DYNAMIC RECOGNITION: AN OUTLINE OF THEORY OF HOW TO RECOGNIZE CONCEPTS WITHOUT MATCHING RULES, INTELLIGENCE SYSTEMS GROUP

by

R. S. Michalski

ISG-16, UIUCDCS-F-86-965., Department of Computer Science, University of Illinois, Urbana, 1986

1720-15

DYNAMIC RECOGNITION:

An Outline of Theory of How to Recognize Concepts without Matching Rules

Ryszard S. Michalski
Intelligent Systems Group
Department of Computer Science
University of Illinois at Urbana-Champaign

Abstract

Existing approaches to the recognition problem assume that to recognize a concept one needs to match the observed data with a stored description of the concept. When the properties of the observed object match the stored description of a concept, then the name of the concept is returned. A match can be strict or approximate; in the latter case the degree of match is computed and used as a measure of confidence in recognizing the concept.

A major problem with such methods is that in order to recognize an object one always needs to measure the same properties of it. Yet, people can recognize the same concept using many different subsets of properties, without storing a correspondingly many different concept descriptions. A theory of *dynamic recognition* is presented that shows how a recognition system can recognize a concept from many different subsets of properties but using only one concept description.

INTRODUCTION

There are few intellectual processes more universal than the processes of recognition. Recognizing concepts in a given input stream of observations is one of the most ubiquitous acts of human mind. Since the early fifties there have been many research activities trying to explain recognition processes and build computational models of them. There now exists a whole field of research bearing the name *pattern recognition*. Statistical decision theory, decision-theoretic methods, syntactic approaches, and logic-based methods are among the basic techniques proposed for solving the recognition problem. Yet, the flexibility and ease with which the human mind can handle recognition problems in all kinds of conditions remains unexplained and mysterious.

To *re*-cognize a concept is to assign a concept name to a stream of observational data that relates in some way to the stored description of the concept. The basic assumption made in concept recognition techniques used now is that in order to assign a concept **C** to a stream of data C one needs to find a match between C and a description of the concept **C**. Statistical distributions, production rules, grammars and logic expressions are among the basic types of concept descriptions used. In expert systems, for example, decisions or actions are decided by matching the input data with the condition parts of production rules. Typically, one or at most a few conjunctive conditions are linked with one decision in a production rule.

This approach does not explain, however, how people can recognize the same concept in so many different ways. For example, one may be able to recognize a given person by seeing just a part of his face, or his back, by hearing his voice, seeing his handwriting, or even by observing a silhouette of this person in the dark. Each of these pieces of information can appear in all kinds of specific forms and manifestations, and thus the number of different subsets of data items that can trigger recognition can practically be infinite.

Clearly, it is completely unsatisfactory to have one rule for each subset of data that is sufficient to recognize an object. But this is, basically, how our current recognition systems work.

This paper proposes a computational theory that explains how a recognition system could recognize a given concept from many different subsets of information about it, but using only one concept description. I assume here that the reader is familiar with basic concepts of variable-valued logic and algorithm AQ for solving the general covering problem employed in various inductive learning systems (e.g., Michalski, 1972, 1977).

PREMISES OF THEORY

The main idea behind dynamic recognition is that the system uses a *characteristic description** of each concept, and recognizes an object by conducting inductive inference on candidate concept descriptions. Thus, recognition is viewed not as a process of matching but as inductive inference that determines the discriminant features between concepts in a given context. The proposed method for dynamic concept recognition involes three stages:

1. REDUCE
2. INDUCE
3. INQUIRE

Given some facts about an object, these facts are used to "reduce" existing concept descriptions (the REDUCE stage). If concept is not recognized at this stage, the inductive procedure AQ is applied to the reduced rules to determine the simplest *discriminant* recognition rules (the INDUCE stage). The difference between the standard form of the AQ algorithm and this form is that in our case events are

---

*** By a characteristic description of a concept is meant a collection of all information known about it. It is expressed as a conjunction of assertions that are true about this concept, or about all known instances of this concept.**

rules. After discriminant rules are generated, a question answering process is activated that lasts until one of the descriptions becomes true. When this happens, the concept associated with this description is recognized. At each step of the question answering process the algorithm INQUIRE determines a question that fosters maximal reduction of the rules. Let us now proceed in more detail.

ALGORITHMS

1. Algorithm REDUCE

Given:

1. Characteristic descriptions of concepts (i.e., collections of assertions about each concept)
2. Zero or more initial facts about the object to be recognized.

Determine:

A set of *candidate concept descripions*. Such a set contains a subset of the original concept descriptions, and individual descriptions may contain fewer assertions (as some of the assertions are replaced by TRUE).

The algorithm determines which selectors are satisfied and which are not by the initial facts. Rules that have dissatisfied selectors are removed. Selectors that are satisifed are removed from rules.

