BEYOND PROTOTYPES AND FRAMES: THE TWO-TIERED CONCEPT REPRESENTATION

by

Ryszard S. Michalski

Reports of the Machine Learning and Inference Laboratory, MLI 91-6,
George Mason University, Sept. 26, 1991.

Beyond Prototypes and Frames:
The two-tiered-Concept Representation

Ryszard S. Michalski

P91-20
MLI91-6

2/16/92 update

Prepared for
Categories and Concepts: Theoretical Views and Inductive Data Analysis,
I. Van Mechelen, J. Hampton, R.S. Michalski and Peter Theuns (eds.)
Academic Press Ltd., 1992.

Beyond Prototypes and Frames: The Two-tiered Concept Representation

R.S. Michalski
Artificial Intelligence Center
George Mason University,
Fairfax, VA 22030

I. Introduction

Cognitive scientists have been, for years, searching for essential ingredients of intelligence. While this issue may not be satisfactorily resolved for quite some time, two abilities are clearly central to intelligent behavior. One is the ability to acquire knowledge or skill through experience, that is, the ability to learn. The second is the ability to apply the knowledge or skill possessed to solve new problems, that is, the ability to reason. The new problems may concern actual events in the real world, for example, when one has to react to a new external stimulus, or may be imaginary, for example, when one creates them for planning purposes.

A precondition for the above abilities is the capability to represent diverse forms of knowledge. As our knowledge is built of individual concepts, to represent knowledge one needs to represent concepts. Consequently, understanding how concepts are represented is a fundamental problem underlying all efforts in the quest to understand intelligence. Although cognitive scientists, psychologists, linguists, philosophers and artificial intelligence researchers have given considerable attention to this problem, no conclusive solution of it exists.

Other chapters of this book have reviewed some of the most widely known approaches to concept representation. These approaches include the classical view, probabilistic view, prototype view and the frame view (see, e.g., chapters by Barsalou and Hale, Hampton, and Murphy). Although these views explain many aspects of concept representation, they do not provide a totally satisfactory and conclusive solution.

The aim of this chapter is to explore a new approach that overcomes some of the weaknesses of these well-known representations. The new approach, called the *two-tiered (TT) concept representation,* considers human concepts not as fixed, well-defined structures, but as flexible, context-modifiable and background-knowledge-dependent units of our knowledge. To represent concepts so understood, the TT view recommends splitting the total concept representation into two components ("tiers"), one that explicitly defines the most stable and common concept properties, and the second that implicitly defines allowable concept modifications, exceptional cases and context-dependency. The latter component consists of procedures for "flexible" matching concept descriptions with instances, and a body of inference rules for characterizing concept variability in different contexts and situations.

As a background for the new approach, the following section outlines our notion of concepts and categories. Later sections provide details about the two-tiered representation, relate it to other views, and finally discuss its application to data analysis.

2. The Meaning of Concepts and Categories

While everybody agrees that concepts and categories are fundamental building blocks for our knowledge and thinking, there seem to be no universal agreement as to their meaning and the definition. There are differences among the views of individual cognitive psychologists on these issues, and differences between the views of cognitive psychologists and researchers on formal theories of learning and artificial intelligence. These issues, however, are important to the proposed two-tiered concept representation, and therefore, we will attempt to clarify our view on them here. We will present a view that appears to us most theoretically satisfying, although it is somewhat different from the views often expressed in cognitive science literature. The main advantages of the proposed view is that it allows us to more rigorously define the meaning of concepts and categories, to distinguish "concepts" from "concept representations," and to clarify the meaning of concept learning.

Let us start with views expressed recently by some of the most prominent cognitive scientists. Barsalou (1990) states: "Whereas most theorists would probably reserve 'concept' for categorization rules, I will also use 'concept' for people's conceptualization of categories." Medin (1989) expressed the view that "a concept is an idea that includes all that is characteristically associated with it." He also suggests that "a category is a partitioning or a class to which some assertion or set of assertions might apply." Smith (1989) views a concept as a "mental representation of a class or individuals and deals with what is being represented and how that information is typically used during the categorization."

The above characterizations do not seem to be all equivalent, and from a viewpoint of a formalist raise certain questions. For example, if concepts are "people's conceptualization" (or "mental representation") of categories, then this implies that there are as many different concepts for the same category as there are different cognitive representations of that category. Everyone has an individual representation of a given category, since people's knowledge about concepts changes, as any new fact adds to the representation. If a concept "square" is the same thing as a cognitive representation of the class of squares, then we have so many concepts of "square" as there are minds (or machines) that hold some representation of "square."

For similar reasons, one could argue with the view that identifies concepts with a concept description or a classification rule. For any given concept one can usually think of many different concept descriptions or classification rules that specify how to distinguish between concept instances and non-instances. To see this point, suppose that a concept was defined as a certain classification rule, expressed in terms of specific properties and/or relations linked by specific operators. If one would replace this rule by a syntactically different one but logically equivalent (i.e., one that produces exactly the same classification), then we would have a contradiction: since the initial and the newly created rule are different from each other, then they would constitute different concepts, yet—since they generate exactly the same classification of instances—they would have to be viewed as the same concept. For example, there can be many different medical procedures ("classification rules") for distinguishing a specific type of cancer from other types of cancer. Although these medical procedures are different, the concept of cancer is all the same. In short, identifying concepts with their classification rules may lead to a contradiction.

Another concern is that the above views seem to disagree with the view of concepts often expressed in artificial intelligence and formal theories of concept learning (e.g., Sowa, 1984; Fulk and Case, 1990). This view is that concepts are certain sets of entities (the entities are called concept instances). A *concept description* is an expression that characterizes this set. Thus, a human or machine can learn a concept, if it can determine a representation that can be effectively used to recognize concept instances. Thus, a concept is not the same as its representation. The formal status of concepts as sets seems to be also confirmed by our linguistic expressions that uses them. It is common to say "This example belongs to a concept" or "The instances of the concept are..." Such expressions seem to be more consistent with the usage of "concept" as a set than as an "idea," "conceptualization," etc. The view of concepts as sets has not been common in cognitive science, and therefore may cause difficulties for cognitive psychologists. It allows one, however, to more rigorously define the meaning of concept representation, the similarity between concepts, and to describe the two-tiered view, and therefore we will use it in this chapter.

We will elaborate this view by stating that *concepts are discrete units that stand for sets of entities grouped together because of some reason*. Since concepts are in one-to-one correspondence to sets, they can be formally treated as sets. The set constituting a concept may consist of just one individual (e.g., the sun in our solar system; since it is a dynamic system changing in time, then, strictly speaking, it can be viewed also as a set of instances corresponding to different time moments), a countable number of individuals (e.g., the set of students taking Cognitive Science 101, or the set of birds called bluejay), or an uncountable number of individuals (e.g., the concept of "real number"). There can be many different reasons for grouping entities to a concept. For example, a reason may be that entities have the same function or purpose, have similar appearance, have similar physical, spatial, or temporal characteristics, share the same abstract property(ies), have the same relation to other entities, or because of some combination of these and/or other factors. Typically, concepts are sets with a name, and are referred to through a name. There can be, however, concepts without a name. For example, entities such as crowds, tanks, speeches, white-red-and-blue flag, and broken monuments of formerly admired communist leaders may constitute an unnamed concept of "things that come to mind when thinking about the August 1991 revolution in the (former) USSR." Different people may associate different objects with this revolution, therefore this concept is a personal construct.

The instances of a concept represent physical or abstract objects. These objects are perceived by (human or machine) through their properties. These properties can be values of certain attributes, relations among objects components, relations among properties of the components, etc. From now one, attributes or relations used to characterize a set will be called *descriptors*. Thus, every concept instance can be viewed theoretically as a point in a *concept description space* spanned over the descriptors through which the objects are perceived. Concept descritpions then map entities into sets in certain description spaces. In the simplest and the most common case, a concept description space is spanned over only attributes, i.e., zero or one argument functions. In a more general sense, a description space can be spanned over any type of descriptors.

