ON THE BEHAVIOUR OF SELF LEARNING CLASSIFICATION ALGORITHMS

J. Aguilar-Martin, Laboratoire d'Automatique et d'Analyse des Systèmes de Toulouse, BP 4-036, 31-Toulouse, (France)

Let us consider a set of items $x \in \mathcal{E}$, $\mathcal{E}$ is supposed to be topologically structured. Our observation of the set $\mathcal{E}$ is given by sequential extraction of items x_t considered exhaustive when t goes to infinite.

At each discrete instant $t_E[0, 1, 2, ...]$ we face the problem of:

1) Classifying x_t according to the existing partition P_{t-1} at time $t - \varepsilon$ on the set $E_{t-1} = \{x_0, x_1 ..., x_{t-1}\}$.

2) To modify the partition P_{t-1} into $\mathcal{P}_t$ using the new information carried by x_t.

Two situations are to be considered:

a) x_t can be classified in $\mathcal{P}_{t-1}$ and it only modifies eventually some attributes of the partition but not the number of classes.

b) x_t needs the creation of a new class thus causing the growing of the structure of partition.

It is obvious that for a hypothetical observer completely aware of the true structure of ε each partition $\mathcal{P}_t$ looks imperfect and the relation induced in ε fails to the axioms of equivalence relations.

Then we study the concept of distance or proximity of an element to a class $c \in \mathcal{P}$.

A classical point of view replaces c by a reprezentative element $x_c \in c$ and then uses the usual concept of distance under any of its classical forms because the symetry of the concept is preserved.

In our classification algorithms we try to use a slightly different concept of proxymity that has not this symmetry. Computer algorithms thus constructed show through numerical examples how this self learning procedure works.