National University of Singapore
School of Computing
CS3244 Machine Learning

Tutorial 2: Concept Learning

Issue: February 3, 2022 Due: February 7, 2022

Important Instructions:

- *Your solutions for this tutorial must be TYPE-WRITTEN.*
- *Make TWO copies of your solutions: one for you and one to be SUBMITTED TO THE TUTOR IN CLASS. Your submission in your respective tutorial class will be used to indicate your CLASS ATTENDANCE. Late submission will NOT be entertained.*
- *Indicate your NAME, STUDENT NUMBER, and TUTORIAL GROUP in your submitted solution.*
- *YOUR SOLUTIONS TO QUESTIONS* **BL 2**, **BL 3(b)** *will be GRADED for this tutorial.*
- *You may discuss the content of the questions with your classmates. But everyone should work out and write up ALL the solutions by yourself.*

BL 2 Sean Tay has constructed a dataset of binary-valued input attributes and binary-valued output/target concept with 4 noise-free training examples: $D = \{\langle\langle 0,0\rangle,1\rangle, \langle\langle 0,1\rangle,0\rangle, \langle\langle 1,0\rangle,1\rangle, \langle\langle 1,1\rangle,1\rangle\}$.

(a) State the hypothesis h_4 that the FIND-S algorithm would output after learning from all 4 training examples.

(b) Is hypothesis h_4 consistent with D? Explain briefly and concisely your answer.

(c) Does your answer to (b) contradict Proposition 2 on page 15 of the "Concept Learning" lecture slides? Explain briefly and concisely your answer.

Solution.

(a) $h_4 = \langle ?, ? \rangle$.

(b) h_4 is not consistent with D since $h_4(\langle 0,1\rangle) = 1 \neq c(\langle 0,1\rangle) = 0$.

(c) No, the answer to (b) does not contradict Proposition 2. The true target concept c (i.e., $c(x)$ has a value of 0 if $x = \langle 0,1\rangle$, and a value of 1 otherwise) is not a member of the hypothesis space H (i.e., $c \notin H$) since c cannot be specified using the hypothesis representation of a conjunction of constraints on the input attributes, as described on page 5 of the "Concept Learning" lecture slides.

BL 3 In the second lecture of CS3244, Bryan was huffing and puffing over page 25 of the "Concept Learning" lecture slides, during which he was explaining why, in the CANDIDATE-ELIMINATION algorithm, each minimal generalization h of s is not just consistent with the $+$ve training example d but also consistent with all $+$ve and $-$ve training examples observed thus far. Can you recall what he was babbling about at that time?

(a) Prove that h is consistent with all $+$ve training examples observed thus far, not including d.
Hints: Consider using Proposition 1 and the $\geq_g$ relation between h and s.

(b) Prove that h is consistent with all $-$ve training examples observed thus far.
Hints: Consider using Proposition 1 and the $\geq_g$ relation expressed in the line of pseudocode on page 25: "some member of G is more general than or equal to h".

You may assume that every member of the specific boundary and general boundary is consistent with all training examples observed thus far.

Solution.

(a)
1. Since $s \in S$, s is consistent with all $+$ve training examples observed thus far (not including d), by the given assumption on S.
2. Therefore, every $+$ve training instance (not including that of d) satisfies s, by Proposition 1.
3. Since $h \geq_g s$, every $+$ve training instance (not including that of d) also satisfies h.
4. Therefore, h is consistent with all $+$ve training examples observed thus far (not including d), by Proposition 1.

(b)
1. We know that $\exists g \in G \;\; g \geq_g h$.
2. Since $g \in G$, g is consistent with all $-$ve training examples observed thus far, by the given assumption on G.
3. Therefore, every $-$ve training instance does not satisfy g, by Proposition 1.
4. Since $g \geq_g h$, every $-$ve training instance also does not satisfy h.
5. Therefore, h is consistent with all $-$ve training examples observed thus far, by Proposition 1.

TM 2.3 Consider the $EnjoySport$ learning task and the hypothesis space H described on page 7 of the "Concept Learning" lecture slides. Let us define a new hypothesis space H' that consists of all pairwise disjunctions of the hypotheses in H. For example, a typical hypothesis in H' is

$$\langle ?, Cold, High, ?, ?, ? \rangle \vee \langle Sunny, ?, High, ?, ?, Same \rangle \; .$$

Trace the CANDIDATE-ELIMINATION algorithm for the hypothesis space H' given the sequence of training examples from the table on page 4 of the "Concept Learning" lecture slides (i.e., show the sequence of S and G boundary sets).