Viewing concepts as subsets of a description space, one can characterize a concept instance as more or less typical, depending on the relationship between the instance and other instances. For example, one simple measure of typicality would be the "distance" between an instance to the "center" of the set in the description space. The similarity between concepts can be defined by the "closeness" between corresponding sets in the concept description space. This definition implies that the similarity between concepts is crucially dependent on what attributes/variables are spanning the concept description space. Without defining the description space, the similarity between two concepts cannot be determined. A description space is often defined only implicitly, e.g., through the properties that humans measure through their senses. For example, an orange and a tangerine are usually viewed as similar fruits. The implied assumption is that they are similar in terms of the properties measured by our unaided senses. However, one could think of a description space in which these two entities would be very dissimilar, e.g., in the space spanned over specially selected atomic level attributes. In view of these difficulties, the presented "two-tiered" approach developes a new view of concept typicality.

To recognize an instance of a concept is to identify an a priori known concept (set) to which the given instance belongs. To do so, one always needs to use some knowledge (explicit or implicit) about the concept and about the instance. The knowledge about the concept is usually expressed in the form of a "concept description," or a "classification rule." Such a description or rule is supposed to capture the reason underlying the concept.

The difficulty in understanding the meaning of a concept is that the reason for creating a concept may be incorporated in a body of knowledge that is encoded in the structure of an agent's (concept bearer's) mind as a result of some learning or creative process. Such a body of knowledge may, or may not, be precisely translatable into an effective classification rule. It may be externalized only by the classifications one makes, e.g., one may not be able to tell why a given painting is beautiful, but may be quite confident in this evaluation. In sum, the reason (or principle) underlying a concept is not necessarily expressible linguistically, and a concept (a set) should be distinguished from a classification rule for that concept. One basic problem is then what is the relationship between a concept and a classification (or categorization) rule used for its recognition. Another problem is what is the relationship between concepts and categories. This section proposes a new view on these issues, somewhat different from this one currently used in cognitive science literature.

A classification rule is an expression in some language (natural or formal) that constitutes an effective procedure for telling which instances belong to a concept and which do not. Such a procedure uses various attributes and/or relations (i.e., descriptors) that are measurable directly or indirectly. As mentioned earlier, there may be many different classification rules for recognizing instances of a given concept. For many concepts, however, only approximate classification rules can be constructed. As mentioned above, this is because the reason for grouping entities to a concept may involve abstract notions expressed in terms of other abstract notions, and these, in turn, may depend on the individual experience of the agent holding the concept.

Consider, e.g., the concept of "friendship" or "revolution." The meaning of such concepts changes with an agent's experience, and thus these concepts are living personal constructs. They are open-ended sets of behaviors that this agent would classify as "friendship" or "revolution." Such concepts would usually be manifested by instances the agent may generate rather then by any formal descriptions. To define a precise classification rule for such concepts, even only for a specific agent and for a specific time period, one would have to inspect all different concept instances that the agent could possibly generate. But this may be impossible because the agent may not be able to classify some instances, and/or because the number of instances may be too large. Consequently, some human concepts cannot, in principle, be defined precisely. Consequently, there cannot be precise classification rules for them. For such concepts there can only be approximate classification rules derived from abstract linguistic or other descriptions, or inductively learned from subsets of instances generated by a reasoning agent.

Concepts that were created by defining an a priori classification rule (e.g., prime numbers), or for which an effective classification rule has been developed (e.g., mammals) are called *categories*. Categories are thus special kinds of concepts. An effective classification procedure for a category must be expressed in terms that are either directly observable or measurable, or themselves have an effective recognition procedure. From this viewpoint, concepts such as "peace" or "freedom" are not categories, because no effective recognition procedures for them have been developed. On the other hand, concepts such as "students who received A in my class" or "GMU employees earning less than 35k in 1992" are categories.

Some concepts are sets that are relatively "stable," that is, more or less permanently useful in describing our world, or our actions in it. Therefore, we develop effective classification rules for them, and they become categories. For example, animal species, plants, diseases, astronomical objects, etc., are such concepts. We call them *natural categories*. Other concepts represent sets that may have only temporary utility, like, for example, the concept of "things to carry out of a burning house" (Barsalou, 1987), or "pins in my drawer." Such concepts are not natural categories, although they are or can be categories. As stated earlier, for a given concept, there may be many different classification rules. Some of them may be simple but only approximate (e.g., those defining the animal species based on the animal's physical appearance), and some may be complex but precise (e.g., those defining the animal species based on animal's genetic code). Depending on the need, different types of rules are employed.

Concepts that are sets with *precise boundaries* that do not change with the context or situation in which the concepts are used, are called *crisp* (or *classical*). Crisp concepts occur primarily in science, where—whenever it is useful—they are given a name (e.g., the concepts such as "even number" or "vertebrates"). These concepts remain crisp usually and only if used within the scope of a given scientific discipline, e.g., "triangle" or "rectangle" in geometry, a "group" in mathematics, "water" in chemistry. Outside of this discipline, however, they may represent different sets of entities. For example, a triangle may mean an arrangement of streets or cities; a group may mean a set of people, etc. There are relatively few crisp concepts outside of science. For example, kin relationships, such as "father" or "grandmother," can be viewed as crisp. But even these concepts are often used in a flexible manner. For example, the founder of a scientific discipline may be called a "father of that discipline" ("Lukasiewicz is a father of multi-valued logic"). Here, the original concept is extended to a larger set through an analogical reasoning.

Concepts that intrinsically do not have precise boundaries, either because they are open-ended and/or because they are context-dependent, are called *flexible*. The boundaries of flexible concepts are imprecise and may dynamically evolve with time, and/or change with the context in which they are used. For example, the concept "computer" (i.e., the set of all entities called "computers") does not have, in principle, a well-defined and context-independent boundary. The entities called computers have been rapidly evolving in time, and it would be unwise to a priori limit the meaning of this concept.

In order to make the above notions more clear, one needs to explain what is meant by context. By *context* of a concept we mean a set of concepts relevant to the intended meaning of the given concept. What concepts are relevant to the intended meaning is specified by the agent's background knowledge (sometimes also called agent's theory). A context can be explicitly specified by evoking the name of a more general concept that includes the concept discussed, or by providing a description characterizing it.

In different contexts, a flexible concept may stand for similar or quite different sets of entities. For example, the concept "printer" would have a different meaning in the context of "old-fashioned office equipment" than in the context of "modern office equipment." As another example, consider the concept "key." In the context of doors, rooms, houses, buildings, etc., "key" would mean a set of entities that serve for opening door locks. In the context of "travel," "suitcases," "briefcases," etc., the concept so named would stand for a somewhat different set of entities, related only at an abstract functionality level to the entities in the previous set. Finally, in the context of "problems," "solutions," "methods," etc., the concept "key" would stand for a quite different set, related only by analogy to the previous sets, and could be considered as being a different concept. This example suggests that sets associated with a given name of a flexible concept consist of elements that may be bounded to each other by different relationships, such as identity, surface similarity, analogy, functionality, or just by a certain relationship to other concepts. The context determines which sets are referred to in a given discourse.

Crisp concepts, by the very fact that thay have precise boundaries, can be represented by logical-style descriptions that evaluate to "true" or "false." Such descriptions may be in the form of disjunctive normal expressions, sets of rules, decision trees, grammars, exemplars, prototypes, etc. These descriptions can always be transformed, by the use of intermediate terms or relations, to a single conjunction of single conditions. Consequently, crisp concepts can be defined by a set of jointly sufficient and individually necessary conditions. Thus, the old controversy associated with the classical view of concept (e.g., Smith and Medin, 1981; Smith, 1989; Sutcliffe, 1992 - this book) can be resolved through the idea of crisp concepts. The classical view simply applies only to crisp concepts. If concepts have precise boundaries, then they can be equally precisely described by a logical description, or by a "prototype-based" description (but not with the same simplicity). Thus, the prototype view of concepts makes a real difference only if one does not assume that concepts have precise boundaries.

Majority of human concepts are not crisp, but flexible. For such concepts, it may not be possible to construct precise logic-style classification rules. The reason is that boundaries of such concepts—as mentioned earlier—are intrinsically imprecise and/or dependent on the current and future contexts in which the concept may be used. Thus, only approximate classification rules are attainable for flexible concepts.

It should be noted, however, that instances of concepts, crisp or flexible, may occur with different frequencies. One way to characterize the "typicality" of a concept instance is associated with such a frequency of occurrence (the "frequency-based" typicality). For example, there could be typical or atypical geometrical triangles, in the sense that they occur frequently or infrequently. Since a triangle in geometry is a crisp concept (with sharp boundaries), every instance is a "prototype" triangle. In other words, all geometrical figures satisfying the definition of a geometrical triangle are "equally good" triangles, although they can occur with different frequency.

