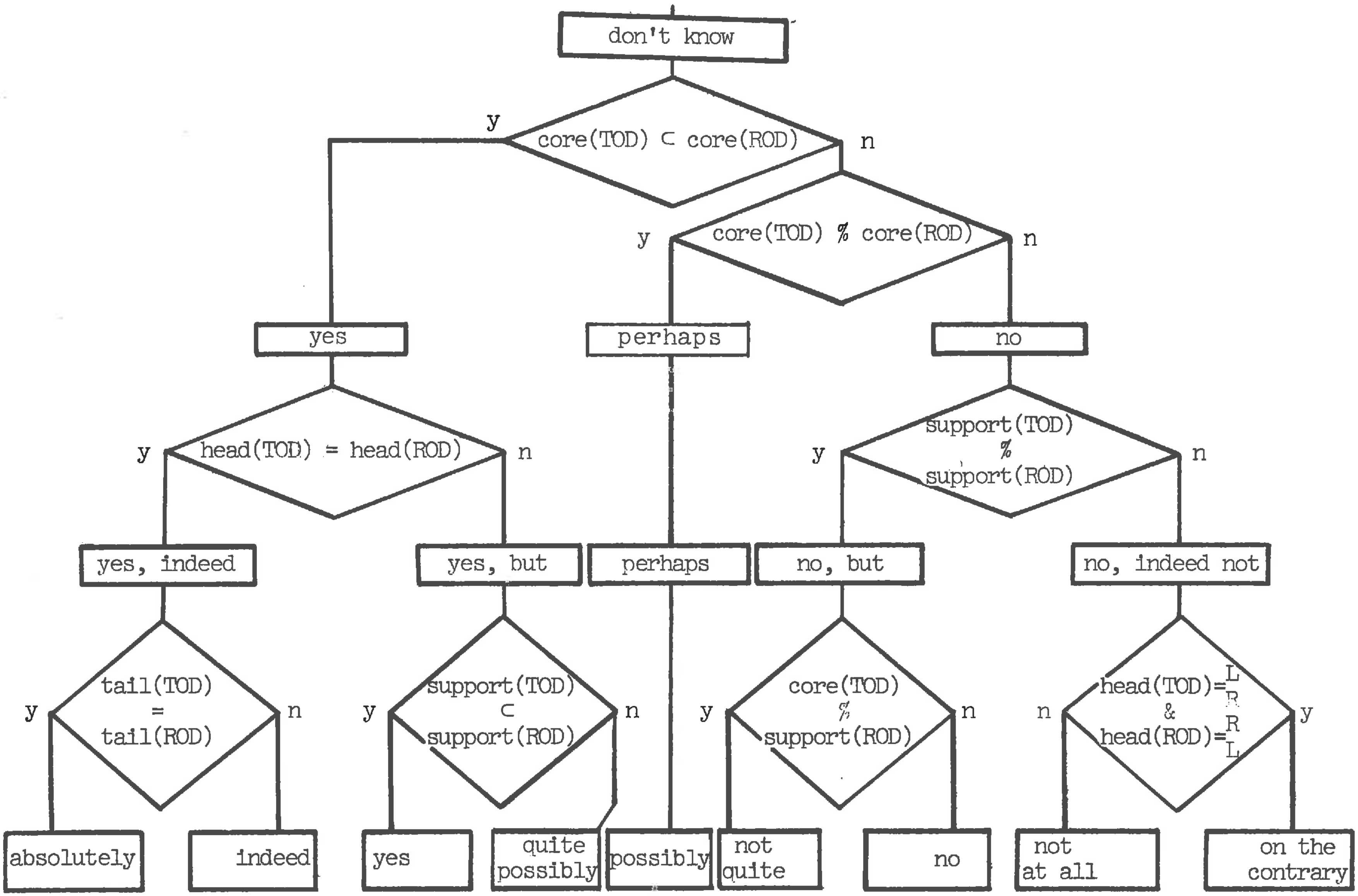
In the case in which the support intervals of the two distributions do not overlap ("no, indeed not") we also can distinguish two situations. We will check whether the heads of TOD and ROD are located at the ends of the domain interval. If so, they must be located at opposite ends of the interval and the response "on the contrary" will be given. Otherwise, the response will be "not at all".

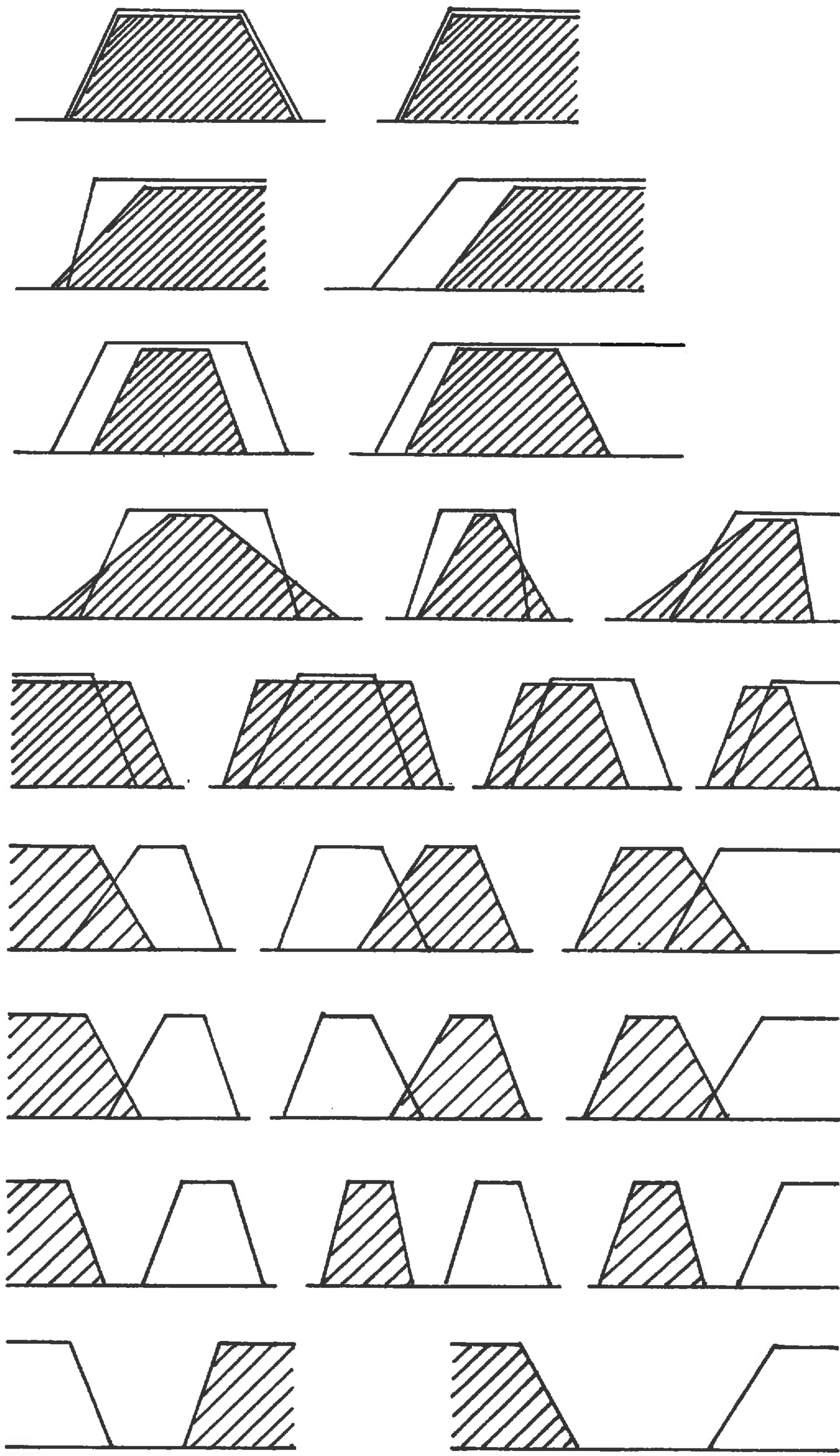
4.2.2.8 Summary Of Matching Procedure -

The complete decision tree for the feature matching procedure is shown in Figure 4.10. In Figure 4.11, relationships between the corresponding possibility distributions are represented graphically.

Fig. 4.10 decision tree for qualitative matching

Fig. 4.11 possibility distribution and corresponding matching results


don't know
core(TOD) ⊂ core(ROD)
y
n
core(TOD) % core(ROD)
y
n
yes
perhaps
no
head(TOD) = head(ROD)
y
n
support(TOD) % support(ROD)
y
n
yes, indeed
yes, but
perhaps
no, but
no, indeed not
tail(TOD) = tail(ROD)
y
n
support(TOD) ⊂ support(ROD)
y
n
core(TOD) % support(ROD)
y
n
head(TOD)=L R & head(ROD)=R L
n
y
absolutely
indeed
yes
quite possibly
possibly
not quite
no
not at all
on the contrary



absolutely

indeed

yes

quite
possibly

possibly

not
quite

no

not
at all

on the
contrary

Notice that we have obtained a gradual scale of matching results even though we have a finite number of possible responses [Bellman & Zadeh (1977)]. Each linguistic answer could be replaced by its neighbor and still would quite well reflect the result of the matching process. This behavior is called "graceful degradation" [Norman & Bobrow (1975)] and is very desirable in order to prevent the matching process from being sensitive to small perturbations in the data.

4.2.3 Possibility Distribution Modifiers -

In section 2.2.5, we introduced operators which serve to modify possibility distributions. We distinguish three basic types of possibility distribution modification: shifting, sharpening, and fuzzification.

4.2.3.1 Shift Operator -

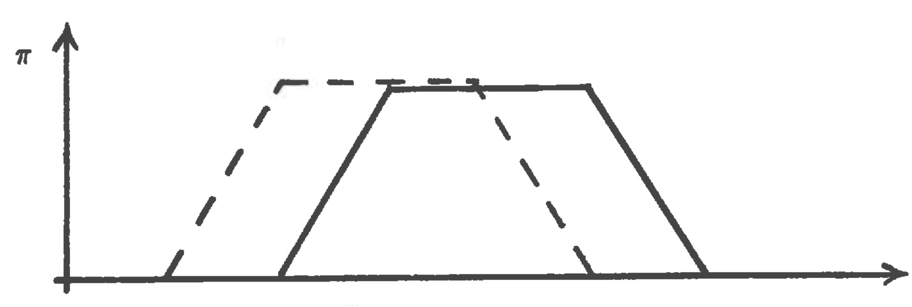


Fig. 4.12 the possibility distribution indicated by the dotted line is shifted to the right.

Shift operators serve to generate feature descriptors whose support values are higher or lower on the feature axis than those of the base distribution. These operators may correspond to linguistic modifiers like "not quite x", "almost x", or "very x". In our approach, we first define the desired effect of operators and then we associate mnemonically suitable linguistic labels.

4.2.3.2 Sharpening Operator -

Sharpening operators are used to precisiate the reference of a descriptor. This is done by selection of a subset of the base distribution.

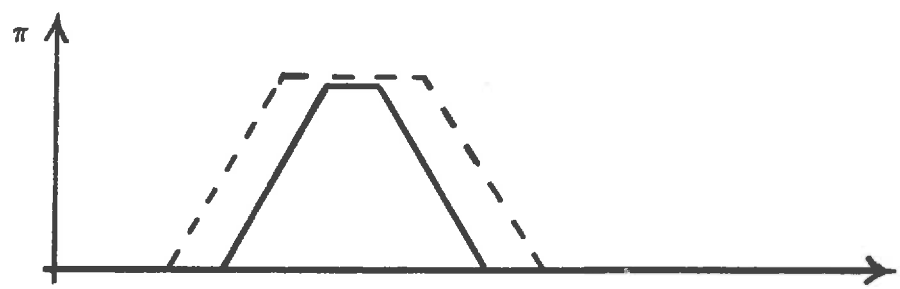


Fig. 4.13 sharpening of a possibility distribution

Corresponding linguistic modifiers may be "exactly x", "perfectly x", "precisely x".

4.2.3.3 Fuzzification Operator -

Fuzzification operators make descriptors less specific. This is done by creating a superset of the base distribution.

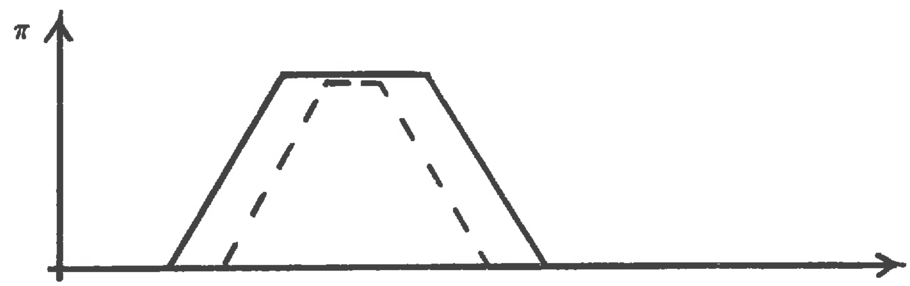


Fig 4.14 fuzzification of a possibility distribution

Corresponding linguistic modifiers may be "more or less x", "sort of x", "approximately x", "around x".

4.2.4 Relational Descriptor Matching -

The qualitative distinctions made in sections 4.2.2 and 4.2.3 suggest a simplified approach to feature matching: instead of detailed comparison of the possibility distributions associated with the descriptors, we can inspect relationships between possibility distributions directly. This is where the hybrid representation of linguistic labels comes into the picture. Useful relationships to be represented include equality, subset, superset, overlap, disjunctness, and "opposite of" relations.

