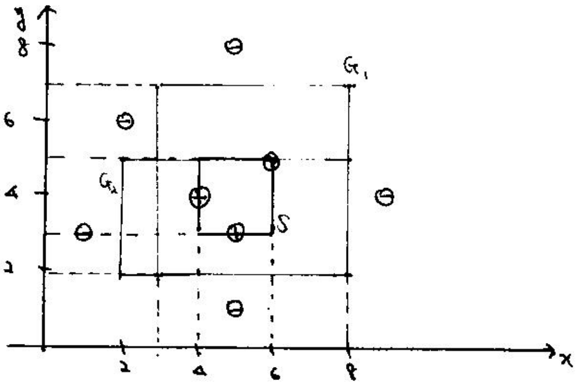
H.W #1.

1. <sol>

For example, A checkers learning problem:

- Task T : playing checkers
- performance measure P : percentage of the games won against opponent.
- experience E : playing practice games against itself.

2. (a)



Special boundary $S = \{ 4 \le x \le 6, \quad 3 \le y \le 5 \}$

General boundary $G_1 = \{ 3 \le x \le 8, \quad 2 \le y \le 7 \}$

$G_2 = \{ 2 \le x \le 8, \quad 2 \le y \le 5 \}$

$G = \{ <3 \le x \le 8, \ 2 \le y \le 7>, \ <2 \le x \le 8, \ 2 \le y \le 5> \}$

(b) - A query guaranteed to reduce the size of V.S :

(4, 6) , (3, 2) (8, 7) .. etc

- A query guaranteed not to reduce the size of V.S :

(5, 4) (5, 6) , .. etc.

3. <proof>

i) We will show that $\{h \in H : (\exists s \in S)(\exists g \in G)(g \geq_g h \geq_g s)\} \subseteq VS_{HD}$

Let $\begin{cases} s \text{ be an arbitrary member of } S \\ g \text{ be an arbitrary member of } G, \\ h \text{ be an arbitrary member of } H. \end{cases}$

Such that $g \geq_g h \geq_g s$.

By definition of S, s must be satisfied by all positive examples in D,
Since $h \geq_g s$, h must be satisfied by all positive examples in D,

By definition of G, g cannot be satisfied by any negative examples in D.
Since $g \geq_g h$, h cannot be satisfied by any negative examples in D.

$\therefore$ h is consistent with D.

$\therefore$ $h \in VS_{HD}$.

ii) We will show that $\{h \in H : (\exists s \in S)(\exists g \in G)(g \geq_g h \geq_g s)\} \supseteq VS_{HD}$.
Let $h \in VS_{HD}$ and assume that $h \notin \{h' \in H : (\exists s \in S)(\exists g \in G)(g \geq_g h' \geq_g s)\}$
Then it means $\forall s \in S$, $\forall g \in G$, $\neg(g \geq_g h \geq_g s)$.

$\iff$ $\forall s \in S$, $\forall g \in G$, either $\neg(g \geq_g h)$ or $\neg(h \geq_g s)$.

$\iff$ $\forall s \in S$, $\forall g \in G$, either $\exists x \in X$, $g(x)=0$ but $h(x)=1$.
or $\exists x' \in X$, $s(x')=1$ but $h(x')=0$.

Since for all negative examples $x \in X$ and for all $s \in S$, $s(x')=1$,
We can calculate that $\forall s \in S$, $\forall g \in G$
h misclassifies a negative example or a positive example.

$\Rightarrow$ $h \notin VS_{HD}$

This is a contradiction.

From (i). (ii). The version space representation thm is proved.

4. <proof>

In CE algorithm, the version space VS_{HD} is determined by the version space representation thm.

But, if there are errors in training examples, then we find a wrong version space $\widetilde{VS}_{HD}$ for the wrong training examples γ.

And if $c \notin H$, the target concept cannot be described in the hypothesis representation.

Thus, if (1) there are no errors in training examples and
(2) the concept $c \in H$.
are satisfied, then CE algorithm can always find the version space VS_{HD}. given the training examples D.

5. (a) - The inductive bias is the minimal set of assumptions which are assumed in the inference for new instances from the training data set.

- So we always need the inductive bias to infer $c(x)$ for new instances x's and if there is no inductive bias, we cannot get any inference for new instances.

(b) (1) Find-S algorithm

① The target concept can be represented in its hypothesis space

② All instances are negative instances unless the opposite is entailed by its other knowledge.

Sheet 1~4 Date

(2) CE algorithm

: The target concept can be represented in its hypothesis space

(3). ID3 algorithm

① Shorter trees are preferred over large trees.

② Trees that place high information gain attributes close to the root are prefered over those that do not.