A different situation is with flexible concepts. Since such concepts do not have precisely defined boundaries, or different boundaries apply in different contexts, to develop a classification rule for them one needs to rely on the distributions of instances over a space of the observable descriptors. If there are some instances that belong to the concept frequently and in many situations and many contexts, then they can be viewed as "highly typical." Here, their typicality can be measured by the degree of similarity (defined in terms of the chosen observable descriptors) between an instance and an "average" instance in the space spanned over these descriptors. Which descriptors are chosen to characterize an instance depends on how relevant they are to the original principle underlying the concept. A judgment of relevance of different descriptors depends on the knowledge or experience that an agent has with the concept.

The prototype view of concepts (see, e.g., Hampton's chapter in this book) concerns flexible concepts whose instances have a "well-behaved" distribution (just one or few modalities) in the space spanned over the chosen descriptors. However, for many concepts, e.g., chair, game, house, music, mechanism, etc., it is very difficult to determine a set of measurable descriptors that would produce a space in which all their instances would be distributed in such a "well-behaved" fashion. The key problem in representing such concepts lies in the difficulty of describing all their possible manifestations and context-dependency in terms of measurable properties.

It is a remarkable challenge to understand how people can use and communicate with such concepts without having an a priori agreed upon classification rules or well-defined descriptions. If specifically asked, different people usually produce different descriptions of such concepts. The proposed below two-tiered representation aims at explaning this phenomenon. It can handle both crisp and flexible concepts, but its primary aim is to represent the latter ones.

3. The Two-tiered Concept Representation

The fundamental assumption underlying the two-tiered concept representation is that total classificatory information about a concept resides not only in its explicit representation, but also in the inference and matching procedures applicable to it. Through the application of these inference and matching procedures the meaning of a concept can be extended, modified, or adapted to various contexts. The two-tiered representation thus constitutes a significant departure from the previous approaches, which attempted to represent the complete meaning of a concept within one explicit knowledge structure (a list of properties, a logical desription, a set of prototypes). Nevertheless, as explained later, the TT view builds upon and relates in various ways to other views.

The TT view was originally proposed in Michalski (1986), and subsequently applied to various problems in concept learning (e.g., Michalski, 1990; Bergadano et al., 1988b, c; 1991). The original formulation of this approach was based on an observation that although most human concepts lack precise definition when used outside of a specific context, they acquire a precise meaning when used in a combination with other concepts in a specific context. Consider, for example, the statement: "This blonde woman near the center of the room is a student at GMU." If there is only one woman with hair of lighter color than any other woman near the center of the room, the statement above precisely defines the person of interest (although the concepts "blonde," "near," and "room" are imprecise by themselves, outside of a specific context). Thus, statements made of imprecise concepts can convey precise meaning. In other words, the fact that individual concepts are flexible (i.e., imprecise and context-dependent) does not prevent us from communicating precise meaning.

The above example illustrates the basic supposition for the TT approach (Michalski, 1986) that the imprecision of human concepts stems not from an undesirable vagueness of our concept definitions, but rather from the universal need for cognitive economy. By allowing individual concepts to be imprecise and context-modifiable, the expressive power of concepts is greatly enhanced. The latter means that one can employ fewer concepts for expressing a greater set of meanings. This leads to a simplification of descriptions of our immensely complex universe. Various experiments reported here and elsewhere (e.g., Michalski, 1990; Bergadano et al., 1988b, c; 1991) have confirmed this idea in a microworld to which it was applied.

The two-tiered approach assumes that (flexible) concepts have a certain central tendency and basic usage, which should be described explicitly, as the "first approximation" of the concept meaning. The complete concept meaning, however, may vary significantly beyond the basic (or typical) usage, especially in different contexts. The TT view postulates that this variability and context-dependency is best represented implicitly, i.e., results from an application of "flexible" matching methods and context- and background knowledge-dependent rules of inference. Accordingly, a two-tiered concept representation consists of two components (tiers):

• the BCR, the *base concept representation,* that represents the basic concept meaning explicitly. This basic meaning may be characterized by specifying the typical function of the concept instances, their usual physical properties and appearance, their structure and the components, their role in a system of concepts by indicating ancestor and descendent concepts, etc. (In the specific method used in the experiments mentioned later, the BCR was expressed by a logical DNF description, or an equivalent set of rules).

• the ICI, the *inferential concept interpretation,* that consists of inference rules and matching procedures, that modify the basic meaning accordingly to any given situation or context. In matching an instance with the concept representation, these inference rules and matching procedures involve the agent's background knowledge (BK) and the context of discourse. The ICI may involve any type of inference—deductive, analogical or inductive.

Thus, the basic or typical concept meaning is expressed by the BCR, and imprecision, exceptions and the context-dependency are handled by ICI. The ICI captures the allowable transformations applicable to a concept in different contexts. Entities that satisfy the base concept representation are called typical, and those that do not are characterized by a degree of typicality. Factors affecting the degree of typicality are the type and the amount of transformation needed for matching the entity with the BCR, and the frequency of the occurrence of an entity. Highly atypical entities are called *exceptions*, and require ICI rules for identification (Bergadano et al., 1991).

Figure 1 illustrates the relationship between the BCR and the ICI in a TT concept representation. It shows that the ICI can, in general, extend the concept meaning beyond the BCR in one area of the description space, and reduce the meaning in another area.

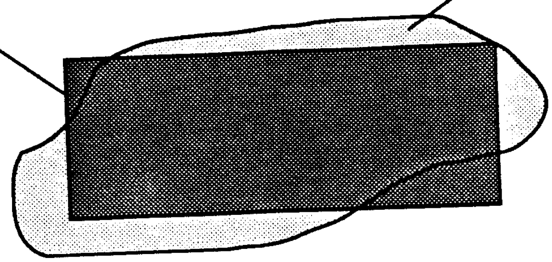

Figure 1. An illustration of the relationship between the Base Concept Representation (BCR) and the Inferential Concept Interpretation (ICI).

Note that BCR is not a "prototype," but a characteristic description of a class of entities that are viewed as typical members of a given concept. Such a description may be a disjunction of conjunctions (i.e., a DNF expression). A "prototype" instance should satisfy the BCR, but there could be a number of distinct "prototype" instances that also satisfy it. For example, a large numer of chairs can be viewed as typical of the concept of "chair." Also, a chair, a desk, or a bookshelf can be viewed as typical examples of the concept "office furniture." The BCR can be expressed at different levels of abstraction. For example, a BCR of the concept "table" can include statements stating their typical function, typical physical properties, the typical material they are made of, etc. In different situations different components of BCR can be used for concept recognition. A method of "dynamic recognition" that uses various subsets of properties for recognizing a concept in different situations (a situation is defined by the set of properties that have been observed or can be observed) have been described in Michalski (1989).

Here is an example of a TT representation of the concept "chair":

BCR: *Function*:: To seat one person.

Structure:: A seat supported by legs and has an attached backrest from one side.

Physical properties:: Made of wood or plastic. The seat is flat and is located about 14-18 inches above the ground. There are usually four legs which are located in the vertices of an imaginary square, etc.

(BCR may also include pictures or 3D models of typical chairs)

Immediate ancestor (parent) concept:: Furniture

Immediate descendents (children):: straight chair, armchair, chaise lounge, folding chair, highchair, lawn chair, rocking chair, barber chair, throne, wheelchair.

ICI: *Possible variations of the properties in BCR*: The number of legs can vary from one to four. The legs may be replaced by any support. The shape and the material of the seat, the legs and the backrest are irrelevant, as long as the function is preserved. The backrest may be very small or missing; the seat may be closer to the ground or much higher above the ground than specified in BCR, etc.

Situation or Context-dependent variations:

Situation:: If no typical chair is available ---> any object on which a person can sit can be called chair.

Context = museum exhibit ---> an object mathing the visual characteristics of a chair is not used for seating persons (any more)

Context = {toys, doll's house, dolls, children playing,...} ---> the size can be much smaller than stated in BCR. The chair does not serve for seating persons, but correspondingly small dolls.

Context = organizational structure, committee, positions of responsibility, University department, etc. ---> chair means a person being in charge.