These relations may be useful for several reasons:

1. they express qualitative properties which may be more easily described than a possibility distribution. For example, it is much simpler to state that "cold" refers to lower temperatures than "warm" and "warm" to lower temperatures than "hot" than it is to state that a particular temperature value is a typical or a marginal instance of "warm". Thus, there may be situations in which only relational information is available and it may suffice to answer a given request;

2. the effect of linguistic modifiers may be more easily described in terms of qualitative relations than by quantitative operators. In many cases, this information may be sufficient to answer a request in a meaningful manner. For example, suppose we view the linguistic modifier "very" in a given context as a subset selector and our data base contains the assertion

   "the water is very warm"

   and we want to answer the request

   "is the water warm?"

   then our system will respond "yes", since all water that is "very warm" can be labeled "warm", as a consequence of the subset relationship.

4.3 Set Descriptors And Quantifiers

The previous section was concerned with reducing a candidate set of objects whose features correspond to a given object description. In this section, we will discuss how the resulting set is matched against combinations of indefinite or definite set determiners and crisp or fuzzy quantifiers. Object descriptions may denote single objects or sets of objects and they may refer to specific objects or to unspecific ones. In natural language, this difference is expressed by the use of definite or indefinite articles. Examples are given below:

1. a big dog
2. the white cat
3. Johnny
4. three chinese men
5. the 3 Cs
6. several cups
7. the bunch of people

The following table classifies determiners and quantifiers into seven categories:

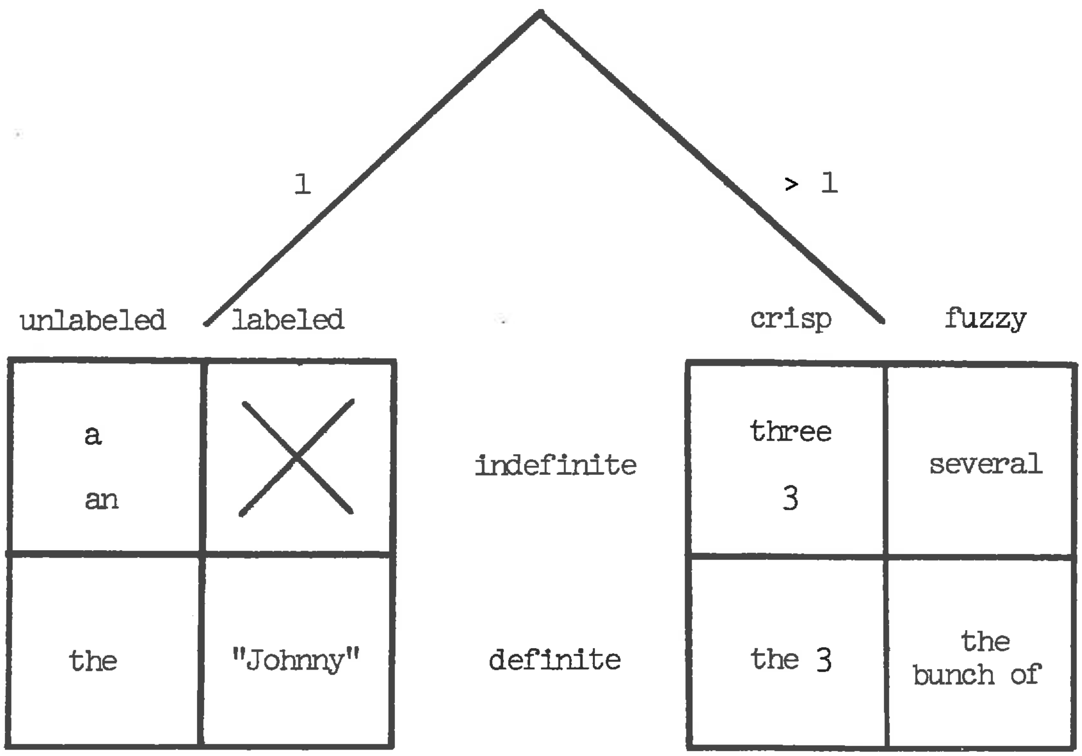


Fig. 4.15 seven different set quantifiers

In our notation for descriptions, determiners and quantifiers are specified by descriptors, namely

[determiner = definite]

or

[determiner = indefinite]

where "indefinite" is assumed by default, and

[quantity = <quantifier>]

where <quantity> can be a crisp or a fuzzy number. "1" is assumed by default.

In the following sections, we will present methods for the interpretation of simple and composite set descriptions.

4.3.1 Singular Indefinite Descriptor -

The feature matching process provided us with a set of objects which conform with the corresponding object description to a higher or lesser degree. In this and in the following sections we will assume that the objects are sorted in decreasing order of agreement with the description. The objective now is to select a subset which satisfies both the feature description and the set description as best as possible.

An example for a singular indefinite object identification request embedded into an instruction is:

"find an x"

where x is an object from the sorted list. This instruction is satisfied best with an object from the top of the list, since this object satisfies best the features of the description and satisfies completely the singular indefinite set reference. The agreement between response and description corresponds to the agreement between the object and the feature description. In case the object list is empty, the description does not only disagree with the objects, but it also is inadequate.

4.3.2 Singular Definite Set Reference -

An example for this reference category embedded into a retrieval instruction is:

"find the x"

The difference to the previous category is that the description presupposes that there is only one object which satisfies the feature description, since it contains a definite determiner. If we are to determine degree of compatibility for the complete object description and the response, we must verify that only one object conforms with the feature description. Thus, rather than looking for the first object on the list, we will look for the next best occurrence as well. If a second object does not exist or if it has low agreement with the description, compared to the first object, then the description can be considered highly adequate and we may have high confidence in the result of the description matching process. If a second object does exist, however, we compare its degree of agreement with the object descriptor with that of the first object. The greater the difference in agreement, the higher is the adequacy of the descriptor and the greater is the confidence in the matching result.

4.3.3 Singular Labeled Descriptor -

A retrieval instruction of this category is:

"find 'x'"

where 'x' is the name of a specific object. Here, not the object on the top of the list, but the object with corresponding name is requested. As in the indefinite case, the agreement with the description depends on the feature agreement and the adequacy of the description depends on the existence of an object with the correct name.

4.3.4 Crisp Plural Indefinite Descriptor -

An example for this category is

"find n x's"

where n is a natural number. We will assume, for the moment, that the request asks for "exactly n x's", although there are situations in which the interpretation "at least n x's" may be intended. This issue is discussed in the section on fuzzy cardinality. As in the singular case, the best candidates are on the top of the object list. Thus, the top n objects are selected. The agreement with the overall description is taken to be the minimum of the individual agreements. The adequacy of the description depends on the existence of n objects.

4.3.5 Crisp Plural Definite Descriptor -

An example is

"find the n x's"

Here, the first n objects are taken from the top of the list. To determine the adequacy of the descriptor, the n+1st object is compared with the nth object with respect to its feature agreement. If the difference between their agreement with the descriptor is substantial, the description is considered highly adequate.

4.3.6 Fuzzy Indefinite Quantifier -

Consider the instruction

"find several x's"

Here, for the first time the quantifier is not crisp. How many cups are "several cups"? The view that we will take is, that it depends on the particular situation context what is the best answer to this question. As in the crisp case, we will assume here that the intended meaning of "several" is "exactly several" rather than "at least several" or "at most several".

We can define what we mean by "several" in the context of cups by means of a possibility distribution. For example:

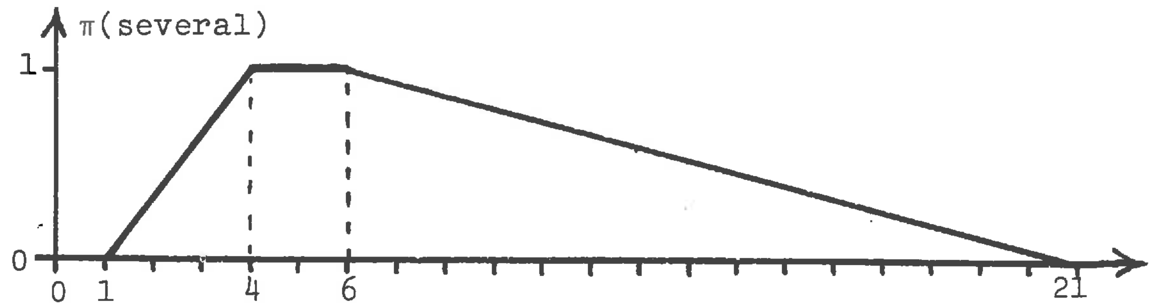


Fig. 4.16 possibility distribution for quantifier "several"

This distribution signifies that a single cup or more than 20 cups cannot be referred to as "several cups". On the other hand, 4, 5, or 6 cups would make up sets to which the label "several cups" fits perfectly.

Now suppose, in the given data base there are many objects which agree perfectly with the concept of a cup. In this case, the answer to the question above is simple: 4, 5, or 6 of these objects would satisfy the request best. But what is the best response if the data base contains less than four perfect cups plus a few objects which can be considered cups to a lower degree. Can the lack of perfect cups be compensated for by responding with a larger set of less perfect cups to satisfy the request? Hardly. Instead, we should maximize the agreement with both concepts: with the concept of cup and the quantifier "several". We do this by arranging the candidate cups in decreasing order of "cupness" and return as many cups from the top of the list as will maximize the minimum of the two individual agreements. This maximum mutual agreement value then can be taken as the adequacy value for the request.

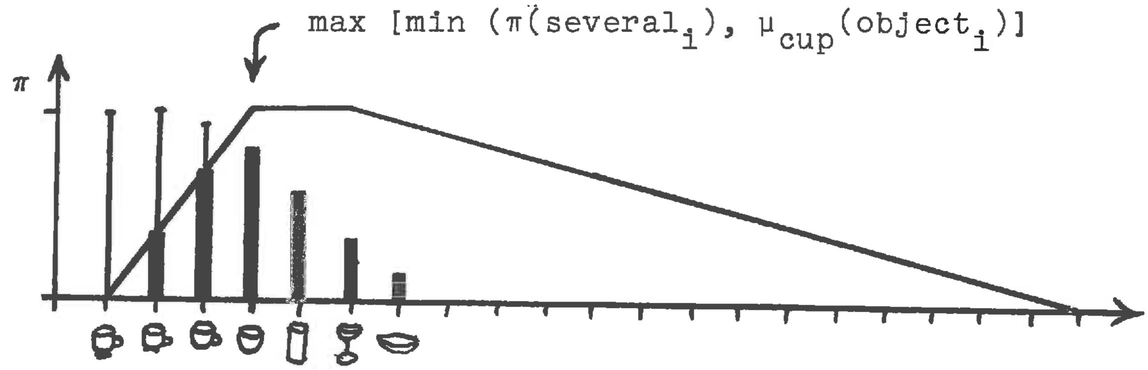


Fig. 4.17 joint possibility distributions for response optimization

4.3.7 Fuzzy Definite Quantifier -

An example retrieval instruction for this case is:

"find the bunch of spoons"

Here, the number is left fuzzy, but it is presupposed that the features distinguishing spoons from non-spoons describe a rather crisp set, in the given context. To determine the best response to the retrieval request, we want to maximize adequacy and agreement of the response. The adequacy is influenced by the difference of agreement between the objects in the response set and the agreement of the next object in the list and by the maximum agreement of the "worst" response object and the quantifier agreement.

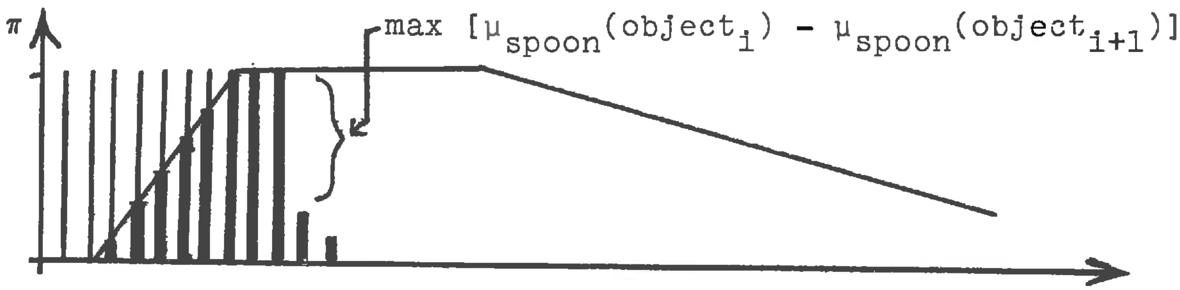


Fig. 4.18 joint possibility distribution for response optimization

4.3.8 Higher-order Quantification -

Interaction between multiple set quantifiers can be treated in much the same way as a combination between set descriptors and object descriptors. Consider, for example, the following instruction:

"pick several bunches of pretty flowers"

Here, we have interaction between three descriptors. Each bunch of flowers is supposed to be characterized by a particular appearance of its elements which may restrict the number of flowers in each bunch. So far, there is no difference to section 4.3.6. But if we ask for several bunches, we may have to take into account an additional trade-off, if the resource of pretty flowers is limited. In the absence of additional information as to which property or quantifier is the most important to satisfy, we will maximize the minimum of all agreement values for a given constellation while maximizing each individual agreement value within the above stated constraint. In case of the example above, we would have a set of constellations which satisfy the description "bunches of pretty flowers" to a higher or lesser degree. From this fuzzy set of sets we select the one which maximizes both its membership in the fuzzy set and the quantifier "several".

