TU Clausthal

Department of Informatics

Prof. Dr. J. Dix
Dr. N. Bulling
Lic. F. Schlesinger

Artificial Intelligence

– Sheet 4: Learning –

Date: 04. May 2011

Points:

_____ of 22

Group / Tutor:

Name(s) & Matr. no.:

Exercise 1 (8 Points, Version-space-learning)

Apply *version-space-learning* to the following examples where it is only allowed to use **conjunctions of attributes** (i.e. no disjunctions, no negations) as hypotheses:

Number	Director	Genre	Language	Duration	Reaction
1	Woody Allen	Comedy	English	Short	yes
2	J. Cameron	Action	German	Long	no
3	S. Spielberg	Comedy	English	Long	no
4	Woody Allen	Drama	English	Long	yes
5	Woody Allen	Drama	German	Short	no

Proceed as follows:

1. Choose the examples in the same order as given in the table.
2. *Each* time an example is added state the **G-set** and the **S-set**.
3. In *each* cycle of the algorithm explain what happened and why.
4. In each generalization (resp. specialization) step at most one attribute can be discarded from (resp. added to) a hypothesis, or a whole hypothesis can be left out.

Exercise 2 (3 Points)

We never test the same attribute twice along one path in a decision tree. Why not?

Exercise 3 (3 Points)

Suppose we generate a training set from a decision tree and then apply decision-tree learning to that training set. Is it the case that the learning algorithm will eventually return the correct tree **as the training set size goes infinity**? Why or why not?

Exercise 4 (4 Points)

In the recursive construction of decision trees, it sometimes happens that a mixed set of positive and negative examples remain at a leaf node, even after all attributes have been used. Suppose that we have p positive examples and n negative examples.

(a) Show that the solution used by Decision-Tree-Learning, which picks the majority classification, **minimizes the absolute error** over the set of examples at the leaf.

(b) Show that the **class probability** $\frac{p}{p+n}$ **minimizes the sum of squared errors**. (The sum of squared errors is given by $E(x) = p(1-x)^2 + nx^2$).

To be submitted:

25. May 2011
before class

Exercise 5 (4 Points)

This exercise concerns the expressiveness of decision lists.

(a) Show that decision lists can represent any Boolean function, if the size of the tests is not limited.

(b) Show that if the tests can obtain at most k literals each, then decision lists can represent any function that can be presented by a decision tree of depth k.