To recognize a concept instance, the instance is matched against the BCR of candidate concepts. This process first attempts to find a BCR that directly matches the instance. A match is direct if the BCR's conditions are precisely satisfied by the instance. If there is no direct match, or there are several direct matches, then the process employs the ICI rules of inference and matching procedures. If there is no direct match with any BCR, the system searches for a deductive match, i.e., tries to determine if the instance is a logical consequence of some BCR. For example, if there is a BCR condition "X is a very good student" and ICI has a rule that "a student who has a high grade average and gets along well with most of the peers and teachers then that student is a very good student," then an instance "John has high grade average and gets along with almost everyone" deductively matches the BCR condition. If there is no deductive match or there is more than one direct match, a *plausible match* procedure is employed. Such a procedure determines the degree to which an instance matches different concept candidates by using matching procedures and rules contained in the ICI. The best matching description defines the concept.

Most of the known concept representations, such as semantic network, decision tree, and frame representation usually employ a direct match. Plausibility matching employs usually a method for determining some degree of similarity or partial match. In a general case, a plausibility matching may employ any type of plausible inference (e.g., approximate deduction, analogy or induction; Collins and Michalski, 1989). Illustrative examples of plausible matching are given in the next section. Performing plausible inferences may involve concept meta-knowledge, e.g., the importance of concepts and their frequencies of occurrence, the relation to other concepts and other relevant domain knowledge. Such meta-knowledge is assumed to be a part of the second tier.

As mentioned earlier, the ICI includes the background knowledge relevant to the proper interpretation of the concept. Since such background knowledge (called by some authors "theory," e.g., see chapter by Murphy) grows with a person's experience, the same concept may have a different meaning for different people. For example, the concept of a computer is quite different for a novice than for a computer expert who witnessed the evolution of computers over many years. Thus, the two-tiered view provides a simple model for explaining the observable personal differences in concept meaning. In sum, by dividing the concept representation into an explicit and implicit part, the TT view combines elements of the classical representation, prototypical representation, and other elements not present in other representations.

In the cognitively-oriented form of the TT representation described here, the distribution of the concept meaning between BCR and ICI is relatively stable. The BCR expresses the general unifying idea, the typical function, and/or measurable properties implied by or correlated with this idea or function. Such a BCR can be viewed as representing the "first approximation of the concept." The ICI, in this case, defines the matching procedures and inference rules for handling less typical instances and context-dependency. In a general approach, the TT theory allows any "distribution" of meaning between BCR and ICI (Michalski, 1990).

An advantage of distributing the concept meaning between the BCR and the ICI is that it permits a learner to flexibly modify or extend the concept meaning by varying matching procedures and inference rules, and/or by changing the context of discourse. The concept meaning can thus be changed without having to alter the base concept representation. As mentioned earlier, by evaluating the type and the amount of inference involved in matching the BCR with an instance, one may produce a qualitative or quantitative estimation of the degree of typicality.

The ability to produce a measure of the degree of match indicates one principal difference between this approach and the fuzzy set approach (e.g., Zadeh, 1965, 1976). In the fuzzy set approach, a set defining a concept is associated with a membership function, which needs to be *defined* to the learner by a person. The influence of the context is hidden in the definition of this membership function. In the proposed approach, a concept is associated with interpretation procedures and context-dependent rules, which implicitly define the membership of an instance in a concept. These rules and procedures can be used to *compute* the membership function in different contexts.

4. Examples Justifying the Two-tiered Approach

There are many examples that show that human concepts are not characterized solely by explicitly stated properties, relations, or prototypes. A full description of most concepts involves general and domain-specific rules and background knowledge, in order to capture the concepts' variations, and context-dependency. These rules tell us what transformations of instances preserve or not their concept membership, which are potentially feasible, and what transformations move them outside of a concept. For example, repainting a chair preserves its membership in the category of chairs, but shredding it does not. Consequently, to learn a concept one needs to learn not only the base concept representation, but also inferential concept interpretation. Let us consider a few examples illustrating and confirming the ideas of the two-tiered representation.

Example 1. A pizza or a coin?

This example is based on recent experiments by Rips (1989), which involved concepts that can be characterized, using the TT terminology, as having significantly different ICIs, and somewhat comparable values of one attribute in their BCRs. Specifically, a subject was presented with a partial description of an object that expressed only the value of one of its attributes (e.g., its diameter). Based on this description, the subject was supposed to classify the object to one of the two categories. The attribute value given in the description was somewhere between the subject's average values of this attribute for the two categories. For example, an object was described as "three inches in diameter," and the categories of choice were pizzas and quarters. The results of experiments were that the subjects consistently classified such an object as likely to be a pizza, although the specified size was closer to the quarter than to a typical pizza.

The two-tiered model explains this behavior simply. The subjects do not only hold the description of the typical values of objects in each class (the BCR), but also have an understanding of the possible variability of the objects in each category (the ICI). This understanding comes either from a generalization of their observations, or from their relevant background knowledge, e.g., about the process that produces objects of the two categories. Consequently, they could admit an object to the category of pizzas, although unusually small for a pizza, but not to the category of quarters, since quarters do not have a variable diameter.

Example 2. Concept of sugar maple

Our prototypical image of a sugar maple is that it is a tree with three- to five-lobbed leaves that have V-shaped clefts. Some of us may also remember that the teeth on the leaves are coarser than those of red maple, that slender twigs turn brown, and that the buds are brown and sharp-pointed. As a tree, of course, a maple has a root, trunk and branches.

Suppose that while strolling on a nice winter day someone tells us that a particular tree is a sugar maple. A simple introspection tells us that the fact that the tree does not have leaves would not strike us as a contradiction of what we know about sugar maples. Yet, clearly, the presence of leaves of a particular shape is deeply embedded in our typical image of a maple tree. The two-tiered theory explains this phenomenon simply: the inferential concept interpretation associated with the general concept of deciduous trees evokes a rule "in winter deciduous trees lose leaves." Since a maple is a deciduous tree, the rule would apply to the maple tree. The result of this inference would override the stored standard information about maple trees, and the inconsistency would be resolved. Matching an instance with the concept requires in this case deductive reasoning from the knowledge associated with a more general concept.

Example 3. A coffee pot used as a bird feeder

This example is taken from a study of children concepts performed by Keil (1989). In this study, children were shown objects representing some known concepts. These objects then have undergone a sequence of transformations that produced an object resembling an instance of a different concept. In one experiment, the transformations were realistic, i.e., corresponded to changes that can be obtained by some simple operations. In the other experiments, the transformations were purely imaginary (they did not correspond to changes that would be allowed by the ICI).

For example, in one case the initial object was a coffee pot. Subsequently, the pot was punched full of holes, and its spout was removed. It was then filled with bird feed and hung from a branch of a tree. In another case, the initial object (presented as a picture) was a raccoon. It was then presented as having been transformed to black with a white tail, and able to produce a super smelly stuff from beneath its tail, like a skunk. The children were then asked if the end result of these transformations was, in the first case, a bird feeder or a coffee pot, or in the second case, a raccoon or a skunk. The experiments thus can be interpreted as testing the children's understanding of the feasibility of various transformations of instances of different concepts.

The results of the first experiment were that all children (5 to 9 years old) indicated that it was a coffee pot that was turned into a bird feeder. In the second experiment, however, all children have experienced a difficulty in accepting the change of the raccoon into a skunk. These experiments indicate that children knew not only physical characteristics of the objects in the involved categories (here, the coffee pots and raccoons), but had also the knowledge of transformations that are applicable to objects in these categories.

Example 4. Concept of a triangle

Let us consider the concept of "triangle." In geometry, a triangle is a plane figure consisting of three non-colinear points connected by straight lines. Using the notation of *annotated predicate calculus* (APC), which is like predicate calculus, but employs more compact forms of some logical expressions (Michalski, 1983), one can write:

Triangle(T, P1, P2, P3) $\Leftarrow$ Consists(T, P1 & P2 & P3) & Type(P1 & P2 & P3, point)& Connected_by_straight-line(P2 & P1,P3 & P2,P1 & P3) & Non-colinear(P1, P2, P3)

In the above expression, symbol "&" is used in two related meanings: first, to denote an ordinary *(external)* conjunction connecting predicates, and second, to denote an *internal conjunction,* i.e., a conjunction of terms, which is treated as a *compound* argument of a predicate. For example, the predicate "Type(P1 & P2 & P3, point) states that P1 and P2 and P3 are points, thus is equivalent to a conjunction of three predicates: Type(P1, point) & Type(P2, point) & Type(P3, point).