4.4 Note On Fuzzy Cardinality

As alluded to earlier, different interpretations of natural language descriptions from the ones discussed above, are possible. Here, we will review the concept of fuzzy cardinality and point out a variety of interpretations of fuzzy quantifiers. The search for an appropriate measure of the cardinality of a fuzzy set boils down to the question how we can best characterize the intuitive notion of the magnitude of a fuzzy set. At first glance, the sum of the individual membership values of the elements of a fuzzy set [Zadeh (1977d)] provides an appealing measure, but in many cases it yields undesired results.

For example, the cardinality of the fuzzy set of very rich people would be rather large if everybody who owns something contributes to the measure. More appropriately, cardinality of a fuzzy set is characterized by a fuzzy number, as we did in the foregoing sections. Then the meaning of a quantifier can be related directly to the domain to which it is applied.

As in the non-fuzzy case, fuzzy quantifiers in natural language may have an implied meaning deviating from the most straightforward interpretation. Specifically, a quantifier Q may be intended to mean [Zadeh (1979c)]:

1. at least Q

2. exactly Q

3. at most Q

Examples are:

1. can you lend me <u>5</u> dollars?
   don't worry, I have <u>another</u> pen

2. she has <u>two</u> jobs
   he has <u>several</u> children

3. he can do <u>fifteen</u> push-ups
   yes, I can handle <u>a few</u> more students in my class

In our work, we assume that the appropriate interpretation has been obtained in the translation process from natural language and is made explicit in the meaning representation.

4.5 Integrated Search For Best Response

In the preceding sections, we presented the different aspects of matching separately, for easier conceptualization. In practice, the steps we have presented will not be carried out entirely sequentially. In particular, candidate sublists do not always have to be determined entirely. The lists must only be long enough to satisfy the quantifiers, since all the elements are taken from the top of the list.

The result of an object description interpretation may reflect the adequacy and the agreement measure obtained during the matching process: the adequacy may be manifested in a summarized, qualitative answer and the agreement measure may be reflected in the detailed matching result. This point is discussed in more detail in the discussion of L-FUZZY.

CHAPTER 5

FROM FUZZY TO L-FUZZY

This chapter describes a computer implementation of some of the ideas presented in the previous four chapters. This implementation is based on the AI-language FUZZY [LeFaivre (1974b, 1977)] which presented a first step towards the integration of fuzzy sets into programming languages. We describe a dialect of FUZZY, called L-FUZZY which substitutes linguistic modifiers for numerical modifiers. This allows for direct representation of fuzzy sets and possibility distributions instead of representation of (crisp) elements of fuzzy sets.

This method of "linguistic representation" is an alternative to "linguistic approximation" [Zadeh (1975c), Bonissone (1979a,b)]. In linguistic approximation, numerical values or distributions of high resolution are identified with linguistic labels of low resolution. In contrast, in linguistic representation, we delay reference to high-resolution information as long as possible. By relying more heavily on higher-level relationships between linguistic labels than on low-level definitions, the computational effort can be reduced.

We are using the language FUZZY as basis for our implementation since FUZZY provides a rich system of control mechanisms that have been tested elaborately and stand the test [Wahlster (1978), Hahn et al. (1979)]. FUZZY runs under control of the LISP interpreter. Therefore all procedures and interactive debugging tools of the LISP language are available to the FUZZY user. The FUZZY source code is well-structured and transparent. It can be understood by FUZZY programmers with the help of the LISP interpreter even though no program code documentation is available.

5.1 Descriptors In FUZZY

A summary of design considerations and language features of FUZZY is given in Appendix B. In the present section, we will compare the representation of descriptors in FUZZY and L-FUZZY in some detail.

FUZZY descriptors may contain both a "value" and a numerical value modifier, called "Z-value". The Z-value typically is used to associate a fuzzy set membership value, a degree of possibility, probability, certainty, etc., with the value. In this way, a fuzzy set can be characterized by a collection of fuzzy set elements. The fuzzy set elements are represented by a pair (<value> . <Z-value>). If we view the numerical membership value of a fuzzy set element as representing a point on a 1-dimensional scale*, then the descriptor of a fuzzy set element in FUZZY is crisp.

The information we would like to represent is not well enough defined to allow this rigid representation, in general. Rather than representing fuzzy sets and possibility distributions in terms of well-defined instances, we like to directly represent linguistic descriptors which can stay by themselves, without definition of their meaning in precise terms. If more precise interpretation of a given descriptor is required, its meaning can be approximated in terms of a possibility distribution, for example. In this way, the higher-level concept, namely the linguistic label, serves as primary reference and its interpretation in terms of a more precise representation is an approximation. (In contrast, in linguistic approximation, the possibility distribution serves as primary reference and the linguistic label is viewed as approximation.)

Representations of this type are necessary if we believe that higher-level reasoning is possible without understanding of underlying lower-level mechanisms or that lower-level mechanisms can be inferred from higher-level

----

* this appears to be the most natural interpretation which is commonly used.

observations. For example, we may talk about movement of airplanes without understanding the forces that make airplanes fly. Thus, to predict a movement of the plane without this detailed knowledge, we need a high-level representation for airplane movements. In the process of understanding the underlying forces, we build models which constitute lower-level approximations to actual forces, but the high-level knowledge may still accurately describe the movements.

5.2 Linguistic Modifiers In L-FUZZY

The conversion from numerical modifiers in FUZZY to linguistic modifiers in L-FUZZY involves the introduction of some conventions. This is to maintain advantages that stem from the well-defined structure of the number system. In particular, we want to be able to compare unequal linguistic modifiers, without over-interpreting their meaning. In our view, representing linguistic modifiers by real numbers is an over-interpretation, even if the particular numerical value is not considered very significant.

In FUZZY, data base entries are ordered according to the value of their numerical modifier (Z-value). This allows for efficient information retrieval. Linguistic modifiers are not sequenced as rigidly as numbers are; however, they can be partially ordered according to their relative effect on the descriptors they modify. For example, the linguistic modifiers

absolutely, very, :* , somewhat, not,
not at all, on the contrary

are ordered according to strength of amplification of their associated base statement [compare Lenneberg (1975, pp.29f.)]. By arranging linguistic modifiers in this way, the data base can be pre-screened for an appropriate descriptor. To determine the compatibility between two descriptors with different modifiers, further analysis is necessary.

Substitution of linguistic modifiers (L-values) for numerical Z-values requires replacement of the rules governing these values. In L-FUZZY this is done as follows:

1. L-values are LISP atoms made up by lower case characters and digits. Blanks are written as hyphens.

2. L-values evaluate to themselves (like LISP numbers).

3. L-values may have qualitative properties in a property list. These relate them to other L-values. These properties include equivalence, subset, superset, emphasizing, and weakening

----

* The colon (":") is used in L-FUZZY to denote the neutral modifier called "unitor" [Zadeh (1978b)].

relationships.

4. L-values may have context-adaptive procedures associated which modify possibility distributions. In this way, absolute references of descriptors can be manipulated.

Since L-values do not correspond to points, but rather to fuzzy ranges, they are treated in the language more like ranges of Z-values than like Z-values themselves. We have not provided for ranges of L-values, in L-FUZZY. The effect can be achieved, however, by defining a new L-value whose reference extends over the range of the two original L-values.

5.3 Possibility Distributions In L-FUZZY

As in FUZZY, provisions for semantic interpretation of assertions are left up to the user, in L-FUZZY. A standard format is used to specify possibility distributions in accordance with section 4.2. A 5-tuple

(F, U, S, E, LIST)

sets up specifications for the distributions: F designates the feature dimension that the distribution refers to, U designates the units that are used on the abscissa, S indicates whether the possibility distribution has standard representation (S = 'P) or is inverted (S = 'I). In the inverted representation, the possibility values 0 and 1 are interchanged so that we obtain an "impossibility distribution", in effect [Sanchez (1978)]:

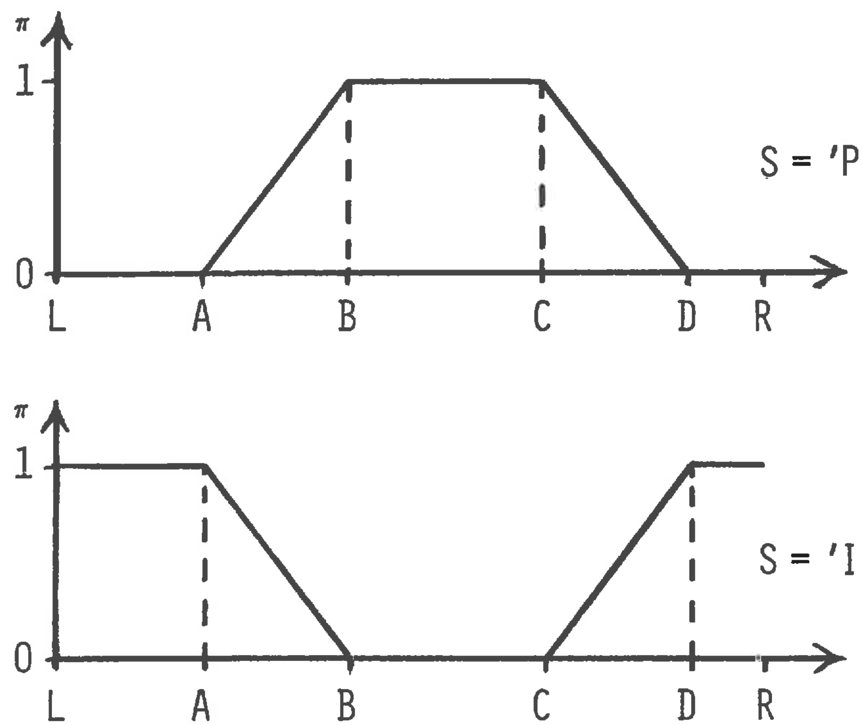


Fig. 5.1 standard and inverted representations of possibility distributions

E refers to the environment or context in which the distribution is applicable. LIST is a 6-tuple:

(L, A, B, C, D, R)

[L , R] denotes the discourse interval. In standard representation, A marks the transition point of the possibility, n, from n = 0 to n > 0, B the transition from n < 1 to n = 1, C from n = 1 to n < 1, and D from n > 0 to n = 0 when proceeding through the discourse interval from L to R. In the inverted representation, n = 0 and n = 1 are interchanged. If S = 'P and n (L) = 1 or S = 'I and n (R) = 1 then we set A:=B:=L, and if S = 'P and n (R) = 1 or S = 'I and n (R) = 0 then we set C:=D:=R.

As an example, the possibility distribution indicating warm water temperatures, which is depicted in Figure 3.3, could be represented by

(TEMPERATURE CELSIUS P WATER (0 15 30 100 100 100))

This representation allows for unimodal possibility distributions, or in the case of inverted distributions, for unimodal "impossibility distributions".

5.4 Fuzzy Matching

The multiple representation of linguistic modifiers allows for matching of assertions on various levels. Suppose, the fuzzy associative net contains assertions of the following form:

(<basic assertion> <linguistic assertion modifier>), for example:

((JOHN IS TALL) very),

((BOB IS TALL) :),

((TOM IS MEDIUM-SIZED) more-or-less).

A request to the data base has the following form:

(GOAL <basic request> <linguistic request modifier>),

for example:

(GOAL (BOB IS TALL) very).

The request modifier has the function of specifying the range of possibilities which should be searched to arrive at the answer to the request. We can distinguish three types of requests to the data base:

1. basic assertion = basic request
   assertion modifier = request modifier

2. basic assertion = basic request
   assertion modifier $\neq$ request modifier

3. basic assertion $\neq$ basic request

In the first case, the request can be satisfied by "trivial matching", i.e., only the labels must be compared, the possibility distributions involved are irrelevant.

In the second case, the relative effect of modifiers must be considered. For example, in Figure 3.4, the modifier "very" has the effect of precisiation (or selection of a fuzzy subset of possibilities) with respect to the unitor (i.e., the identity modifier). Thus, an assertion which holds when modified by "very" also holds when not modified. Conversely, an assertion which holds when not modified, possibly may hold when modified by "very". The following table shows the compatibility between four modifiers in linguistic terms:

assertion \ request	very	:	more or less	not
very	absolutely	indeed	no	not at all
:	possibly	absolutely	indeed	on the contrary
more or less	not quite	possibly	absolutely	no
not	not at all	on the contrary	no	absolutely

In the third case, where the basic assertion does not agree with the basic request, the possibility distributions must be analyzed in detail. For example, given that the data base contains the assertion

((TOM IS MEDIUM-SIZED) more-or-less),

the answer to the request

(GOAL (TOM IS TALL) :)

requires a comparison of the possibility distributions of "medium sized" and "tall" over the height of a person. In this case, the comparison yields the answer "not quite". Ten standard answers can be generated this way: "absolutely", "indeed", "yes", "quite possibly", "possibly", "not quite", "no", "not at all", "on the contrary", and "don't know". These answers are obtained according to the decision procedure for qualitative matching which is described in section 4.2.2 and depicted in Figure 4.10.

Observe that there is relatively little qualitative difference from one possible response to the next. This is an indication of "graceful degradation" of performance when program data degrades gradually.