Solution.

$$
\begin{aligned}
G_0 &= \{\langle ?,?,?,?,?,? \rangle \vee \langle ?,?,?,?,?,? \rangle\} = \{\langle ?,?,?,?,?,? \rangle\} \\
S_0 &= \{\langle \emptyset,\emptyset,\emptyset,\emptyset,\emptyset,\emptyset \rangle \vee \langle \emptyset,\emptyset,\emptyset,\emptyset,\emptyset,\emptyset \rangle\} = \{\langle \emptyset,\emptyset,\emptyset,\emptyset,\emptyset,\emptyset \rangle\} \\
G_1 &= G_0 \\
S_1 &= \{\langle Sunny, Warm, Normal, Strong, Warm, Same \rangle \vee \langle \emptyset,\emptyset,\emptyset,\emptyset,\emptyset,\emptyset \rangle\} \\
&= \{\langle Sunny, Warm, Normal, Strong, Warm, Same \rangle \vee \\
&\quad \langle Sunny, Warm, Normal, Strong, Warm, Same \rangle\} \\
&= \{\langle Sunny, Warm, Normal, Strong, Warm, Same \rangle\} \\
G_2 &= G_1 \\
S_2 &= \{\langle Sunny, Warm, ?, Strong, Warm, Same \rangle\} \\
S_3 &= S_2 \\
G_3 &= \{\langle Sunny,?,?,?,?,? \rangle \vee \langle Cloudy,?,?,?,?,? \rangle, \\
&\quad \langle Sunny,?,?,?,?,? \rangle \vee \langle ?,Warm,?,?,?,? \rangle, \\
&\quad \langle Sunny,?,?,?,?,? \rangle \vee \langle ?,?,Normal,?,?,? \rangle, \\
&\quad \langle Sunny,?,?,?,?,? \rangle \vee \langle ?,?,?,Weak,?,? \rangle, \\
&\quad \langle Sunny,?,?,?,?,? \rangle \vee \langle ?,?,?,?,Cool,? \rangle, \\
&\quad \langle Sunny,?,?,?,?,? \rangle \vee \langle ?,?,?,?,?,Same \rangle, \\
&\quad \langle Cloudy,?,?,?,?,? \rangle \vee \langle ?,Warm,?,?,?,? \rangle, \\
&\quad \langle Cloudy,?,?,?,?,? \rangle \vee \langle ?,?,?,?,?,Same \rangle, \\
&\quad \langle ?,Warm,?,?,?,? \rangle \vee \langle ?,?,Normal,?,?,? \rangle, \\
&\quad \langle ?,Warm,?,?,?,? \rangle \vee \langle ?,?,?,Weak,?,? \rangle, \\
&\quad \langle ?,Warm,?,?,?,? \rangle \vee \langle ?,?,?,?,Cool,? \rangle, \\
&\quad \langle ?,Warm,?,?,?,? \rangle \vee \langle ?,?,?,?,?,Same \rangle, \\
&\quad \langle ?,?,Normal,?,?,? \rangle \vee \langle ?,?,?,?,?,Same \rangle, \\
&\quad \langle ?,?,?,Weak,?,? \rangle \vee \langle ?,?,?,?,?,Same \rangle, \\
&\quad \langle ?,?,?,?,Cool,? \rangle \vee \langle ?,?,?,?,?,Same \rangle\} \\
G_4 &= \{\langle Sunny,?,?,?,?,? \rangle \vee \langle Cloudy,?,?,?,?,? \rangle, \\
&\quad \langle Sunny,?,?,?,?,? \rangle \vee \langle ?,Warm,?,?,?,? \rangle, \\
&\quad \langle Sunny,?,?,?,?,? \rangle \vee \langle ?,?,Normal,?,?,? \rangle, \\
&\quad \langle Sunny,?,?,?,?,? \rangle \vee \langle ?,?,?,Weak,?,? \rangle, \\
&\quad \langle Sunny,?,?,?,?,? \rangle \vee \langle ?,?,?,?,Cool,? \rangle, \\
&\quad \langle Sunny,?,?,?,?,? \rangle \vee \langle ?,?,?,?,?,Same \rangle, \\
&\quad \langle Cloudy,?,?,?,?,? \rangle \vee \langle ?,Warm,?,?,?,? \rangle, \\
&\quad \langle ?,Warm,?,?,?,? \rangle \vee \langle ?,?,Normal,?,?,? \rangle, \\
&\quad \langle ?,Warm,?,?,?,? \rangle \vee \langle ?,?,?,Weak,?,? \rangle, \\
&\quad \langle ?,Warm,?,?,?,? \rangle \vee \langle ?,?,?,?,Cool,? \rangle, \\
&\quad \langle ?,Warm,?,?,?,? \rangle \vee \langle ?,?,?,?,?,Same \rangle, \\
&\quad \langle ?,?,?,?,Cool,? \rangle \vee \langle ?,?,?,?,?,Same \rangle\} \\
S_4 &= \{\langle Sunny, Warm, ?, Strong, Warm, Same \rangle \vee \langle Sunny, Warm, High, Strong, Cool, Change \rangle\}
\end{aligned}
$$