Suppose that someone tells us that the tall towers in his/her hometown form a big triangle. Obviously, the meaning of the triangle in this statement differs from that in the formal geometrical description. To match the statement with the concept of triangle, the following assumptions and transformations need to be made:

a. In the context of describing a configuration of physical objects such as towers, the individual objects play the role of points. Thus, the statement implies that towers correspond to points, and that there are three towers in the town. The matching operation involves drawing an analogy between the abstract points and the towers, which can be characterized as consisting of one step of generalization (GEN):

Point ----GEN----> Object

and one step of specialization (SPEC):

Object ----SPEC----> Tower

b. In the context of towers, the presence of a "straight line" is imaginary, i.e., there is no physical connection, but one could imagine a straight line between the objects (towers). The condition "ConnectedBy" is then satisfied in such an abstract sense. This is an operation of abstraction. Thus, matching the statement about a triangular arrangement of towers with the definition of a triangle involves here a generalization, specialization and abstraction. [A reader interested in an analysis of the differences between generalization and abstraction may consult Michalski (1991).]

The examples above show that relating a concept instance to a concept representation is not just a straightforward comparison of attribute values in an instance with those in the concept representation, as done in various mechanized decision processes. They show that such a process may involve different forms of inference, as postulated in the two-tiered representation.

5. Learning Two-tiered Concept Representations

The essence of the two-tiered representation is a recognition that a concept may have explicit and implicit properties. The explicit properties are represented by logical-style expressions, and the implicit ones by a set of rules and/or matching procedures. Consequently, to acquire a TT concept representation, one needs to learn both, the explicit (BCR), and the implicit (ICI) components of a TT representation.

In general, each concept has a different BCR that has to be learned. On the other hand, the ICI is often the same or very similar for various classes of concepts. Consequently, it can be learned only once for a given class of concepts. All concepts of this class, and/or descendant concepts in the type hierarchy will share the same ICI. For example, all diseases in the class of liver diseases may share the same flexible method for matching the BCR descriptions with symptoms. By sharing the ICI among concepts, a significant economy of the concept representation can be achieved.

The two-tiered approach does not impose restrictions on what type of language is used to represent BCR and ICI parts of a concept. Any type of knowledge representation formalism could potentially be employed. It is assumed that the language employed should depend on the goal of the concept representation. The goal implies the degree of abstraction with which the concept needs to be represented, and this sets the criteria for the representation language. For example, in a diagnostic situation, it may be sufficient to represent a disease at a high level of abstraction, e.g., by an attributional description in the form of propositional calculus rules. In a medical research situation, the representation may have to be much more complex, and involve not only structural representation (e.g., using a predicate calculus language), but also images and drawings.

In most of the experiments done so far with the TT representation, the representation language was an attributional calculus, called VL_1. In this calculus, descriptions are expressed in the form of decision rules. A decision rule links a logical product of elementary conditions with a decision (e.g., a class membership). The elementary conditions relate a multiple-valued attribute to a value or a set of values. For example, elementary conditions may be [sex = male], or [color = blue or red] or [weight > 2 kg]. [See, e.g., (Michalski, 1974; Bergadano et al., 1991)] for details on this representation.). In experiments in learning structural descrtiptions, an extension of predicate calculus (called *annotated predicate calculus*) was used (Bergadano et al., 1988b).

The rest of this section outlines an implemented method for learning two-tiered concept representations from concept examples using the above mentioned attributional calculus It also presents results from comparing the method with several other methods that use different concept representations. Full details of the method and the experiments performed are described in Bergadano et al. (1988b, c; 1991).

In the method, the *Base Concept Representation* (BCR) is created in two phases. In Phase I, a complete and consistent concept description is induced from the supplied examples. This is done by an inductive learning program AQ-15 (Michalski et al., 1986). Such a description explains (covers) all positive examples of the concept being learned, and none of its counter-examples (negative examples). The description is in the form of rules that link single conjunctive conditions with a concept name.

In Phase II, the obtained description is optimized according to a *description quality criterion*. The criterion prefers the descriptions that are computationally simple and at the same time have a high predictive accuracy, i.e., correctly classify new concept instances. The optimization process involves making various changes in the original description, such as removing a rule, removing a condition, or modifying a condition in a rule. These changes correspond to either generalization or specialization of the original rules.

For example, rule removal is a specialization operation, and condition removal is a generalization operation. A condition modification can be either a specialization or generalization operation depending on the type of modification. The rules obtained as a result of the optimization process may be significantly simpler than the original rules.

The obtained rules, however, may not directly match all the positive examples, i.e., they may constitute an *incomplete* concept description, and/or may directly match some negative examples, i.e., may constitute an inconsistent concept description. It turns out, however, that the incompleteness and/or inconsistency is often well compensated by the inferential concept interpretation (ICI), so that the resulting rules may actually have a higher predictive accuracy than the original rules.

Depending on the type of description learned, the ICI may consist of just a procedure for *flexible matching*, or the procedure plus a set of rules that capture the concept exceptions. In the latter case, the rules are acquired through an interaction with a teacher.

The method has been implemented in POSEIDON (AQ16), and experimentally tested on two real-world problems: learning the concept of an acceptable union contract, and learning voting patterns of Republicans and Democrats in the U.S. Congress. For comparison, a few other learning methods were also applied to the same problems. These methods included simple variants of exemplar-based learning, an ID3-type decision tree learning (ASSISTANT program; Chestnik et al., 1987) a method that utilizes complete and consistent descriptions, and a method that uses a very simple BCR (the "Top rule" description) and flexible matching.

To give the reader some sense of the results obtained, Table 1 presents a summary of the results from testing these methods (based on Bergadano et al., 1988b, c; 1991).

	Labor Contract	Congress Voting
Simple exemplar-based method		
Predictive accuracy (%Correct)		
1-nearest neighbor	77%	86%
3-nearest neighbor	83%	84%
5-nearest neighbor	80%	84%
Complexity (#Rules / #Conds)	27 / 432	51 / 969
Pruned decision tree (ASSISTANT + PRUNING)		
Performance (%Correct)	86%	86%
Complexity (#Leaves / #Nodes)	29 / 53	19 / 28
Complete and consistent description (AQ15 without rule truncation)		
Performance (%Correct / %Unrecognized)	80% / 3%	86% / 0%
Complexity (#Rules / #Conds)	11 / 29	10 / 32
"Top rule" description (AQ15 with rule truncation)		
Performance (%Correct)	83%	85%
Complexity (#Rules / #Conds)	2 / 6	2 / 6
Optimized two-tiered description (AQ16 or POSEIDON with rule truncation and condition optimization)		
Performance (%Correct)	90%	92%
Complexity (#Rules / #Conds)	9 / 12	10 / 21

Table 1. Summary of the results for testing descriptions learned by different methods.

The "Simple exemplar-based method" illustrates the exemplar view of concepts, i.e., represents concepts by sets of their examples. An unidentified concept instance is recognized by determining examples that are most "similar" or "close" to it. In the "k-nearest" procedure, k most similar examples are determined, and the concept associated with the majority of them is assigned to the instance (in the experiments reported, k was 1, 3 and 5).

The predictive accuracy of each description expresses the percent of classification of testing examples that are evaluated as correct by an expert. The complexity of the description was measured by the number of rules (#Rules) in it, and the total number of elementary conditions in all the rules (#Conds).

The "Complete and consistent" concept description is a disjunctive normal expression (equivalent to a set of rules) that describes all positive examples, and none of the negative examples. The "Top rule" description represents a very simple variant of two-tiered representation. The BCR is just a single rule that covers the largest number of examples among all the rules in the consistent and complete description. Such a rule can be viewed as representing the most typical concept examples. The ICI is a procedure for *flexible matching* that determines a similarity between an instance and the BCR of candidate concepts. The concept associated with the most similar BCR is assigned to the unidentified instance. It is interesting to notice although this form of two-tiered representation is extremely simple, it still has a relatively high predictive accuracy.

The "Optimized two-tiered description" is a more advanced variant of the TT representation. The BCR is a set of rules generated by removing some rules and/or conditions from the consistent and complete description. This is done by the so-called TRUNC/SG optimization procedure (Bergadano 1988b, c; 1991). The procedure generalizes some parts and specializes other parts of the consistent and complete description in order to maximize a *general description quality criterion*. The ICI of the concept was a set of expert-generated rules for covering exceptional cases identified by the system. As shown in Table 1, the "Optimized TT description" had the highest predictive accuracy, and was also the next simplest, after the "Top rule" description.