5.5 Informative Output

In FUZZY, a retrieval request returns an assertion which best fulfills the request specification, or it fails. In L-FUZZY, the quality of the match between the request descriptor and the target descriptor is returned in linguistic terms [compare Tong & Bonissone (1979)]. For example, if request descriptor and target descriptor are related as in Figure 5.2,

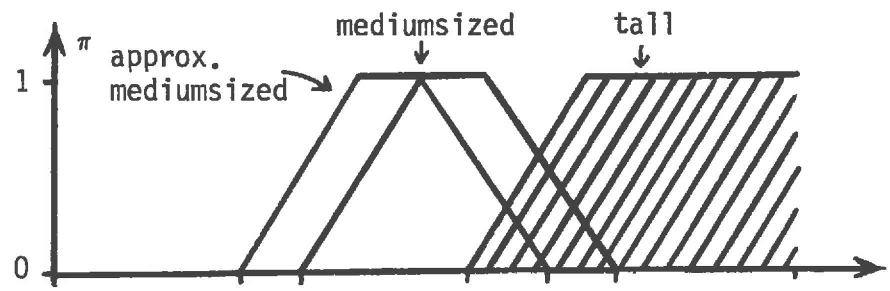


Fig. 5.2

the request

(GOAL (TOM IS TALL))

yields

(not-quite: ((TOM IS MEDIUM-SIZED) approximately))

(compare with Figure 4.11). Thus, the qualitative matching result indicates to what extent the response fulfills the request. This information can be useful to decide whether a search process can be stopped (if the request is sufficiently satisfied), whether a better response should be searched for, or whether a different goal should be pursued.

5.6 Implementation Of L-FUZZY Procedures

The integration of L-values into the control structure of FUZZY was done by systematic replacement of Z-values by L-values. The default ZHIGH and ZLOW have been replaced by LDEFAULT and LCONTRA, respectively, and have linguistic values of ":" (unitor) and "-" (antonym of unitor), respectively. ZRANGE is replaced by LRANGE and has a default value of ":", as well. The variables in the fuzzy source code which require numerical Z-values are typically prefixed by "Z". The prefix of these variables has been changed to "L" where the underlying operations had to be changed to deal with linguistic values. Also, procedures in the FUZZY source which were affected by the changes have been relabeled accordingly. The code modification was done in such a way that the logical control flow of the procedures was not affected. Instead, if additional operations are required in the linguistic reasoning process, an additional procedure level was added which roughly corresponds to the decomposition of possibility distributions into qualitative statements.

The following language primitives require L-values in place of Z-values: ACCUM:, ADD, ASSERT, BACK, EXIT, NEXT, SUCCEED, SUCCEED?, SUCCEED!, LVAL: (formerly ZVAL:). The following primitives require L-values in place of Z-ranges: DEDUCE, FETCH, FOR, DEDUCE:, FOR FETCH, FOR GOAL, FOR TRY, GOAL, TRY.

5.7 Future Developments

At this point, L-FUZZY only contains the structure for representing and manipulating linguistic information. This has been done in such a way that numerical values can be treated as a special case of linguistic values. What is missing is a system of linguistic labels, linguistic modifiers, and linguistic inference rules.

L-FUZZY has lost the conveniently structured system of numerical modifiers which FUZZY has available. For this reason, we need a tool for defining linguistic values, preferably by interaction with a computer, by means of a dialogue. As one of the next steps, we intend to implement a knowledge acquisition component in L-FUZZY [Lopez de Mantaras (1980a)].

As a next step, we can implement inference rules, both on the level of possibility distributions and on the relational level. The deduction capabilities of L-FUZZY should turn out to be useful for deriving heuristic rules from low-level possibility inference and to derive possibility distributions from high-level relational inferences.

As a step towards further development of L-FUZZY, we may generalize the concept of linguistic modifiers in L-FUZZY in such a way that they can modify various dimensions, e.g. truth values, degrees of certainty, degrees of importance, etc. It appears more natural and efficient to do this by expressing the referenced dimension implicitly in the linguistic modifier than by implementing a vector of modifiers, as LeFaivre (1974a) suggested.

CHAPTER 6

EXAMPLES AND APPLICATIONS

We have described ways of describing objects by means of linguistic descriptors and of interpreting these descriptions for object identification. In the present chapter, we will give some examples of practical applications for the described techniques. To this end, we will summarize the features of linguistic descriptions:

- linguistic descriptors are imprecise and fuzzy
- linguistic descriptors are incomplete
- linguistic descriptors are context-adaptive
- linguistic descriptors are subjective

These features suggest that this type of object descriptions may be useful in situations in which

- a relatively large set of features is required to describe all objects

- the feature values cannot all be distinguished clearly

- only a relatively small subset of features is required to uniquely designate most of the objects

In the following sections we will give some examples.

6.1 "Soft Sciences"

In the so-called "soft sciences", namely social sciences, psychology, linguistics, but also in fields like economics, meteorology, medicine, systems analysis, research results are viewed to be significant to a much lower degree of precision than they can be measured. Theories and findings in these areas are commonly communicated and taught by words -- as compared to by mathematical formulations as in the "hard sciences". In order to use these theories and research results in a scientific manner, the meaning of these words must be established, to a certain extent. Traditionally, this is done by translating somewhat crude observations into precise terms (usually points on a scale) and by giving error bounds indicating the precision of the observation. Imprecise observations then are treated as if they were precise and are used in calculations to yield precise results. The error bounds are used to indicate that the actual result may differ, but this is taken merely as a lack of perfection, not as an intrinsic necessity. Very much care has to be taken not to over-interpret the results of such calculations when it is translated back into terms that are commensurable with the original experiments.

An alternative approach is linguistic modeling. Here, somewhat crude observations are represented by somewhat crude linguistic terms, namely words, which correspond to fuzzy possibility distributions rather than points on a scale and directly reflect the precision of the observation. Reasoning processes which make use of these observations then take into account the characteristics of the distribution, rather than merely the extreme error bounds. The result of the reasoning process will correspond to a possibility distribution which may be more crisp or more fuzzy than the original distributions. It will directly reflect the sharpness of the result.

6.1.1 Example: Coding Of Facial Expressions -

A facial action coding system (FACS) has been developed [Ekman & Friesen (1978)] to describe human facial expressions. This is to relate facial expressions to various emotional conditions or different social situations. Obviously, it is not meaningful to describe facial expressions by measuring distances between well-defined points on the faces of the subjects. This is, because it is not possible to agree on well-defined points which are significant for all human faces and because it is not possible to find rules which would use these measurements to allow for comparison of facial expressions of different people in a meaningful way. This is on top of the fact that the effort in determining these measurements would not be justifiable for anything they might be used for, unless perhaps, it would be part of an automatic face description process.

Instead, facial expressions are described in linguistic terms, in this coding system. Aspects of a facial expression are described in terms of "Action Units" which have a strong correspondence to the muscular structure of the face. The facial expression then is described in terms of the contribution of these Action Units to the "distortion" of a "neutral" face:

> "An Action Unit can be totally uninvolved, or it can be: trace ... slight ... marked ... pronounced ... severe ... extreme ... maximum. The meaning of this scale will become apparent as you study the FACS illustrations in each chapter. And you will learn what is meant by trace, slight, etc., in practice scoring where you will apply this scale ..."

Scoring requirements for a particular Action Unit are expressed by rules as follows:

(1) Lateral portion of brow pulled upward <u>slightly</u>, changing the shape of brow.

and (2) Lateral portion of eye cover fold stretched <u>slightly</u>.

If you did not see the brow <u>move</u>, then the additional requirement below must be met, and at least one requirement (1), (2) or (3) must be <u>marked</u> with the other two <u>slight</u>.

(3) Horizontal or curved wrinkles above lateral portion of brow. If these wrinkles are in the neutral face, they must increase either <u>slightly</u> or <u>markedly</u>.

In a child, you might never see requirement (3). In such instances, if you did not see the brow <u>move</u>, then you must rely just upon requirements (1) or (2), but one of them must be <u>marked</u> and the other <u>slight</u>.

6.1.2 Medical Consultation -

A better known example for use of imprecise knowledge in artificial intelligence is the MYCIN system. In MYCIN, qualitative information about biological cultures is evaluated to obtain a measure of evidence for or against the existence of certain microorganisms. Degrees of evidence are expressed in numerical terms which are translated for the user into linguistic terms. Measures of belief and of disbelief can be combined by numerical rules to yield "certainty factors".

Similar as in the previous example, the linguistic labels are represented by points rather than by fuzzy ranges, in MYCIN [compare Wechsler (1976)]. A problem with the numerical inference rules in MYCIN is that they require independence of pieces of evidence. In a complex system, in particular in a system whose underlying mechanisms are not completely known, this assumption never can be guaranteed to hold. For example, if the same MYCIN rule was applied twice during the diagnosis process, the resulting diagnosis were different than if it were applied only once.

With possibilistic information we can develop inference rules which do not require independence of observations. If a rule that already was used is applied a second time, the range of possible results is not further restricted. Initially, it will be more difficult to set up inference rules and they will not be as general as in the MYCIN system. However, they may be more acceptable to the user, since they can be written in terms of examples as they appear in a text book. Different variables can be used as entries of a table which takes linguistic values. The expert can specify (in linguistic terms) by how much a given variable must change such that a given diagnosis does not apply any longer. Such a diagnosis system can be programmed in such a way that it gives very conservative advise, i.e., it rather answers "don't know" than speculate erroneously.

6.2 Person - Machine Interface

In developing expert systems, we bridge the gap between high-level knowledge and low-level consequences. The first step of formulating high-level knowledge can be by words. Relationships between fuzzily described causes and uncertain consequences can be loosely specified [compare Ragade (1976)]. A computer system as described as described in this thesis can be used to gain better knowledge about what we mean if we use certain linguistic terms. This can be done most efficiently by interactive systems which give feedback about the interpretation of the descriptions to the user and allow him or her to modify either the description or its interpretation. In such a way, an artificial language can be constructed which uses English labels and has a fairly straightforward interpretation to the human user.

6.2.1 Robot Control -

An interactively agreed-upon language can be used to command a manipulator or a robot [Munson (1971), Uragami et al. (1976)]. This can be done in much the same way as we tell a car driver who is unfamiliar with the particular surroundings how to get to a certain destination. If we sit in the car, we can use feedback and modify our instructions interactively. Accordingly, we can direct a robot or interact with a scene analysis system with linguistic commands instead of formulating mathematical procedures to achieve the task.

6.2.2 Identification Of Natural Products -

Consider the description of different kinds of wood.* It appears impossible or at least extremely impractical to classify different grains of wood in numerical terms. However, a wood expert may be able to distinguish two different kinds of wood easily by comparing grain, texture, hue, hardness, etc. in linguistic terms. These linguistic descriptors can be used for automatic classification procedures. In the interactive approach, denotation of descriptors can be taught by examples by presenting typical and marginal combinations of feature values together with an expert classification.

6.2.3 Route Finding -

Suppose, someone gives you instructions to find a certain place. Typically, these instructions are given in linguistic terms, rather than by means of precise

----

* This example was suggested by Professor L.A. Zadeh.

coordinates or measurements. Nevertheless, in many cases we are able to find exactly the place we have been looking for [Riesbeck (1980)]. We do this by combining constraints from the linguistic description. Accordingly, we can implement linguistic decision procedures which use fuzzy constraints to identify a sequence of landmarks.

6.2.4 Process Automation -

Consider a complex system like a power plant or a production facility. Typically, many of the component processes are understood well-enough that they can be controlled by means of a mathematical model. The interaction of these processes frequently must be supervised by a human expert, however, because they are unfit for mathematical formulation. The expert uses heuristic rules or intuition to decide on specific control actions [Mamdani & Assilian (1975), Mamdani (1976), Kickert & Mamdani (1978)]. These rules and intuitions can be put into words and linguistic algorithms. This may be very useful 1) for gaining better understanding of the process interaction, 2) for training new "to-be experts", 3) for further automation of the control process.

6.3 Communication

Consider the transmission of a television image. Great bandwidth is required to transmit a TV image in terms of low-level features, i.e., black and white image points. For the transmission of more specialized information, this bandwidth can be greatly reduced. For example, if we limit our system to transmit text only, we do not have to send information describing the shape of characters in the alphabet. Instead, we can transmit codes representing the characters and give the receiver some general knowledge about text, namely the shape of characters represented by each code, the sequential nature of characters in a text, left to right and top to bottom arrangement of text. This is done in computer terminals capable of displaying text: a character generator interprets the transmitted code and displays the corresponding character in its appropriate position. Observe that transmitter and receiver become "conceptually detached" (as discussed in section 2.2.7) by providing the receiver with "world knowledge". This general principle can be extended by providing the receiver with higher-level knowledge. For example, we may want to build a system which specializes in transmission of cartoon figures. The receiver will be equipped with a cartoon generator instead of a "character" generator. Different facial expressions and spatial relationships between the figures, etc. may be given by a linguistic description which is decoded by the receiver for cartoon generation. We may have receivers with different cartoon generators (as we have computers with different character generators). Thus, one linguistic description may invoke different images.

CHAPTER 7

DISCUSSION AND CONCLUSIONS

This chapter summarizes the main features of the object characterization and interpretation approach that has been presented. The approach is compared with existing computer implementations. A comparison with human descriptions is made and an outlook for future work in this area is presented.

7.1 Summary

We developed a family of object description languages, beginning on a primitive unambiguous level and subsequently introducing additional features to deal with increasingly complex situations. This leads us to L7, the most advanced level we discuss. L7 incorporates linguistic descriptors for multi-dimensional feature spaces. They are designed to be context-adaptive and allow for successful communication even if they have subjective denotation within the communication context.

