Training Experience, E

Example	Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

X

$c \in \{0, 1\}$

Most general hypothesis : $\langle ?, ?, ?, ?, ?, ? \rangle$

Most specific hypothesis : $\langle \phi, \phi, \phi, \phi, \phi, \phi \rangle$

X : set of instances

c : target concept

$c : X \rightarrow \{0, 1\}$

Problem Specification:

Given

- Instances X
- Hypotheses H (Each hypothesis described as constraints on attributes)
- Target concept c : EnjoySport : $X \rightarrow \{0, 1\}$
- Training Examples D : +ve and -ve examples of the target function

Determine

- a hypothesis $h \in H$ s.t. $h(x) = c(x)$ $\forall x \in X$

Inductive learning hypothesis

Any hypothesis found to approximate the target function well over a sufficiently large set of training examples will also approximate the target function over other unobserved examples.

2.3 Concept learning as search

concept learning - the task of searching thro a large space of hypotheses.

goal - find the hypothesis that best fits the training data.

General-to-specific ordering of Hypotheses

searching through hypothesis space relies on general-to-specific ordering of hypotheses.

$h_1 = \langle Sunny, ?, ?, Strong, ?, ? \rangle$ (specific)

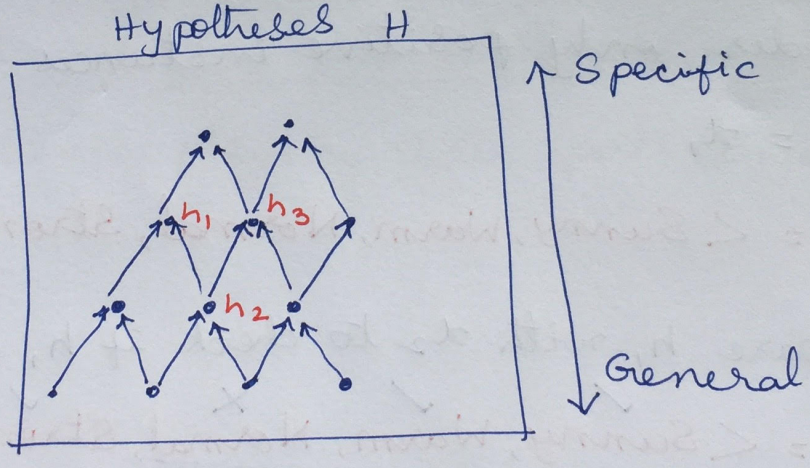
$h_2 = \langle Sunny, ?, ?, ?, ?, ? \rangle$ (general)

Let h_1, h_2 be two hypotheses defined over X

h_1 is more-general-than-or-equal-to h_2

$h_1 \geq_g h_2$ iff

$$\forall x \in X \left[h_2(x) = 1 \rightarrow h_1(x) = 1 \right]$$



$h_1 = \langle Sunny, ?, ?, Strong, ?, ? \rangle$

$h_2 = \langle Sunny, ?, ?, ?, ?, ? \rangle$ } $h_2 \geq_g h_1$

$h_3 = \langle Sunny, ?, ?, ?, Cool, ? \rangle$ } $h_2 \geq_g h_3$

Find-S Algorithm (Finding a maximally specific hypothesis)

1. Initialize h to the most specific hypothesis in H
2. For each positive instance x
   - For each attribute constraint a_i in h
     If the constraint a_i is satisfied by x
     do nothing
     else replace a_i in h by the next more general constraint satisfied by x
3. Output hypothesis h

Find-S on EnjoySport Training data

* consider only positive instances

(1) $h_1 = x_1$

$h_1 = \langle$ Sunny, Warm, Normal, Strong, Warm, Same $\rangle$

(2) Compare h_1 with x_2 to check if h_1 is satisfied by x_2

$h_1 = \langle$ Sunny ✓, Warm ✓, Normal ✗, Strong ✓, Warm ✓, Same ✓ $\rangle$

$x_2 = \langle$ Sunny, Warm, High, Strong, Warm, Same $\rangle$

Update hypothesis:

$h_1 = \langle$ Sunny, Warm, ?, Strong, Warm, Same $\rangle$

③ ~~Compare~~ x_3 is a negative example ⇒ ignore

④ Compare h_1 with x_4 to check if h_1 is satisfied by x_4

$h_1 = \langle$Sunny, Warm, ?, Strong, Warm, Same$\rangle$

$x_4 = \langle$Sunny, Warm, High, Strong, Cool, change$\rangle$

Final hypothesis (∵ all training examples are now parsed)

$\langle$Sunny, Warm, ?, Strong, ?, ?$\rangle$

So, by Find-S algorithm, our concept is:

Given that today is sunny, air temperature is warm and Wind speed is strong, Aldo is going to Enjoy playing water sports (EnjoySport = yes)