6. The Tiered View vs. Other Views

As mentioned earlier, cognitive science literature distinguishes among several *views* of concept representation. The major ones are the classical view, the prototypical (or probabilistic) view, an exemplar view, and the theory view (Smith and Medin, 1981; Medin and Smith, 1984; Nosofsky, 1987, Allen et al., 1988; Murphy and Medin, 1985; and chapters by Barsalou and Hale, by Hampton, and by Murphy in this book). Some authors also distinguish between different *types of formalisms* (or representation languages) employed in the concept representation, e.g., feature lists, decision trees, connectionist models, predicate logic expressions, simple and complex frames (e.g., Barsalou and Hale, this book). The distinction between a "view" and a "representational formalism" is usually not well explained in the literature. Our interpretation is that a view of a concept concerns inherent aspects of concepts as sets, while the representational formalism is a language in which those aspects are expressed. In principle, a sufficiently rich representational formalism, e.g., predicate calculus, with a proper interpretation, could be used for expressing any of the concept views.

The classical view assumes that concepts have well-defined borders, and can be represented by singly necessary and jointly sufficient conditions. This implies that they can be represented.by logic-style conjunctive definitions. The inadequacy of this view has been widely recognized, e.g., Wittgenstein, 1922; McCloskey and Glucksberg, 1978; Barsalou and Medin, 1986; and Lakoff, 1987. To overcome the limitations of this view, other views have been advanced, such as the *prototype view* and the *exemplar view* (e.g., Smith and Medin, 1981; Medin and Smith, 1984; Nosofsky, 1987, Allen et al., 1988).

The prototype view represents concepts by prototypes, and uses the so-called *family resemblance* principle (e.g., Rosch and Mervis, 1975). Unlike the classical view, this view does not require that all concept instances share the same properties. It can, therefore, explain the observed phenomenon that instances of the same concept may differ in their *typicality*. The exemplar view claims that concepts can be adequately represented by their exemplars (e.g., Smith and Medin, 1981; Bareiss, Porter and Craig, 1990). Both these views can also be criticized on various grounds. The prototype view, which formally is based on the idea of linear separability, disregards the existence of correlations between the attributes, the context-dependency, and other information that has been shown to be relevant to human concept understanding (e.g., Estes, 1986; Flannagan, Fried and Holyoak, 1986).

The exemplar view employs similarity-based and context-sensitive matching, which has received support in the cognitive science literature. It is also supported by an observation that people do store examples of concepts, especially exceptional examples. This view disregards, however, the importance of general concept descriptions that clearly play a role in human concept representations. Such general descriptions are useful in many ways, for example, for quickly identifying the differences between concepts, for recognizing them from partial information, for handling context-dependency, for efficiently storing invariant information about concepts, etc. The above operations would be difficult to perform, if concepts were represented only by examples. In some work using the exemplar view, general aspects of concepts are captured under the idea of "category structure," which is a network of domain knowledge that specifies the relevance of exemplars to the concept they define (Bareiss, Porter and Craig, 1990).

Some recent work has advocated a theory view (or knowledge-based view) which emphasizes the need to define concepts through their role in theories in which they exist as interrelated components (Carey, 1985; Hofstadter, 1985; Murphy and Medin, 1985; Schank, Collins and Hunter, 1986; Murphy - this book).

The two-tiered (TT) view constitutes a departure from the existing approaches, although it has a close relationship to most of them. It assumes that concepts have a certain central tendency, and proposes to describe this tendency explicitly, in the form of a logic style description. Such a description is the "first approximation" of the concept. On the other hand, it also assumes that concepts may have a great variability and be context-dependent, and that these aspects are best handled not explicitly, but implicitly, by a flexible matching method and context- and background knowledge-dependent rules of inference.

Since the TT view proposes to represent basic concept properties (BCR) in the form of explicit logic-style expressions, it is related to the classical view (it admits, however, that the BCR can be in the form of conjunctive descriptions, as well as disjunctive descriptions). It also recognizes that most concepts have a central tendency, and therefore one needs to distinguish between more or less typical concept examples. In this sense, the TT view is similar to the probabilistic view and the fuzzy set concept representation (e.g., Zadeh, 1965 and 1976). Unlike these views, it proposes to handle this aspect by the inferential concept interpretation (ICI) that employs a flexible concept matching procedure (Bergadano et al., 1991).

The TT view has also a relationship to the exemplar view, as it postulates the use of sophisticated matching procedures and inference rules in classifying new instances. Also, the TT view recognizes the usefulness of storing individual examples, especially in the case of exceptions (e.g. Zhang, 1990a, b). The TT approach is also closely related to theory view, as it stresses the role of background knowledge and inference in concept representation, especially in the case of non-typical or borderline instances. The main differences between the TT view and the theory view are that the former makes a distinction between the explicit and implicit representation, and postulates that the agent's knowledge or "theory" plays a role primarily in the inferential concept representation, rather than in the basic concept representation.

7. Summary and Open Problems

The most significant aspect of the presented two-tiered view is that it represents concepts as a combination of an explicit and implicit representations. This stands in contrast to other views that attempt to compact concept representation into a single monolithical structure, whether as a conjuction of properties (as in the classical model), or as a linear weighted function (as in the prototype model). In the two-tiered representation, the first tier, the base concept representation (BCR), captures the explicit and common concept meaning, and the second tier, the inferential concept interpretation (ICI), defines allowable modifications of the base meaning and exceptions. The typical concept instances strictly match the BCR, and thus can be recognized efficiently. The exceptional and/or context-dependent cases involve the inferential concept interpretation, which takes more time, but is done relatively less often. Such a two-tiered representation is particularly suitable for learning flexible concepts, i.e., concepts that inherently lack precise definition and are context-dependent. The TT view asserts that it is impossible, in principle, to give such concepts a precise meaning in an explicit and context-independent form. Therefore, it proposes to describe precisely only their central tendency, and to use flexible matching procedures and inference rules to characterize the concept's variability, to resolve borderline cases, and to express their context-dependency.

The TT view seems to be confirmed by an observation that people usually can recognize rapidly and without any difficulty typical concept instances, e.g., that chair is an example of furniture. When, however, they need to decide about special borderline cases (e.g., if a lamp or painting is an example of furniture), they try to resolve the problem by reasoning and/or an evaluation of the similarity to the cases with clear membership. This seems to contradict approaches that try to resolve such cases by weighting features or statistical evaluation. In the TT view, the reasoning and similarity eventuation is conducted by matching procedures, inference rules and agent's background knowledge residing in the second tier of the concept representation.

The TT view also assumes that the imprecision of human concepts stems not from an undesirable vagueness of our concept definitions, but rather from the universal need for cognitive economy. By allowing concepts to have a context-modifiable meaning, and making them precise only to the extent to which a given situation and/or context requires them to be precise, the expressive power of concepts is greatly enhanced. This means that one can employ fewer concepts for expressing more meanings, and this helps us to simplify our descriptions of our immensely complex universe. Thus, by compacting the variability of concepts into rules and matching procedures, rather than trying to express it explicitly, one can reduce the overall complexity and improve the effectiveness of concept representation. The experiments on learning TT representations have demonstrated that the obtained TT representations tend to be considerably simpler and have better predictive accuracy than other representations.

The ideas of TT representation are recent, and many problems remain unsolved. Among especially interesting and important problems is how to represent not just one, but a complex system of concepts using the TT representation. As indicated in the paper, the ICI usually can be shared or inherited from higher level concepts. This problem is then how to allocate different parts of the ICI within a structure of TT representation. There is also a need for a clearer analysis of the relationship between the TT view and other views, and experimental studies with human subjects to determine the adequacy of the TT model as a cognitively viable concept representation. There are also many problems related to the development of systems capable of learning TT representations, and efficiently using them for concept recognition processes. In particular, an important problem is how people learn individual tiers of different concepts.

Finally, a theoretical analysis and more experimental investigations are needed to confirm or disconfirm the initial finding that the TT representation can lead to significant memory savings and a simultaneous improvement of the predictive accuracy of concept representations over other representations.

ACKNOWLEDGEMENTS

The author thanks Larry Barsalou, James Hampton, Iven Van Mechelen, Zenon Kulpa, and an unkown reviewer for the critisim and useful comments on the earlier version of this chapter. He also expresses his gratitude to George Tecuci, Mike Hieb, and members of the Machine Learning Doctoral Seminar at George Mason University for useful discussions and suggestions.

