LEARNING (COMPLEX) STRUCTURAL DESCRIPTIONS FROM EXAMPLES

REGINE LOISEL, YVES KODRATOFF
G.R.22 du C.N.R.S., Institut de Programmation, Tour 55-65
4 Place Jussieu 75230 Paris Cedex 05 France

SUMMARY
We present a formalization of an intuitively sound strategy for learning a description from examples : within a partition examples are grouped according to greatest resemblances and examples not in the same subset show a maximum of differences.

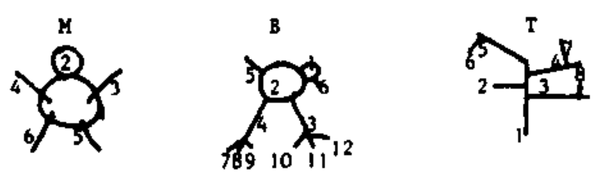
I. INTRODUCTION
WINSTON [4] has demonstrated the importance of the near-miss concept in a context of learning descriptions from examples. His methodology is practical when a few simple scenes are dealt with. We have extended it to include numerous complex examples.

The definition of the near miss concept [3 ,4] will be summarized in section 3.

A problem arises from the fact that a large number of near-misses can be obtained which do not convey the same type of information. Our experience shows that at least three types of near-misses must be introduced : highly ambiguous, ambiguous and discriminant near-misses, each conveying a different type of information. A second problem concerns the building of a structural description when several examples of several concepts are given, i.e. there are so many possible descriptions that a choice must be made of the most suitable as a recognition device. This leads us to define "promising" partitions of a set of examples. Given a set of examples, the description will be a tree recursively constructed by the rule : divide the set into its most promising partitions. An example of our methodology is given below. This example is actually too simple for the system and must be considered only as an illustration of the definitions we propose.

2. EXAMPLES
Given 3 concepts, each of them illustrated by only one example :

"Man" **"Bird"** **"Tree"**



Each labelled stroke is one element part of the example.

The first step consists of defining the reference stroke : e.g, that which is most centrally located, or that which contains other element, or possibly that which is directly above the central stroke.

This choice is very important as the whole process depends on it, but an explanation of the reasons underlying the choice is not within the scope of this paper.

Once a reference or "most important" stroke has been chosen, its relationships with the other strokes are computed using rules very similar to those usually used in formal descriptions [1,21. For example, in "Man" M the relationship between 1 and 2 is expressed by : 1(1,0,0,-1)2, meaning that, a) 1 and 2 are tangent, b) 1 and 2 are not inside each other, c) 1 and 2 have their centers on the same vertical axis, d) 1 is directly below 2. Here 1 is the first stroke, (1,0,0,-1) is the relationship between 1 and 2,2 is the second stroke. The only important point about this description is that it is already a generalization of the description since, for example, relationships between 3 and 5 will never be looked for.

We thus obtain a list of sublists which is a description of these three examples (this has to be accepted by the reader) :

M : ((1(1,0,0,-1)2),(1(3,0,-1,-5)3),(1(3,0,1,-5)4), (1(3,0,-5,1)5,(1(3,0,5,1)6))

B : ((2(2,4,-1,0)1),(2(3,0,-5,1)3),(2(3,0,5,1)4), (2(5,0,1,-5)5), (2(0,0,1,5)6), (2(0,0,4,3)7), (2(0,0,5,1)8),(2(0,0,5,1)9),(2(0,0,-5,3)10), (2(0,0,-5,1)11), (2(0,0,-4,1)12)).

T : ((1(1,0,1,-1)2) ,(1(1,0,-5,-1)3),(1(1,0,-5,-1) 4),(1(1,0,5,-1)5),(1(0,0,4,-3)6),(1(0,0,0,-1)7), (1(0,0,-4,-3)8))

3. NEAR-MISSES
Given that M1,B2 and TI are considered as reference strokes of the 3 concepts, the definition of a <u>near-miss</u> is :

If two sublists of two different examples refer to a stroke of the same importance and if the relationships match except for one and only one of their elements, then the conjunction of the two relationships is a near-miss between the two examples.

For example, considering the lists of M and T, the sublists (1(1,0,0,-1)2) and (1(1,0,5,-1)5) constitute a near-miss : central (1,0,0,-1/1,0,5,-1), since they both refer to the central stroke and their relationship matches except for one of its values.

Let EI and E2 be two examples. Let r1 a relationship of E1 which constitutes a near-miss with a relationship r2 of E2.

A near-miss is <u>highly ambiguous</u> when rl also belongs to E2 and r2 also belongs to EI.

For example, central **(3,0,5,1/3,0,-5,1)** is highly ambiguous .

A near-miss is ambiguous when rl also belongs to E2 or (exclusive or) r2 also belongs to E1.
For example, central (3,0,5,1/0,0,5,1) is ambiguous.
A near-miss is discriminant when rl does not belong to E2 an r2 does not belong to E1.
For example central (3,0,1,-5/5,0,1,-5) is discriminant. From the data given in section 2, it follows that : near-misses between M and B are :
highly ambiguous : (3,0,5,1/3,0,-5,1)
ambiguous : (3,0,5,1/0,0,5,1) : (3,0,-5,1/0,0,-5,1)
separating : (3,0,1,-5/5,0,1,-5)
near-misses between M and T are :
discriminant (1,0,0,-1/1,0,5,-1) : (1,0,0,-1/1,0,-5,-1), (1,0,0,-1/0,0,0,-1)
near-misses between B and T are :
discriminant : (0,0,4,3/0,0,4,-3) ;(0,0,-4,1/0,0,

4. DEFINITIONS

4.1. Relationships between two examples Ei and Ej : Ei and Ej are highly comparable if there exists at least one separating near-miss and one highly ambiguous near-miss between them.
Ei and Ej are comparable if there exists no highly ambiguous near-miss between them and there exist at least one separable near-miss and one ambiguous near-miss beween them.
Ei and Ej are **separable** if there exists no highly ambiguous or ambiguous near-miss between them and there exists at least one separating near-miss between them.