The semantics of L7 is discussed in detail: the descriptors specify a fuzzy range of possible feature values. The concept of graded possibility is discussed and contrasted to the concept of probability. The advantage of fuzziness in sparsely occupied feature spaces is shown. The role of linguistic labels and operators is discussed. Different methods of specifying the meaning of linguistic descriptors are given and it is explained how these methods can be combined for comprehensive feature representation.

It is discussed how object descriptions in L7 can be interpreted with reference to the given situation context. For comparison of two descriptions, adequacy of the comparison and agreement of the descriptors are distinguished to determine a compatibility value. Matching of two descriptions involves comparison of object descriptors and of set descriptors. It is shown how qualitative matching results can be obtained which can be refined successively and how this matching process can be integrated into a search process to yield the best response.

Design considerations and implementation of linguistic modifiers into the programming language FUZZY are discussed. The objective was to make maximal use of existing control structures while substituting fuzzy possibility distributions for fuzzy set membership values. Examples for this system are suggested. They include applications for research in the "soft sciences" and alternative methods for person-machine and machine-machine communication.

7.2 Comparison With Other Systems

The best known computer system for identification and manipulation of objects is SHRDLU [Winograd (1972, 1973)]. However, since SHRDLU operates in a well-defined, clear-cut artificial blocks world, the problems emphasized in this thesis do not appear. In particular, SHRDLU does not address the issues of a suitable language for complex environments, context-adaptability, and fuzzy descriptors. For this reason, no feature matching problem exists. SHRDLU is implemented in PLANNER [Hewitt (1969, 1972)] one of the crisp models of the language FUZZY.

Shaket (1975) picks up on the SHRDLU paradigm and accounts for the fact that features are determined by fuzzy values, in real applications, rather than by completely crisp ones. His blocks world is described by fuzzy descriptors. Shaket represents linguistic labels by fuzzy set membership functions and determines compatibility of descriptors merely by comparing the membership functions involved. His approach is the "classical" fuzzy set approach in which it is assumed that membership functions for linguistic labels are available in all cases. Shaket deals only with crisp object set descriptors. Thus, his system may yield a definite answer to a retrieval request even if the request is not very well put. The quality of the match is not transparent to the user. The system is implemented in APL, a programming language which invites representation of fuzzy sets as number vectors [compare Wenstop (1976) and Bonissone (1979a)]. APL has built-in operations to manipulate vectors and arrays.

The HAM-RPM system [Hahn et al. (1979)] aims to deal with a multitude of aspects of natural language on a homogeneous level. One of these aspects is fuzziness of referential meaning in adjectives, another is the complexity of real-world problems with respect to number of features and number of objects in the world. Meaning of adjectives is represented and compared much the same way as in Shaket's system, the complexity of the environment is dealt with by focussing mechanisms and spatial arrangement of data, such that only relevant data must reside in main memory of the computer at any given time. HAM-RPM is implemented in FUZZY and the fact that it succeeded very well in becoming a rather flexible, well-segmented system with complex structure and fast response times contributed to the choice of FUZZY as base language for the implementation of linguistic modifiers.

The SWYS system [Hanssmann (1980)] is designed as a natural language interface for real scenes or photographic images. It is also implemented in FUZZY and distinguishes "applicability" and "truth" of descriptors (compare section 4.1).

SOGI [McDermott (1980)] is a program which represents imprecise knowledge of the relative position of objects using a low-resolution, crisp map. McDermott calls this a "fuzzy map", but in fact locations are represented by crisp intervals and his spatial inference system cannot distinguish positions within the interval or recognize that two neighboring locations, one inside the interval and one outside, actually may be very close together, unless the two relations are explicitly related to one another.

7.3 Comparison With Human Approach

An important characteristic of human object descriptions appears to be the fact that in many cases, a description is not intended to represent a precise concept and would serve its purpose better if it were more specific. This becomes apparent if we replace one linguistic descriptor by another without changing the reference of the whole description. We may feel perfectly comfortable with both descriptions even though we would not consider the two descriptors identical with respect to their denotation. Instead, we seem to have reached a limit of resolution in the given context. The precision and crispness of the description is limited by the knowledge about the object and the domain. In order to give a more detailed description, more knowledge has to be taken into account.

On the other hand, linguistic descriptors in human language are not homogeneous entities which can be arranged to form a homogeneous "space of language". This becomes apparent if we take two linguistic descriptors which have similar reference in one context and compare their reference in a variety of contexts. Numerous different aspects may distinguish the two descriptors.

This suggests that reference of linguistic terms has many different levels -- some are rather crude, some are rather fine with respect to a given situation context. We have attempted to accommodate this aspect of human descriptions in a very modest way by allowing the reference of linguistic descriptors to be indicated by reference to related linguistic descriptors rather than exclusively by low-level definitions.

Another aspect of human communication is that so-called "yes - no questions" only rarely are answered by "yes" or "no". Nevertheless, the response that is actually given may be much more meaningful than a "yes - no" response could be. This is, because an approximate question can be given a more precise reference by an appropriate qualifying answer. This allows the questioner to modify his or her model of the environment of the responder, rather than merely filling in an empty slot.

We account for this aspect of the "human approach" with "qualitative linguistic matching" described in chapter 4. Enough answers along the gradual "yes - no" axis are available to reach the limit of "linguistic resolution", i.e., for any answer given by the system, the neighboring answer would be acceptable as well.

7.4 Outlook

Two important aspects of communication about objects by means of linguistic descriptions have been left out in the present study: 1) acquisition of meaning of object descriptors [e.g. Winston (1970), Nevins (1978)], and 2) generation of "good" object descriptions.

In terms of our representation of referential meaning, "meaning acquisition" refers to the process of establishing relations between linguistic labels and modifiers and of setting up meaning approximations in terms of possibility distributions. These relations and approximations may have to be revised as part of a continuous adaptation process. Ideally, we would like to have computer programs which would adapt referential meaning from examples. Lopez de Mantaras (1980a) is working on such a system which is to be implemented in L-FUZZY.

Automatic generation of good object descriptions in complex environments is a very hard but very interesting task. A comprehensive model of common-sense knowledge is required to pick from a large number of features a combination of those which effectively denote the target object in the given situation. This must be done by taking into account features of other objects in the context, preferably without exhaustive search.

8 APPENDICES

APPENDIX A

GLOSSARY

This appendix serves to clarify the exact usage of terms in this dissertation and to point out a different use of the same terms by other authors. The terms that will be exposed are those related to description characteristics, that is to say, to aspects which characterize the relation between a representation and that which it represents. We will define five basic aspects, namely precision, fuzziness, accuracy, descriptiveness, and brevity. First we will discuss each of these aspects and their opposites separately, then we will define our use of terms whose meaning is composed of the five basic aspects, and finally we will point out how the meaning of some of the terms is defined differently in the literature.

PRECISION - IMPRECISION

The terms "precise" and "imprecise" are used in this dissertation to denote a relative degree of granularity. A descriptor is called precise if its feature resolution capability agrees with the feature resolution of its reference data and it is called overprecise if it has higher resolution capability than could be meaningful for referencing corresponding data. Thus, precision is treated as a measure of specificity, in our work.

FUZZINESS - CRISPNESS

The labels "fuzzy" and "crisp" are used to denote the rate of change of applicability with respect to a variation of the reference feature. A descriptor is called crisp (or clear-cut) if there is an abrupt transition from the feature range to which the term applies to the range to which it does not apply. It is called fuzzy, if this transition is gradual. A crisp description does not have to be precise, but a fuzzy descriptor requires a certain degree of imprecision to allow for the gradual transition.

ACCURACY - INACCURACY

We use the terms "accurate" and "inaccurate" to express whether a description is correct or incorrect, respectively. In our usage of the concept, a descriptor can be accurate without being either precise or crisp. This implies that a very general descriptor with ambiguous denotation will be considered accurate if it covers the actual reference features.

DESCRIPTIVENESS - AMBIGUITY

We call a description descriptive if it has definite (or unambiguous) reference in the given domain of discourse. To be descriptive, a reference does not have to be precise, crisp, or accurate.

BREVITY - ELABORATENESS

Brevity of a description has to do with the number of descriptors it contains. A description is called brief or short if it contains few descriptors and it is called elaborate or long if it contains many descriptors.

We view the five aspects discussed above as gradable rather than absolute concepts. They can be combined to form composite aspects, namely sharpness, exactness, vagueness, conciseness, and redundancy. The meaning that we relate to these terms is given below.

SHARPNESS

We call a description sharp if it is both precise and crisp. An extremely precise descriptor always is sharp, since precise descriptors are bound to be non-fuzzy.

EXACTNESS

We call a description exact if it is both precise and accurate, and we call it inexact otherwise.

VAGUENESS

We call a description vague if it is both fuzzy and ambiguous [compare Zadeh (1978b, p.396)]. A fuzzy description per se may be completely descriptive, i.e. unambiguous, with respect to its reference object if the situation context allows for only one interpretation.

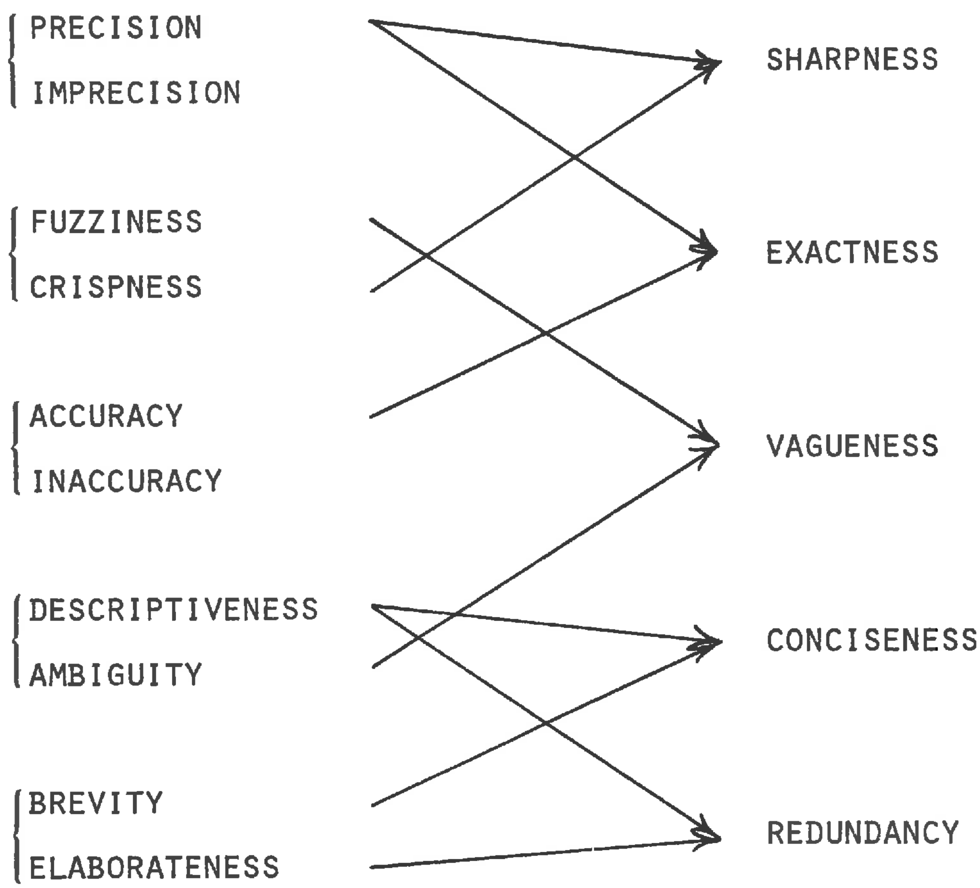
CONCISENESS

We call a description _concise_ if it is both _descriptive and brief_.

REDUNDANCY

We call a description _redundant_ if it is _descriptive and elaborate_, i.e., if a shorter description could be found which would be descriptive as well.

Our terminology for characterizing descriptions is summarized in the following table:



On the left, we have pairs of contrasting basic aspects. Within each pair, the terms refer to the same aspect dimension. On the right you find composite aspects which are derived from the basic ones.

Of particular interest are the following relationships between some of the characteristics:

$$(1) \qquad \text{vague} \subset \text{fuzzy} \subset \text{imprecise}$$

and

$$(2) \qquad \text{ambiguous} < \text{concise} < \text{redundant}$$

(1) indicates that each vague descriptor is fuzzy and each fuzzy descriptor is imprecise, and (2) indicates that by enriching a description, an ambiguous descriptions may become concise, and a concise description may become redundant.

Russell's terminology

In his paper on "vagueness", Russell (1923) employs a different terminology which we can translate into the terms above. Russell states (p.91), "a belief is <u>precise</u> when only one fact would verify it; it is <u>accurate</u>, when it is both precise and true." His concept of precision agrees with our concept, but he calls <u>accurate</u> what we would call <u>exact</u>. Russell's concept of accuracy does not allow for ambiguity of reference, but ours does.

Russell uses the term vague to indicate that a representation is not (Russell-) accurate (p.89), which in his terms would mean, that it is either imprecise or untrue, or both, i.e. "inexact", in our terms. In addition, Russell states that in a vague representation, "there is not one definite fact necessary and sufficient for its truth, but a certain region of possible facts, any one of which would make it true. And this region is itself ill-defined: we cannot assign to it a definite boundary. This is the difference between vagueness and generality" (p.88). Thus, what Russell calls "vague" we would call "fuzzy", if it is imprecise and does not have clear-cut boundaries, and we would call it "vague" if it is ambiguous in reference, in addition.