This research was done in the Artificial Intelligence Center of George Mason University. Research activities of the Center are supported in part by the Defense Advanced Research Projects Agency under grants No. N00014-87-K-0874 and N00014-91-J-1854 administered by the Office of Naval Research, and in part by the Office of Naval Research under grants No. N00014-88-K-0226, No. N00014-88-K-0397 and No. N00014-91-J-1351, and the National Science Foundation under grant No. IRI-9020266.

BIBLIOGRAPHY

Allen, S.W., Brooks, L.R. and Norman, G.R., "Effect of Prior Examples on Rule-based Diagnostic Performance," *Proceedings of the 29th Annual Meeting of the Psychological Society,* November 1988.

Bareiss, E.R., Porter, B.W. and Craig, C.W., "PROTOS: An Exemplar-Based Learning Apprentice," chapter in *Machine Learning: An Artificial Intelligence Approach Volume III*, Y. Kodratoff and R.S. Michalski (eds.), Morgan Kaufmann Publishers, 1990.

Barsalou, L.W., "The Instability of Graded Structure in Concepts," in *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, U. Neisser (Ed.), pp. 101-40, Cambridge University Press, 1987.

Barsalou, L. W. *Cognitive Psychology: An Overview for Cognitive Scientists,* to appear in D.A. Norman and A. Ortony's (Eds.), Tutorial series on Cognitive Science, Lawrence Erlbaum Associates, 1992.

Barsalou, L.W. and Hale, C.R., "Components of Conceptual Representation," in *CATEGORIES AND CONCEPTS: Theoretical Views and Inductive Data Analysis,* I. Mechelen, J. Hampton, R.S. Michalski and P. Theuns, Academic Press, to appear in 1992.

Barsalou, L.W. and Medin, D.L., Concepts: Fixed Definitions or Dynamic Context-Dependent Representations, *Cahiers de Psychologie Cognitive*, Vol. 6, pp.187-202, 1986.

Bergadano, F., Matwin, S., Michalski, R.S. and Zhang, J., "Measuring Quality of Concept Descriptions," *Proceedings of the Third European Working Session on Learning,* D. Sleeman, Pitman, London, pp. 1-14, 1988a [An extended version was published under the title "A Measure of Quality of Descriptions," *Reports of the Machine Learning and Inference Laboratory,* Vol. 88-3, Artificial Intelligence Center, George Mason University, 1988a.]

Bergadano, F., Matwin, S., Michalski, R.S., and Zhang, J., "Learning Flexible Concept Descriptions Using a Two-tiered Knowledge Representation: Implementation and Experiments," *Reports of the Machine Learning and Inference Laboratory*, Vol. 88-35, Artificial Intelligence Center, George Mason University, 1988c.

Bergadano, F., Matwin, S., Michalski R. S. and Zhang, J., "Learning Two-tiered Descriptions of Flexible Concepts: The POSEIDON System, *Machine Learning Journal and Reports of the Machine Learning and Inference Laboratory, MLI 91-3*. Center for Artificial Intelligence, George Mason University, Fairfax, VA, May 1991. (Accepted for publication in *Machine Learning Journal,* 1992.)

Bergadano, F., Matwin, S., Michalski, R.S., and Zhang, J., "Learning Flexible Concept Descriptions Using a Two-tiered Knowledge Representation: Ideas and a Method," *Reports of the*

Machine Learning and Inference Laboratory, Vol. 88-4, Artificial Intelligence Center, George Mason University, 1988b.

Bratko, I., Mozetic, I. and Lavrac, L., *Kardio: A Study in Deep and Qualitative Knowledge for Expert Systems,* MIT Press, Cambridge, MA, 1989.

Carey, S., *Conceptual Change in Childhood,* MIT Press, Cambridge, MA, 1985.

Chestnik, B., Kononenko, I. and Bratko, I., ASSISTANT 86: A Knowledge Elicitation Tool for Sophisticated Users, *Proceedings of the 2nd European Working Session on Learning,* I. Bratko and N. Lavrac (ed.), Bled, Yugoslavia, May 1987.

Collins, A. and Michalski, R.S., "The Logic of Plausible Reasoning: A Core Theory," *Cognitive Science,* Vol. 13, 1989.

Collins, A. and Smith, E.E., *Readings in Cognitive Science: A Perspective from Psychology and Artificial Intelligence,* Morgan Kaufmann, 1988.

Estes, W.K., "Memory Storage and retrieval Processes in Category Learning," *Journal of Experimental Psychology: General,* Vol. 115, pp.155-175, 1986.

Flannagan, M.J., Fried, L.S. and Holyoak, K.J., "Distributional Expectations and the Induction of Category Structure," *Journal of Experimental Psychology,:Learning, Memory and Cognition,* Vol. 12, pp. 241-56, 1986.

Fulk, M.A. and Case, J., *Proceedings of the Thrid Annual Workshop on Computational Learning Theory* (COLT 90), University of Rochester, Rochester, New York, August 6-8, 1990.

Gentner, D., and Landers, R., "Analogical Reminding: A Good Match is Hard to Find," *Paper prepared for the Panel on Commonsense Reasoning at the International Conference on Systems, Man and Cybernetics,* Tucson, Arizona, 1985.

Gross, P.K., "Incremental Multiple Concept Learning Using Experiments," *Proceedings of the Fifth International Conference on Machine Learning,* Ann Arbor, June 12-14, 1988.

Hoff, W., Michalski, R.S., and Stepp, R.E., "INDUCE.3: A Program for Learning Structural Descriptions from Examples," *Report ISG 83-4, UIUCDCS-F-83-904,* Dept. of Computer Science, University of Illinois, Urbana, 1983.

Hofstadter, D.R., "Analogies and Roles in Human and Machine Thinking," chapter in *Metamagical Themas: Questing for the Essence of Mind and Pattern,* Basic Books, Inc., 1985.

Holyoak, K.J., "The Pragmatics of Analogical Transfer," *The Psychology of Learning and Motivation,* 19, 1985.

Hong, J., Mozetic, I., and Michalski, R.S., "AQ15: Incremental Learning of Attribute-Based Descriptions from Examples, The Method and User's Guide," *Report ISG 86-5, UIUCDCS-F-86-949,* Dept. of Computer Science, University of Illinois, Urbana, 1986.

Iba, W., Woogulis, J. and Langley, P., "Trading Simplicity and Coverage in Incremental Concept Learning," *Proceedings of the Fifth Intern. Conference on Machine Learning,* Ann Arbor, June 12-14, 1988.

Keil, F. "Concepts, Word Meanings, and Cognitive Development," Harvard University Press, Cambridge, MA, 1989.

Kempler-Nelson, D.G., "The Effect of Intention on What Concepts are Acquired," *Journal of Verbal Learning and Verbal Behavior,* Vol. 23, pp. 734-59, 1984.

Lakoff, G., *Women, Fire and Dangerous Things: What Categories Tell UA About the Nature of Thought,* Chicago University Press, 1987.

Lenat, D.B., Hayes-Roth, F. and Klahr, P., "Cognitive Economy in Artificial Intelligence Systems," *IJCAI* , Vol. 6, Tokyo, Japan, 1979.

McCloskey, M. and Glucksberg, S., "Natural Categories: Well-defined or Fuzzy Sets?" *Memory and Cognition*, Vol. 6, pp. 462-472, 1978.

Medin, D.L. and Smith, E.E., "Concepts and Concept Formation," in *Annual Review of Psychology*, M.R. Rosenzweig and L.W. Porter (Eds.), Vol. 35, pp. 113-118, 1984.

Medin, D., "Concepts and Conceptual Structure," *American Psychologist,* Vol. 44, pp. 1469-1481, 1989.

Mehler, G., Bentrup, J., and Riedesel J., "INDUCE.4: A Program for Incrementally Learning Structural Descriptions from Examples," *Reports of Intelligent Systems Group*, Vol. 86, Department of Computer Science, University of Illinois, Urbana, 1986.

Michalski, R.S., "AQVAL/1--Computer Implementation of a Variable-Valued Logic System VL_1 and Examples of its Application to Pattern Recognition," *Proceedings of the First International Joint Conference on Pattern Recognition,* pp. 3-17, Washington, DC, October 30 - November 1, 1973.

Michalski, R.S., "Variable-Valued Logic: System VL_1," *Proceedings of the 1974 International Symposium on Multiple-Valued Logic,* West Virginia University, Morgantown, pp. 323-346, May 29-31, 1974.