4.2. Relationships between one example Ei and a set of examples A : Ei does not belong to A. Example : Ei - T,A » {M,B}
Let |A| be the number of examples in A

- High Ambiguity Value of Ei relative to A :
$HAV_A(Ei)$ • number of examples of A highly comparable to Ei divided by |A|

  Example : $HAV_{B\}}(T)$ - 0

- Ambiguity Value of Ei relative to A :
$AV_A(Ei)$ • HAV (Ei)+(number of examples of A comparable to Ei * number of examples of A belonging to the same concept as Ei)/ |A|

  Example : $AV_{\{MjB\}}(T)$ - 0

- Discrimination power of Ei relative to A :
S.(Ei) • AV.(Ei)+ number of examples of A separable from Ei/ IAI

  Example : S^jtt) - 1

4.3. Relationships between two disjoint sets of examples A and B :
Definitions are those of section 4.2 except that all examples of B must be summed over, for each Ei belonging to B.

5. MOST PROMISING PARTITION OF A SET

5.1. Partition of a set A by a near-miss n : n is defined by the relationships rl and r2. Let X be the subset of A which contains all the examples to which rl or (non-exclusive or) r2 belong. Let X' be the subset A-X.
The partition (X,X') is the partition of A by n.

5.2. Indexes of ambiguity and discrimination power of a set A of examples :
the definitions of section 4.1 are computed for all element pairs of A. Let ||A|| be the number of element pairs of A.

- High Ambiguity Index of A :
HAI(A) = number of highly comparable example pairs of A/ ||A||

  Example : HAI({M,B,T}) = 1/3

- Ambiguity Index of A :
AI(A) = HAI(A)+(number of comparable pairs of A + number of pairs belonging to the same concept)/ ||A||

  Example : AI({M,B,T}) = 1/3

- Discrimination Index of A :
SI(A) = IA(A) + number of discriminable pairs of A/||A||

  Example : SI({M,B,T}) = 1

5.3. Definition of a most promising partition
Let n1...np be the near-misses of the example pairs of A. We obtain p partitions of A by n1...np.
Let(X,X') and (Y,Y') be two of these partitions.
(X,X') is considered more promising than (Y,Y') when :
SI(X) > SI(Y) ifeqthen SI(X') > SI(Y') ifeqthen
$HAV_X(X') < HAV_Y(Y')$ ifeqthen $AV_X(X') < AV_Y(Y')$
ifeqthen $S_X(X') < S_Y(Y')$
ifeqthen HAI(X) > HAI(Y) ifeqthen HAI(X') > HAI(Y')
ifeqthen
AI(X) > AI(Y) ifeqthen AI(X') > AI(Y') ifeqthen |X|> |Y|
where the operator ifeqthen means : if the preceding values are equal then...

  Example : Let us consider n1=(3,0,5,1/3,0,-5,1) which creates the partition P1 : X = M,B, X'=T, and n2=(1,0,0,-1/1,0,5,-1) which creates P2 : Y=M,T, Y'=B. SI(X)=SI(Y)=SI(X')=SI(Y')=1
$HAV_X(X')=0$, $HAV_Y(Y')=1/2$. It follows that P1 is more promising than P2.

6. CONCLUSION

Comments on the intuitive meaning of the most promising partition : while ambiguous and discriminant near-misses both emphasize differences and resemblances between two examples, their respective proportions of information relative to resemblance vs. information relative to differences are not the same. Therefore, ambiguous and discriminant near-misses do not convey the same type of information. The role of the above defined indexes is to express this difference-resemblance play among examples. This means that our structural description will be built as follows. We first try to find discriminable sets (and their associated near-misses). When we have several discriminable partitions, we choose among them by minimizing the ambiguity between the two subsets of the partition. If a choice still remains, we maximize the ambiguity of each of the subsets. The last test simply tells which one has the greatest number of elements verifying the given near-miss.
This means that we want to keep together examples which are discriminable as long as possible.

In fact, we describe our work as a formalization of this sound strategy.

BIBLIOCRAPHIE

1. T.G. DIETTERICH, R.S. MICHALSKI
   Learning and generalization of characteristic descriptions : evaluation criteria and comparative review of selected methods
   Proc IJCAI 79 Tokyo 1979 pp 223-231

2. N.S. FU
   Syntactic methods in pattern recognition. Academic Press (197A)

3. Y. KODRATOFF
   Non statistical learning to analyse scenes by the use of the "near-miss" concept.
   Proc 4th IJCPR Kyoto (1979) pp 384-386

4. P.H. WINSTON
   Learning Structural descriptions from examples
   Ph. D. Thesis MIT(1970).

Note added in proof.

Using a quite different description than the above one, but using the same near-miss definition and our "most promising partition" concept, we very recently built an efficient recognition tree for a very different problem, namely the carcinogenic power of big chejnical molecules.