The reason that Russell does not discriminate between general fuzziness and the special case of vagueness stems from the fact that he relates descriptions to conceivable features in a universe of discourse rather than to actual features in the particular domain of reference. Thus, an imprecise description is for him automatically ambiguous, since it could refer to several conceivable instances while we would consider an imprecise description ambiguous only if in the given context several features actually exist to which the description could refer.

APPENDIX B

INTRODUCTION TO FUZZY

8.1 Central Ideas

>>FUZZY is an attempt to synthesize many of the 'good ideas' of previous AI languages while providing facilities for the efficient storage, retrieval, and manipulation of knowledge which is vague and uncertain in nature. The language was constructed in such a way that it could be directly embedded in LISP (unlike previous systems), and is therefore much more efficient than languages such as MICRO-PLANNER which require a run-time monitor. LISP and FUZZY primitives may be freely intermixed, and FUZZY functions may be called from compiled LISP code if desired.<< [LeFaivre (1977, p.2)].

FUZZY was motivated by the theory of fuzzy sets [Zadeh (1965-1973)]. Fuzzy set theory is a generalization of Boolean set theory and allows for graded membership of elements in a set rather than all-or-none set membership. This is to account for the fact that many concepts dealt with by natural language are unsharp in the sense that there is no sharp boundary between situations for which the concept applies and situations for which it does not apply. Consider, for example, the concept "young". If we talk about people, we may say that persons of less than 10 years of age are young and persons above 70 years are not young. However, there is no particular day at which a person's age switches from 'young' to 'not young'; rather, this is a gradual transition. In fuzzy set theory the concept "young" in this context could be expressed by a "membership function" representing the degree to which a person of a particular age can be considered to be young:

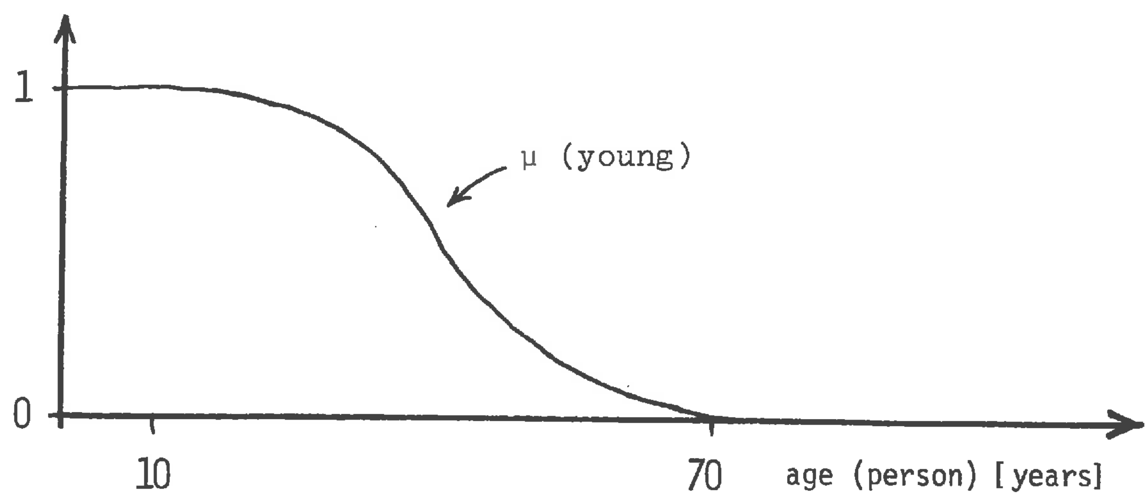


Fig. B.1 Membership function for the fuzzy set of young people

Several AI systems deal explicitly with unsharp information (see for example MYCIN [Shortliffe et al. (1973)]) and with incomplete information (heuristic programs). The language FUZZY is designed to relieve the user from explicitly dealing with both, the available information and its associated modifier (be it a degree of set-membership or certainty, a possibility or probability measure, or an indication of confidence in a decision).

FUZZY has been implemented on a UNIVAC 1110 [LeFaivre (1974a)], on DECsystem-10 [LeFaivre (1977)] DECsystem-20 computers [Freksa (1977)], and most recently on a CDC CYBER 73 computer [Okseniuk (1980)]. It is in use for various AI projects, for example for the AIMDS/BELIEVER system at Rutgers University and for HAM-RPM, a knowledge-based conversationalist at the University of Hamburg.

8.2 Description Of The Language

8.2.1 Fuzzy Expressions And Local Control -

As in PLANNER and other AI languages [Bobrow & Raphael (1974)], expressions in FUZZY may either succeed or fail. In case of success, however, a FUZZY expression may not only return a value, but also a so-called Z-value. This is to allow for the representation of vague concepts by means of implicitly weighting expressions. Internally, a FUZZY expression is represented by a LISP dotted pair, e.g.,

((CHANCE OF RAIN) . 0.3)

may represent a 30% chance of rain. Failure of an expression is indicated by simply returning FAIL instead of the dotted pair. Functions are available to extract the value portion or the Z-value portion of an arbitrary expression. In order to make use of standard LISP predicates, a value of either FAIL or NIL is interpreted as failure.

Like PLANNER, FUZZY has automatic backtracking facilities. However, in many cases, automatic backtracking is undesirable. Therefore, FUZZY allows the user to specify where automatic backtracking should occur and where not. Language primitives are available to compute union, intersection, and complement of fuzzy sets.

8.2.2 The Associative Net -

FUZZY maintains an associative network of assertions quite similar to PLANNER or CONNIVER. Any arbitrary LISP list structure may be entered into this net. In addition, an assertion may have a Z-value associated with it, if desired. The Z-value of the assertions can be used to control success and failure of the FETCH statement or subsequent actions.

8.2.3 Pattern Matching -

As in PLANNER/CONNIVER, a FUZZY variable is assigned a value via pattern matching. For example,

```
(MATCH (?X ??Y) ((A B) C D))
```

binds the FUZZY variable !X to (A B) and !Y to (C D). A greater variety of functions than in PLANNER/CONNIVER is available to place restrictions on the structure of the pattern or the set of items which can match successfully. For example,

```
(*R ?OBJECT (FETCH (RED !OBJECT)))
```

will only match an object which is known to be red.

8.2.4 Contexts And Backtrack States -

It is often convenient to maintain several different contexts with the ability to easily switch from one to another. FUZZY has such a "context mechanism" which activates and deactivates associative nets of assertions. It is also possible to save the state of the entire system in order to allow for later restoration. Functions are available to compute differences between states and to add differences to a state. State changes can be finalized such that they cannot be undone by a subsequent restoration. This feature is useful to prevent backtracking getting out of control. Several FUZZY primitives exist in backtrackable and in finalizing versions to give the programmer easy control over the global control mechanism.

8.2.5 Procedures -

Similar as a MICRO-PLANNER theorem or a QLISP QLAMBDA function, a FUZZY procedure takes an argument which is matched against a procedure pattern. If the match is successful the procedure is entered, otherwise it fails. A FUZZY procedure may be called by name or may be invoked by pattern (-> section 8.2.7). All FUZZY variables are assumed to be local to the procedure in which they appear unless they are declared to be global in the procedure head or by an external GLOBAL declaration. With each procedure there is a procedure demon associated (-> section 8.2.6). A FUZZY procedure either succeeds or fails; in case of success, the procedure pattern is returned as value along with a Z-value supplied by the procedure demon unless specified otherwise. Any other pattern may be returned with or without Z-value.

8.2.6 Global Control -

FUZZY procedures have a more general global control mechanism than PLANNER theorems. The procedure demons are given control not only upon failure of an expression (as in MICRO-PLANNER) but also upon successful termination of a (top-level) expression. This makes it possible to globally evaluate the results returned by the expressions of a procedure. With each procedure a variable is associated which maintains an "accumulated Z-value" for the demon's calculations. The procedure demon is called a last time when the procedure is exited in order to make any necessary final computations (for example concluding statistics).

8.2.7 Deductive Mechanisms -

There are several levels of accessing information in a knowledge base:

1. request of explicitly available information
2. invocation of an explicit procedure to retrieve the desired information
3. specifying a goal and leaving it up to the system how to achieve it

All three methods are possible in FUZZY:

1. the FETCH statement retrieves assertions which are explicitly stored in the associative net by pattern matching
2. FUZZY procedures can be called by name

3. If FUZZY procedures have been stored in the associative net, they can be invoked by pattern matching via the DEDUCE statement. This reliefs the programmer from keeping track of which particular procedures may be used to achieve a certain task and it allows for easy addition of new procedures to the associative net which automatically can be utilized by existing programs without change.

4. The GOAL statement combines the FETCH and DEDUCE statements: first it looks whether the desired information is explicitly available in the net of assertions. If it fails it attempts to deduce the goal by invoking DEDUCE procedures which match the specified pattern.

In addition, FUZZY supports ASSERT and ERASE procedures which are automatically invoked when assertions of corresponding patterns are added and removed from the associative net, respectively.

8.3 Example

The following program illustrates how FUZZY may deal with both vague and incomplete information. The vagueness is expressed here by Z-values associated with assertions.

Incomplete information in this example is manifested by the absence of useful assertions. This 'missing information' does not force the procedure into failure, but rather lowers the confidence in the result obtained by the procedure:

```
 === NET ===

((CHANCE OF RAIN) . 0.8)
((DRYNESS DESIRED) . 0.7)
((BLUE SKY) . 0.4)

 === DEDUCE ===

(PROC NAME: UMBRELLA DEMON: CONFIDENCE (RAIN PROTECTION)
      (BIND ?SK (FETCH ((*OR CLEAR BLUE) SKY)))
      (BIND ?BU (FETCH (BURDENSOME UMBRELLA)))
      (BIND ?DD (FETCH (? DESIRED)))
      (IF (ZAND THRESH: 0.9 (ZNOT !SK) !DD !BU)
    THEN: (SUCCEED '"STAY HOME!")
    ELSE: T)
      (BIND ?CR (FETCH (CHANCE ??)))
      (IF (MINUSP (DIFFERENCE (PLUS (ZVAL !CR) (ZVAL !DD))
                              (PLUS (ZVAL !SK) (ZVAL !BU))))
    THEN: (SUCCEED '"DON'T TAKE UMBRELLA" ZACCUM)
    ELSE: (SUCCEED '"TAKE UMBRELLA" ZACCUM)))

 ===========

(DEFPROP CONFIDENCE
 (LAMBDA (RESULT THRESHOLD C-LEVEL)
  (COND [(EQ RESULT FAIL) (COND [(*GREAT C-LEVEL 0.25)
                                  (DIFFERENCE C-LEVEL 0.25)]
                                 [T (FAIL)])]
        [(EQ RESULT DONE) C-LEVEL]
        [(*LESS (ZVAL RESULT) THRESHOLD) (FAIL)]
        [T C-LEVEL]))
 EXPR)
     (;  RESULT    = result of last top-level expression
         THRESHOLD = criterion for forcing procedure to fail
         C-LEVEL   = current confidence level)
```

The computer listing shows in the beginning the associative net, containing some declarative knowledge about a potential umbrella carrier and his situation. Below you see the procedural associative net containing a deduce procedure to give advise whether or not to carry an umbrella in a given situation. On the bottom there is a LISP procedure which is used by the deduce procedure UMBRELLA as a procedure demon. The procedure UMBRELLA uses assertions and their modifiers to calculate the projected pay-off for carrying an umbrella. The demon CONFIDENCE watches this calculation and determines a confidence measure for the result obtained by UMBRELLA. This is done as follows: UMBRELLA looks for four types of assertions in the associative net:

1. information about the blueness or the clearness of the sky

2. information about the burden to carry an umbrella

3. information about a desired goal which can be satisfied with an umbrella

4. information about the chance of the occurrence of an event which would make an umbrella desirable

The most reliable advise can be given by UMBRELLA if all four pieces of information can be found. If a piece of information cannot be found (i.e., if the corresponding FETCH returns FAIL), the demon reduces the confidence level ZACCUM which is returned as Z-value of UMBRELLA. Observe that the Z-value can be used in a single program to do different kinds of qualifications.

REFERENCES

Nullum est iam dictum,
quod non sit dictum prius.

Eunuch, Prolog, 41

ADAMO, J.M., 1980. L.P.L. A fuzzy programming language: 1. syntactic aspects. Fuzzy Sets and Systems 3, 151-179.

AKIN, O., REDDY, R., 1977. Knowledge acquisition for image understanding research. Computer Graphics and Image Processing 6, 307-334.

BAJCSY, R., ROSENTHAL, D., 1975. Visual focusing and defocusing - an essential part of pattern recognition process. Proc. Comp. Graph., Patt. Rec., & Data Struct. Conf., UCLA.

BAJCSY, R., TAVAKOLI, M., 1975. Image filtering - a context dependent process. IEEE Trans. CAS-22, 5, 463-474.

BALLARD, D.H., 1978. Strip trees: a hierarchical representation for map features. TR 32, Computer Science Department, University of Rochester.

BARROW, H.G., AMBLER, A.P., BURSTALL, R.M., 1972. Some techniques for recognizing structures in pictures. In: Watanabe, S. (ed.), Frontiers of pattern recognition, Academic Press.

BARROW, H.G., TENENBAUM, J.M., 1975. Representation and use of knowledge in vision. TN 108, AI Center, Stanford Research Institute, Menlo Park.

BARTRAM, D.J., 1973. The effects of familiarity and practice on naming pictures of objects. Memory & Cognition 1, 2, 101-105.

BELLMAN, R., KALABA, R., ZADEH, L., 1966. Abstraction and pattern classification. J. Math. Anal. and Applic. 13, 1-7.