2. Algorithm **INDUCE**

Given:

1. Candidate Concept Descriptions
2. Background knowledge (containing specification of variables and their types, background concepts constructive generalization rules, and preference criterion)

Determine:

Discriminant concept recognition rules that optimize the given criterion of preference.

The program AQ15 is applied to the characteristic descripions obtained by algorithm REDUCE. These descriptions play the role of learning events for the program. The output are most preferred discriminant concept recognition rules (Hong, Mozetic and Michalski, 1986).

3. Algorithm **INQUIRE**

Given:

> A set of discriminant concept recognition rules.

Determine:

> A question that maximally reduces the concept recognition rules, or in general, a variable that represents the best choice for maximally reducing rules with minimum-cost information (the *utility-based* control schema).

In selecting the question, one may also apply the criterion of *conceptual closeness* between the previously selected and current variables, and/or take into consideration that determining just one or all variables from some set may carry the same cost. A very simple method for variable selection is to determine the one that occurs in the maximum number of rules.

EXAMPLE

Let us consider an example of the complete recognition process. Suppose the following characteristic concept descriptions are given:

[D = 3] <=== [X_1 = 5][X_3 = 7][X_8 = 2..8]

[D = 1] <=== [X_3 = 3][X_4 > 5][X_6 = 7]

[D = 2] < === [X_1 = 7 v 8][X_3 = 6 v 8][X_8 = 1..7] .

The operator <=== is interpreted: the concept specified in the LHS is characterized by assertions in the RHS.

Case 1. No initial facts

Because there are no initial facts, the REDUCE stage is skipped. One of the sets of discriminant concept recognition rules produced at the INDUCE stage is:

[X_3 = 7] ===> [D = 3]
[X_3 = 3] ===> [D = 1]
[X_3 = 6 v 8] ===> [D = 2].

The operator ===> is interpreted: the statement in the LHS is a sufficient condition to recognize the concept specified in the RHS.

The INQUIRE stage generates a question: What value does X_3 have? If the answer is, for example, X_3 = 3, then we have
TRUE ===> [D = 1], and the system returns the answer 3 as the recognized concept.

Case 2. Given are initial facts:

$X_8 = 7$; $X_4 = 3$

The REDUCE stage produces the following reduced characteristic descriptions:

$[X_1 = 5][X_3 = 7]$ ===> $[D = 3]$
$[X_1 = 7 \text{ v } 8][X_3 = 6 \text{ v } 8]$ ===> $[D = 2]$.

These descriptions are input to an inductive program that determines the minimal cost discriminant descriptions (concept recognition rules). One of the two alternative discriminant descriptions produced by the INDUCE step is:

$[X_3 = 7]$ ===> $[D = 3]$
$[X_3 = 6 \text{ v } 8]$ ===> $[D = 2]$.

Another one is:

$[X_1 = 5]$ ===> $[D = 3]$
$[X_1 = 7 \text{ v } 8]$ ===> $[D = 2]$.

To decide which ruleset to select, the system uses *costs* associated with determining values of individual variables. The ruleset that contains variables with the smalest total cost is selected. For example, suppose that $\text{cost}(X_1) = 10$ and $\text{cost}(X_3) = 1$. The system selects the first ruleset.

The value of variable X_3 is then determined - by asking a question, by conducting an experiment, or by inferring it from known information. Once the value is known, the rule that fires indicates the decision.

CONCLUSION

The presented method of concept recognition replaces the traditional rule matching by an inductive process of discrimination. The whole process of concept recognition consists of three phases: REDUCE, INDUCE and INQUIRE. In the REDUCE step the context, and an a priori supplied information about an object to be recognized is used to determine a candidate set of hypotheses, and reduce the characteristic concept descriptions for these hypotheses.

The second step, INDUCE, constructs goal-oriented discriminant concept descriptions from the reduced characteristic descriptions. The last step involves determining the information needed for final recognition.

Among topics for further research is an extension of the presented method to deal with partial satisfaction of characteristic descriptions in the REDUCE step, and partial satisfaction of discriminant recognition rules in the INQUIRE step, so that the recognition process may generate the *certainty degree* of recognition.

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation under grant No. DCR-8406801, by the Defence Advanced Research Projects Agency (DARPA) under grant No. N00014-K-85-0878, and by the Office of Naval Research under grant No. N00014-12-K-0186.

REFERENCES

Michalski, A Variable-Valued Logic System as Applied to Picture Description and Recognition, in *Graphic Languages*, F. Nake, A. Rosenfeld (Eds.), North-Holland, 1972.

R. S. Michalski, Variable-valued Logic and Its Application to Pattern Recognition and Machine Learning, chapter in *Computer Science and Multiple-valued Logic: Theory and Applications*, 1977.

R. S. Michalski, A Theory and Methodology of Inductive Learning, *Machine Learning: An Artificial Intelligence Approach*, TIOGA, 1983.

Hong, J., Mozetic I. and Michalski, R. S., AQ15: An Incremental Learning System from Examples, Reports of Intelligent Systems Group, 1986.