

Homework 1: Machine Learning 2D5362

Handed out: Thursday, 9.11.00

Due: Thursday, 16.11.00 : 13:30

Name:

Consider the instance space X consisting of integer points in the x, y plane ($0 \leq x \leq 9, 0 \leq y \leq 9$) and the set of hypotheses H consisting of axes-parallel rectangles. More precisely, hypotheses are of the form $(a \leq x \leq b, c \leq y \leq d)$, where a, b, c and d can be integers.

1. Consider the version spaces with respect to the set of positive(+) and negative(o) training examples shown on the second page. Trace the S- and G - boundaries of the version space using the CANDIDATE-ELIMINATION algorithm for each new training instance. Write out the hypotheses that belong to the S- and G-boundary and draw them into the diagram.
2. Suppose the learner may now suggest a new x, y instance and ask the trainer for its classification. Suggest a query guaranteed to reduce the size of the version space, regardless of how the trainer classifies it. Suggest one that will not.
3. Now assume that you are a teacher, attempting to teach a particular target concept (e.g. $3 \leq x \leq 5, 2 \leq y \leq 7$). What is the smallest number of training examples you can provide so that the CANDIDATE-ELIMINATION algorithm will perfectly learn the target concept?