BELLMAN, R.E., ZADEH, L.A., 1977. Local and fuzzy logics. In: Dunn, J.M., Epstein, G. (eds.), Modern uses of multiple-valued logic, 103-165.

BERKELEY, G., 1709. An essay towards a new theory of vision. Dublin.

BEZDEK, J.C., CASTELAZ, P.F., 1977. Prototype classification and feature selection with fuzzy sets. IEEE Trans. SMC-7, 2, 87-92.

BLACK, M., 1937. Vagueness. An exercise in logical analysis. Philosophy of Science 4, 427-455. Also in: Black (1949), 23-58.

BLACK, M., 1949. Language and philosophy. Cornell University Press.

BLACK, M., 1963. Reasoning with loose concepts. Dialogue 2, 1, 1-12. Also in: Black (1970), 1-13.

BLACK, M., 1970. Margins of precision. Cornell University Press.

BOBROW, D.G., COLLINS, A., 1975. Representation and Understanding. Academic Press, New York.

BOBROW, D.G., NORMAN, D.A., 1975. Some principles of memory schemata. In: Bobrow & Collins (1975), 131-149.

BOBROW, D.G., RAPHAEL, B., 1974. New programming languages for AI research. Computing Surveys 6, 153-174.

BODEN, M.A., 1977. Artificial intelligence and natural man. Basic Books, New York.

BOLINGER, D., 1972. Degree words. Mouton, Paris.

BONISSONE, P.P., 1979a. The problem of linguistic approximation in system analysis. Ph.D. thesis, EECS Department, University of California, Berkeley.

BONISSONE, P.P., 1979b. A pattern recognition approach to the problem of linguistic approximation in system analysis. IEEE Conf. System, Man, and Cybernetics, Denver.

BONISSONE, P.P., 1980. Linguistic approach and its related problems. General Electric Corp. Res. & Develop., Schenectady.

BREEDING, K.J., AMOSS, J.O., 1972. A pattern description language - PADEL. Pattern Recognition 4, 19-36.

BREMERMANN, H.J., 1972. Artificial intelligence, pattern recognition and neurobiology. Int. J. Neuroscience 3, 147-154.

BUCHANAN, B.G., SUTHERLAND, G., FEIGENBAUM, E., 1969. Heuristic DENDRAL: A program for generating explanatory hypotheses in organic chemistry. In B. Meltzer & D. Michie (eds.) Machine Intelligence vol. 4, 209-254, American Elsevier, New York.

BUCKLES, B.P., 1979. Fuzzy relational databases: a foundational framework and information-theoretic characterization. Internal Memo. IM-2265, General Research Corporation, Huntsville.

CARROLL, J.M., 1979a. Natural strategies in naming. Research Report RC 7533, IBM Watson Research Center, Yorktown Heights, N.Y.

CARROLL, J.M., 1979b. Naming and describing in social communication. Research Report RC 7685, IBM Watson Research Center, Yorktown Heights, N.Y.

CARROLL, L., 1865. Alice's adventures in wonderland. Macmillan.

CLIFF, N., 1959. Adverbs as multipliers. Psych. Review, vol. 66, No.1, 27-44.

CRITCHLEY, M., 1975. Language. In: Lenneberg & Lenneberg (1975) p. 3 - 16.

DEERING, M.F., 1980. Qualifying examination proposal. EECS Dept., University of California, Berkeley.

DOYLE, A.C., 1901. The hound of the Baskervilles. Ballantine Books, New York 1975.

DREYFUS, H.L., 1979. What computers can't do. Harper & Row, New York.

DUDA, R.O., HART, P.E., 1973. Pattern classification and scene analysis. John Wiley.

DUDA, R.O., HART, P.E., BARRETT, P., GASCHNIG, J.G., KONOLIGE, K., REBOH, R., SLOCUM, J., 1978. Development of the PROSPECTOR consultation system for mineral exploration. Final Report, SRI International, Menlo Park.

EKMAN, P., FRIESEN, W.V., 1978. Facial action coding system. Consulting Psychologists Press, Palo Alto.

FEIGENBAUM, E.A., 1977. The art of artificial intelligence: 1. themes and case studies of knowledge engineering. Memo HPP-77-25, Heuristic Programming Project, Stanford. Also in: IJCAI-5, 1014-1029, MIT, Cambridge.

FEIGENBAUM, E.A., 1980. Panel discussion on cognitive science methodology. Cognitive Science Conference, Yale.

FEIGENBAUM, E.A., FELDMAN, J., 1963. Computers and thought. Mc Graw Hill, New York.

FIRSCHEIN, O., FISCHLER, M.A., 1971. Describing and abstracting pictorial structures. Pattern Recognition 3, 421-443.

FIRSCHEIN, O., FISCHLER, M.A., 1972. A study in descriptive representation of pictorial data. Pattern Recognition 4, 361-377.

FREGE, G., 1879. Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens, L. Nebert, Halle. English translation in: Conceptual notation and related articles, T.W. Bynum (ed.), Oxford 1972.

FREGE, G., 1892. Ueber Sinn und Bedeutung. Zeitschrift fuer Philosophie und philosophische Kritik, vol. 100, 25-50, Leipzig.

FREKSA, C., 1977. KI-Rundbrief 10, p.9.

FREKSA, C., 1980a. L-FUZZY -- an AI-language with linguistic modification of patterns. Memo No. UCB/ERL M80/10 Electronics Research Lab, University of California, Berkeley.

FREKSA, C., 1980b. Communication about visual patterns by means of fuzzy characterizations. XXIInd Intern. Congress of Psychology, Leipzig.

FREUDER, E.C., 1976. A computer system for visual recognition using active knowledge. AI-TR-345 (thesis). MIT, Cambridge.

GAINES, B.R., 1976. Foundations of fuzzy reasoning. Int. J. Man-Machine Studies 8, 623-668.

GAINES, B.R., KOHOUT, L.J., 1975. Possible automata. Proc. Int. Symposion on multiple-valued logic, University of Indiana, Bloomington, 183-196.

GAINES, B.R., KOHOUT, L.J., 1977. The fuzzy decade: a bibliography of fuzzy systems and closely related topics. Int. J. Man-Machine Studies 9, 1-68.

GARVEY, T.D., 1976. Perceptual strategies for purposive vision. TN 117, AI Center, Stanford Research Institute, Menlo Park.

GOGUEN, J.A., 1967. L-fuzzy sets. J. Math. Anal. and Applic. 18, 145-174.

GOGUEN, J.A., 1969. The logic of inexact concepts. Synthese 19, 325-373.

GOGUEN, J.A., 1974. Concept representation in natural and artificial languages: axioms, extensions and applications for fuzzy sets. Int. J. Man-Machine Studies 6, 513-561.

GOGUEN, J.A., 1976. Robust programming languages & the principle of maximal meaningfulness. The Milwaukee Symposium on Automatic Computation and Control.

GUPTA, M.M., RAGADE, R.K., YAGER, R.R., 1979a. Advances in fuzzy set theory and applications. North Holland.

GUPTA, M.M., RAGADE, R.K., YAGER, R.R., 1979b. Report on the IEEE symposium on fuzzy set theory and applications. Fuzzy Sets and Systems 2, 105-111.

GUPTA, M.M., SARIDIS, G.N., GAINES, B.R., 1977. Fuzzy automata and decision processes. North Holland, New York.

GUZMAN, A., 1968. Computer recognition of three-dimensional objects in a visual scene. MAC-TR-59 (thesis), Project MAC, MIT, Cambridge.

HAACK, S., 1979. Do we need "fuzzy logic"? Int. J. Man-Machine Studies 11, 437-445.

HAAR, R.L., 1977. A fuzzy relational data base system. Technical Report TR-586, Computer Science Center, University of Maryland.

HAHN, W.v., HOEPPNER, W., JAMESON, A., WAHLSTER, W., 1979. The anatomy of the natural language dialogue system HAM-RPM. In: Bolc, L. (ed.) Natural language based computer systems. Hanser/Macmillan Verlag Muenchen, London.

HAMMING, R.W., 1950. Error detecting and error correcting codes. Bell Sys. Tech. J. 29, 147-160.

HANSON, A.R., RISEMAN, E.M., 1974. Processing cones: a parallel computational structure for scene analysis. COINS Tech. Report 74C-7, University of Massachusetts, Amherst.

HANSSMANN, K.-J., 1980. Sprachliche Bildinterpretation fuer ein Frage - Antwort System. Mitteilung IfI-HH-M-74/80, Institut fuer Informatik, Universitaet Hamburg.

HARMON, L.D., HUNT, W.F., 1977. Automatic recognition of human face profiles. Computer Graphics and Image Processing 6, 135-156.

HEMPEL, C.G., 1939. Vagueness and logic. Philosophy of Science 6, 2, 163-180.

HERSH, H.M., CARAMAZZA, A., 1976. A fuzzy set approach to modifiers and vagueness in natural language. J. Exp. Psych: Gen. 105, 3, 254-276.

HEWITT, C., 1969. PLANNER: a language for proving theorems in robots. IJCAI-1, Washington.

HEWITT, C., 1972. Description and theoretical analysis (using schemata) of PLANNER. Ph.D. thesis, Department of Mathematics, MIT, Cambridge.

HOFFMANN, J., 1980. Conceptual coding, representation and process. XXIInd Intern. Congress of Psychology, Leipzig.

HUGHES, G.E., CRESSWELL, M.J., 1968. An introduction to modal logic. Methuen, London.

IMAOKA, H., SUGENO, M., 1979. A model of dialog based on fuzzy set concept. IJCAI-6, 419-421, Tokyo.

JAIN, R., NAGEL, H.-H., 1977. Analyzing a real world scene sequence using fuzziness. Proc. of the 1977 IEEE Conf. on Decision and Control, 2, 1367-1372, New Orleans.

KANDEL, A., BYATT, W.J., 1978. Fuzzy sets, fuzzy algebra, and fuzzy statistics. Proc. IEEE 6, 12, 1619-1639.

KAY, P., 1979. The role of cognitive schemata in word meaning: hedges revisited. Unpublished paper, Department of Anthropology, University of California, Berkeley.

KELLY, G.A., 1955. The psychology of personal constructs. Norton & Co., New York.

KEMPSON, R.M., 1977. Semantic theory. Cambridge University Press.

KICKERT, W.J.M., 1979. Towards an analysis of linguistic modelling. Fuzzy Sets and Systems 2, 293-307.

KICKERT, W.J.M., KOPPELAAR, H., 1976. Application of fuzzy set theory to syntactic pattern recognition of handwritten capitals. IEEE Trans. SMC-6, 2, 148-151.

KICKERT, W.J.M., MAMDANI, E.H., 1978. Analysis of a fuzzy logic controller. Fuzzy Sets and Systems 1, 29-44.

KLING, R., 1974. Fuzzy - PLANNER: reasoning with inexact concepts in a procedural problem-solving language. J. Cybernetics 4, 2, 105-122.

KOCHEN, M., 1978. Enhancement of coping through blurring. Fuzzy Sets and Systems 2, 37-52.

KOCHEN, M., BADRE, A.N., 1974. On the precision of adjectives which denote fuzzy sets. J. Cybernetics 4, 1, 49-59.

KUIPERS, B.J., 1977. Representing knowledge of large-scale space. Ph.D. thesis, AI-TR-418, AI Lab, MIT.

KUIPERS, B., 1978. Modeling spatial knowledge. Cognitive Science 2, 129-153.

LAKOFF, G., 1973. Hedges: a study in meaning criteria and the logic of fuzzy concepts. J. Philos. Logic 2, 458-508.

LeFAIVRE, R.A., 1974a. Fuzzy problem solving. TR 37, Department of Computer Science, University of Wisconsin, Madison.

LeFAIVRE, R.A., 1974b. The representation of fuzzy knowledge. J. Cybernetics 4, 2, 57-66.

LeFAIVRE, R.A., 1977. FUZZY reference manual. Department of Computer Science, Rutgers University.

LENNEBERG, E.H., 1975. The concept of language differentiation. In: Lenneberg & Lenneberg (1975) 17 - 33.

LENNEBERG, E.H., LENNEBERG, E., 1975. Foundations of language development -- a multidisciplinary approach. Vol. 1, Academic Press, New York.

LOPEZ de MANTARAS, R., 1980a. Learning the meaning of imprecise concepts by adaptation of their possibility distributions. Actes du table ronde du CNRS "Quelques applications concretes utilisant les derniers perfectionnements de la theorie du flou", Universite Claude Bernard, Lyon.

LOPEZ de MANTARAS, R., 1980b. Translation of natural language descriptions of simple objects into PRUF. XXIInd Intern. Congress of Psychology, Leipzig.

MacVICAR-WHELAN, P.J., 1976. Fuzzy sets for man-machine interaction. Int. J. Man-Machine Studies 8, 687-697.

MAMDANI, E.H., 1976. Advances in the linguistic synthesis of fuzzy controllers. Int. J. Man-Machine Studies 8, 669-678.

MAMDANI, E.H., ASSILIAN, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man-Machine Studies 7, 1-13.

MANDIC, N., ESHRAGH, F., MAMDANI, E.H., 1980. Multi-criteria decision making using fuzzy sets. Department of Electrical & Electronic Engineering, Queen Mary College, London.

McCARTHY, J., 1960. Recursive functions of symbolic expressions and their computation by machine. Part I, CACM 3, No. 4, 184-195.

McDERMOTT, D., 1980. Metric model of spatial processing. Cognitive Science Conference, Yale. Also in: Proc. AAAI-1, Stanford.