Michalski, R.S., "A Planar Geometrical Model for Representing Multi-Dimensional Discrete Spaces and Multiple-Valued Logic Functions," *Report No. 897*, Department of Computer Science, University of Illinois, Urbana, January 1978a.

Michalski, R.S., Designing Extended Entry Decision Tables and Optimal Decision Trees Using Decision Diagrams, *Report No. 898*, Department of Computer Science, University of Illinois, Urbana, March 1978b.

Michalski, R.S., "Theory and Methodology of Machine Learning." In R.S. Michalski, J.G. Carbonell, T.M. Mitchell (Eds.), *Machine Learning: An Artificial Intelligence Approach,* Morgan Kaufmann, Palo Alto, Tioga, 1983.

Michalski, R.S., "Concept Learning," *Encyclopedia of Artificial Intelligence*, John Wiley & Sons, 1986.

Michalski, R.S., "Two-tiered Concept Representation, Analogical Matching and Conceptual Cohesiveness," *Invited paper for the Workshop on Similarity and Analogy*, Allerton House, University of Illinois, June 12-14, 1986. [An extended and improved version is in *Analogy and Similarity*, S. Vosniadou and A. Ortony (eds.), Cambridge University Press, 1989]

Michalski, R.S., "How to Learning Imprecise Concepts: A Method for Employing a Two-tiered Knowledge Representation in Learning," *Proceedings of the Fourth International Workshop on Machine Learning*, University of California at Irvine, June 22-25, 1987.

Michalski, R.S., "Dynamic Recognition: An Outline of a Theory on How to Recognize Concepts without Matching Rules," *Reports of Machine Learning and Inference Laboratory,* Center for Artificial Intelligence, George Mason University, June 1989.

R. S. Michalski, "Learning Flexible Concepts: Fundamental Ideas and a Method Based on Two-tiered Representation", in *Machine Learning: An Artificial Intelligence Approach, Vol. III*, Y. Kodratoff and R.S. Michalski (Eds.), Morgan Kaufman Publishers, Inc., San Mateo, 1990.

Michalski, R.S. "Toward A Unified Theory of Learning: An Outline of Basic Ideas," *Proceedings of the First World Conference on the Fundamentals of Artificial Intelligence*, Paris, France, July 1-5, 1991.

Michalski, R.S. and Chilausky, R.L., "Learning by Being Told and Learning from Examples: An Experimental Comparison of the Two Methods of Knowledge Acquisition in the Context of Developing an Expert System for Soybean Disease Diagnosis," *International Journal of Policy Analysis and Information Systems*, Vol. 4, No. 2, pp. 125-16, 1980.

Michalski, R.S. and Larson, J., "AQVAL/1 (AQ7) User's Guide and Program Description," *Report No. 731*, Dept. of Computer Science, University of Illinois, Urbana, 1975.

Michalski, R.S. and Larson, J.B., "Selection of Most Representative Training Examples and Incremental Generation of VL1 Hypotheses: the Underlying Methodology and the Description of Programs ESEL and AQ11," *Reports of the Department of Computer Science*, Vol. 867, University of Illinois, Urbana, May 1978.

Michalski, R.S. and McCormick, B.H., "Interval Generalization of Switching Theory," *Reports of the Department of Computer Science*, Vol. 442, University of Illinois, Urbana, 1971.

Michalski, R.S. and Negri, P., "An Experiment on Inductive Learning in Chess End Games," *Machine Representation of Knowledge, Machine Intelligence*, Vol. 8, pp. 175-192, E. W. Elcock and D. Michie (eds.), Ellis Horwood Ltd., New York, 1977.

Michalski, R.S. and Stepp, R.E., "Learning from Observations: Conceptual Clustering," in *Machine Learning - An Artificial Intelligence Approach,* R.S. Michalski, J.G. Carbonell, T.M. Mitchell (eds.), Palo Alto, Tioga, 1983.

Michalski, R.S. and Winston, P.H., "Variable Precision Logic." *Artificial Intelligence Journal,* 29, 1986.

Michalski, R.S., Mozetic, I., Hong, J. and Lavrac, N., "The AQ15 Inductive Learning System: An Overview and Experiments," *Proceedings of the American Association for Artificial Intelligence Conference (AAAI),* 1986.

Mumdani, E.H. and Gaines, B.R. (eds.), *Fuzzy Reasoning and its Applications*, Academic Press, 1981.

Murphy, G.L. and Medin, D.L., "The Role of Theories in Conceptual Coherence," *Psychological Review*, Vol. 92, 1985.

Nosofsky, R.M., "Examplar-based Accounts of Relations between Classification, Recognition and Typicality," *Journal of Experimental Psychology: Learning, Memory and Cognition*, 1988.

Popper, K.R., *Objective Knowledge: An Evolutionary Approach,* Oxford at the Clarendon Press, 1979.

Quinlan, R.J., "Learning Efficient Classification Procedures and their Application to Chess End Games," in *Machine Learning: An Artificial Intelligence Approach*, R.S. Michalski, J.G. Carbonell, T.M. Mitchell (eds.), Palo Alto, Tioga, 1983.

Quinlan, R.J., "Probabilistic Decision Trees," in *Machine Learning: An Artificial Intelligence Approach III,* Y. Kodratoff and R.S. Michalski (eds.), Morgan Kaufmann Publishers, 1990.

Reinke, R.E. and Michalski, R.S., "Incremental Learning of Decision Rules: A Method and Experimental Results." in J.E. Hayes, D. Michie, J. Richards (eds.), *Machine Intelligence,* Oxford University Press, 1988.

Rips, L.J. "Similarity, Typicality, and Categorization," in *Analogy and Similarity*, S. Vosniadou and A. Ortony (eds.), Cambridge University Press, 1989.

Rosch, E. and Mervis, C.B., "Family Resemblances: Studies in the Structure of Categories," *Cognitive Psychology*, Vol. 7, 1975.

Schank, R.C., Collins, G.C. and Hunter, L.E., "Transcending Induction Category Formation in Learning," *The Behavioral and Brain Sciences*, 1986.

Smith, E.E., "Concepts and Induction," in *Foundation of Cognitive Science*, Michael I. Posner (ed.), A Bradford Book, The MIT Press, Cambridge, MA, pp. 501-526, 1989.

Smith, E.E. and Medin, D.L., *Categories and Concepts*, Harvard University Press, Cambridge, MA, 1981.

Sutcliffe, J.P., "Concept, Class and Category in the Tradition of Aristotle," in *Categories and Concepts: Theoretical Views and Inductive Data Analysis,*. Iven Van Mechelen, James Hampton, Ryszard S. Michalski and Peter Theuns (eds.), Academic Press Ltd., 1992.

Sowa, J.F., *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley Publishing Company, 1984.

Stillings, N. A., Feinstain, Mark, H., Garfield J. L., Rissland E. L., Rosenbaum D. A., Weisler S. E. and Baker-Ward Lynne, Cognitive Science: An Introduction, The MIT Press, Cambridge, MA, 1987.

Stuart, E., "Computerized Construction in Fortran of a Discriminant Function for Categorical Data", *Applied Statistics*, Vol. 30, pp. 213-22, 1981.

Watanabe, S., "Knowing and Guessing, a Formal and Quantitative Study," Wiley Publishing Co., 1969.

Winston. P.H., "Learning Structural Descriptions from Examples," *The Psychology of Computer Vision,* Winston P.H. (Ed.), McGraw Hill, New York, Ch. 5, 1975.

Wittgenstein, L., *Tractatus Logico-Philosophicus*, London, Routledge and Kegan, 1922.

Zadeh, L.A., "A Fuzzy-algorithmic Approach to the Definition of Complex or Imprecise Concepts," *International Journal of Man-Machine Studies*, Vol. 8, 1976.

Zadeh, L.A., "Fuzzy Sets," *Information and Control*, Vol. 8, 1965.

Zhang, J. and Michalski, R.S., "A Description Preference Criterion in Constructive Learning," *Proceedings of the Sixth International Workshop on Machine Learning*, Cornell University, Ithaca, New York, June 26-27, 1989.

Zhang, J., "A Method that Combines Inductive Learning with Exemplar-Based Learning" In *Proceedings of Tools for Artificial Intelligence*, IEEE Computer Society Press, 1990b.

Zhang, J., "Learning Flexible Concepts from Examples: Employing the Ideas of Two-Tiered Concept Representation." *Ph.D Thesis*, Department of Computer Science, University of Illinois at Urbana-Champaign, 1990a.