McDERMOTT, D.V., SUSSMAN, G.J., 1972. The CONNIVER reference manual, MIT AI lab Memo No. 259.

MILLER, G.A., 1967. The psychology of communication. Basic Books.

MINSKY, M., 1974. A framework for representing knowledge. AI Memo No. 306, MIT, Cambridge. Also in: Winston (1975).

MUIR, J., GREGG, T., 1975. How to keep your Volkswagen alive! John Muir Publications, Santa Fe.

MUNSON, J.H., 1971. Robot planning, execution, and monitoring in an uncertain environment. Technical Note 59, Artificial Intelligence Group, Stanford Research Institute, Menlo Park. Also in: IJCAI-2, London.

NAGEL, H.-H., 1976. Formation of an object concept by analysis of systematic time variations in the optically perceptible environment. Bericht IfI-HH-B-27/76, Institut fuer Informatik, Universitaet Hamburg.

NEISSER, U., 1967. Cognitive psychology. Appleton-Century-Crofts, New York.

NEISSER, U., 1976. Cognition and reality. Freeman, San Francisco.

NEVINS, A.J., 1978. Learning patterns from image descriptions. Tech. Report No. 1, Department of Information Systems, Georgia State University, Atlanta.

NEWELL, A., SHAW, J.C., SIMON, H.A., 1960. Report on a general problem-solving program for a computer. Information processing: Proc. Intl. Conf. Info. Proc. 256-264, UNESCO, Paris.

NEWELL, A., SIMON, H., 1961. GPS, a program that simulates human thought. In: H. Billing, ed., Lernende Automaten, 109-124, Oldenbourg, Muenchen. Reprinted in: Feigenbaum & Feldman (1963).

NEWELL, A., SIMON, H.A., 1972. Human problem solving. Prentice Hall.

NEWELL, A., SIMON, H.A., 1975. Computer science as empirical inquiry: symbols and search. ACM Turing Award Lecture, CACM 19, 3, 113-126.

NILSSON, N.J., 1971. Problem solving methods in artificial intelligence. McGraw-Hill, New York.

NILSSON, N.J., 1974. Artificial intelligence. In: Rosenfeld, J.L. (ed.), Applic. in the soc. sci. and the humanities, Proc. IFIP Congr. 4, 778-801, American Elsevier, New York.

NORMAN, D.A., BOBROW, D.G., 1975. On data-limited and resource-limited processes. Cognitive Psychology 7, 44-64.

NORMAN, D.A., RUMELHART, D.E., 1975. Explorations in cognition. Freeman, San Francisco.

OKSENIUK, A., 1980. FUZZY implementation for CDC CYBER 73. KI Rundbrief 22, p.33.

PALMER, S.E., 1975. The effects of contextual scenes on the identification of objects. Memory & Cognition 3, 5, 519-526.

PALMER, S.E., 1977. Hierarchical structure in perceptual representation. Cognitive Psychology 9, 441-474.

PALMER, S.E., 1978. Fundamental aspects of cognitive representation. In: Rosch, E., Lloyd, B. (eds.), Cognition and categorization, Erlbaum.

POPPER, K.R., 1965. Of clouds and clocks. In: K.R. Popper (1972). Objective knowledge. Oxford University Press, 206-255.

POPPER, K.R., 1976. Unended quest. Fontana, London.

RAGADE, R.K., 1976. Fuzzy sets in communication systems and in consensus formation systems. J. Cybernetics 6, 21-38.

RAGADE, R.K., GUPTA, M.M., 1977. Fuzzy set theory: introduction. In: Gupta, Saridis, Gaines (1977).

RAPHAEL, B., 1976. The thinking computer - mind inside matter, Freeman, San Francisco.

REBALL, S., 1978. Redundante Indexiersprachen. Forschungsinstitut Manfred v. Ardenne, Dresden.

RHODES, M.L., KLINGER, A., 1977. Conversational text input for modifying graphics facial images. In: SIGART (1977), 58-61.

RIEGER, B.B., 1980. Fuzzy word meaning analysis and representation in linguistic semantics. 8th Int. Conf. on Comp. Ling. (COLING-80), Tokyo.

RIESBECK, C.K., 1980. "You can't miss it!": judging the clarity of directions. Cognitive Science 4, 3, 285-303.

ROBERTS, L.G., 1963. Machine perception of three-dimensional solids. Technical report No. 315, Lincoln Laboratory, MIT. Also in: Optical and Electro-optical information processing, J.T. Tippett et al. (eds.), MIT Press, Cambridge, Mass. 1965, 159-197.

ROSCH, E., MERVIS, C.B., 1975. Family resemblances: studies in the internal structure of categories. Cognitive Psychology 7, 573-605.

RUSSELL, B., 1905. On denoting. MIND, N.S. 14, 479-493.

RUSSELL, B., 1923. Vagueness. Australasian J. Psych. & Phil. 1, 1, 84-92.

RUSSELL, B., 1948. Human knowledge. Its scope and limits. Simon & Schuster, New York.

SANCHEZ, E., 1978. On possibility qualification in natural languages. Information Sciences 15, 45-76.

SCHANK, R.C., COLBY, K.M., 1973. Computer models of thought and language. Freeman, San Francisco.

SCHEFE, P., 1980. On foundations of reasoning with uncertain facts and vague concepts. Int. J. Man-Machine Studies, 12, 35-62.

SHAKET, E., 1975. Fuzzy semantics for a natural like language defined over a world of blocks. Master thesis, Comp. Sci. Dept., UCLA.

SHANNON, C.E., WEAVER, W., 1949. The mathematical theory of communication. Univ. of Illinois Press, Urbana.

SHORTLIFFE, E.H., et al., 1973. An artificial intelligence program to advice physicians regarding antimicrobial therapy. Computers & Biomedical Research vol. 6, 544-560.

SIGART, 1977. Newsletter No. 61 of the ACM special interest group on artificial intelligence.

SUSSMAN, G.J., McDERMOTT, D.V., 1972. From PLANNER to CONNIVER -- a genetic approach, Proc. AFIPS FJCC 41, 1171-1180.

STARK, L., DICKSON, J.F., 1966. Mathematical concepts of central nervous system function. In: Schmitt, F.O. & Melnechuk, T., eds. Neurosciences research symposium summaries, vol. 1, 109-178, MIT Press, Cambridge.

TENENBAUM, J.M., 1973. On locating objects by their distinguishing features in multisensory images. TN 84, AI Center, Stanford Research Institute, Menlo Park.

TONG, R.M., BONISSONE, P.P., 1979. Linguistic decision analysis using fuzzy sets. Memo No. UCB/ERL M79/72 Electronics Research Lab, University of California, Berkeley.

TRIBUS, M., 1979. Comments on "Fuzzy sets, fuzzy algebra, and fuzzy statistics." Proc. IEEE, 67, 8, 1168.

TURNER, K.J., 1974. Computer perception of curved objects using a television camera. Ph.D. thesis, Department of Machine Intelligence, School of Artificial Intelligence, University of Edinburgh.

UMANO, M., MIZUMOTO, M., TOYODA, J., TANAKA, K., 1975. On interpretation and execution of fuzzy programs. Technical Report on Automation and Language of IECE, AL 75-27.

URAGAMI, M., MIZUMOTO, M., TANAKA, K., 1976. Fuzzy robot controls. J. Cybernetics 6. 39-64.

VALENTIN, K., 1929. Der Radfahrer. In: Alles von Karl Valentin, Piper, Muenchen, 1979.

WAHLSTER, W., 1977. Die Repraesentation von vagem Wissen in natuerlichsprachlichen Systemen der kuenstlichen Intelligenz. Bericht IfI-HH-B-38/77, Institut fuer Informatik, Universitaet Hamburg.

WAHLSTER, W., 1978. Die Simulation vager Inferenzen auf unscharfem Wissen: eine Anwendung der mehrwertigen Programmiersprache FUZZY. Bericht Nr. 5, Proj. Simulation von Sprachverstehen, Germanisches Seminar, Universitaet Hamburg.

WAHLSTER, W., 1979. Implementing fuzziness in dialogue systems. Bericht Nr. 14, Projekt Simulation von Sprachverstehen, Germanisches Seminar, Universitaet Hamburg.

WATANABE, S., 1978. A generalized fuzzy set theory. IEEE Trans. SMC-8, 10, 756-760.

WECHSLER, H., 1976. A fuzzy approach to medical diagnosis. Int. J. Bio-Medical Computing 7, 191-203.

WEISSMAN, C., 1967. LISP 1.5 primer. Dickenson, Belmont.

WENSTOP, F., 1976. Deductive verbal models of organizations. Int. J. Man-Machine Studies 8, 293-311.

WINOGRAD, T., 1972. Understanding natural language. Academic Press, New York.

WINOGRAD, T., 1973. A procedural model of language understanding. In: Schank & Colby (1973), 152-186.

WINSTON, P.H., 1970. Learning structural descriptions from examples. MAC-TR-76 (thesis). Project MAC, MIT, Cambridge. Also in: Winston (1975), 157-209.

WINSTON, P.H. (ed.), 1975. The psychology of computer vision. McGraw Hill, New York.

WINSTON, P.H., 1977. Artificial Intelligence. Addison-Wesley.

WITTGENSTEIN, L., 1922. Tractatus logico-philosophicus. Suhrkamp, Frankfurt 1964.

WITTGENSTEIN, L., 1951. Ueber Gewissheit - on certainty. Basil Blackwell, Oxford, 1969.

WITTGENSTEIN, L., 1953. Philosophische Untersuchungen, Suhrkamp, Frankfurt 1971.

ZADEH, L.A., 1965. Fuzzy sets. Information and Control 8, 338-353.

ZADEH, L.A., 1970. Fuzzy languages and their relation to human and machine intelligence. Man and Computer Proc. int. Conf., 130-165, Bordeaux.

ZADEH, L.A., 1971. Similarity relations and fuzzy orderings. Information Sciences 3, 177-200.

ZADEH, L.A., 1972. A fuzzy-set-theoretic interpretation of linguistic hedges. J. Cybernetics 2, 3, 4-34.

ZADEH, L.A., 1973. Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. SMC-3, 1, 28-44.

ZADEH, L.A., 1974. On the analysis of large-scale systems. Memo No. ERL-M418, Electr. Res. Lab, Univ. of Calif., Berkeley.

ZADEH, L.A., 1975a. The concept of a linguistic variable and its application to approximate reasoning. I Info. Sciences 8, 199-249. II Info. Sciences 8, 301-357. III Info. Sciences 9, 43-80.

ZADEH, L.A., 1975b. Calculus of fuzzy restrictions. Memo No. ERL-M502, Electr. Res. Lab, Univ. of Calif., Berkeley.

ZADEH, L.A., 1975c. Fuzzy logic and approximate reasoning. Synthese 30, 407-428.

ZADEH, L.A., 1976a. A fuzzy-algorithmic approach to the definition of complex or imprecise concepts. Int. J. Man-Machine Studies 8, 249-291.

ZADEH, L.A., 1976b. The linguistic approach and its application to decision analysis. In: Directions in large-scale systems. Y.C. Ho & S.K. Mitter (eds.), Plenum Press.

ZADEH, L.A., 1977a. Fuzzy sets and their application to pattern classification and clustering analysis. Classification and clustering, 251-299, Academic Press.

ZADEH, L.A., 1977b. Theory of fuzzy sets. Memo No. UCB/ERL M77/1 Electronics Research Lab, University of California, Berkeley. In: Encyclopedia of computer science and technology (J. Belzer, A. Holzman, A. Kant, eds.), Marcel Dekker, New York.

ZADEH, L.A., 1977c. Linguistic characterization of preference relations as a basis for choice in social systems. Memo No. UCB/ERL M77/24, Electr. Res. Lab, Univ. of Calif., Berkeley.

ZADEH, L.A., 1977d. A theory of approximate reasoning (AR). Memo No. UCB/ERL M77/58, Electr. Res. Lab, Univ. of Calif., Berkeley.

ZADEH, L.A., 1978a. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, 3-28.

ZADEH, L.A., 1978b. PRUF - a meaning representation language for natural languages. Int. J. Man-Machine Studies 10, 395-460.

ZADEH, L.A., 1979a. Fuzzy sets and information granularity. In: Gupta, M.M., et al. (1979a), 3-18.

ZADEH, L.A., 1979b. Approximate reasoning based on fuzzy logic. Memo No. UCB/ERL M79/32, Electr. Res. Lab, Univ. of Calif., Berkeley.

ZADEH, L.A., 1979c. Possibility theory and soft data analysis. Memo No. UCB/ERL M79/66, Electr. Res. Lab, Univ. of Calif., Berkeley.

ZADEH, L.A., 1979d. Precisiation of human communication via translation into PRUF. Memo No. UCB/ERL M79/73, Electr. Res. Lab, Univ. of Calif., Berkeley.

ZADEH, L.A., 1980a. Fuzzy sets vs. probability. Proc. IEEE 68, 3, 421.

ZADEH, L.A., 1980b. Inference in fuzzy logic. 10th Annual Symposium on Multiple-valued Logic, 124-131.

ZIMMER, A., 1980a. Eine Formalisierung mnestisch stabilisierter Bezugssysteme auf der Grundlage von Toleranzmengen. In: Thomas, A., Brackhane, R. (eds.), Denken - Urteilen - Handeln, Bern.

ZIMMER, A., 1980b. An application of interrogative logic to fuzzy concepts. 2nd Cognitive Science Conference, Yale.

ZIMMERMANN, H.J., 1979. Rational decision making in fuzzy environments. Course notes.