

# J McLeod Homework 5

October 9, 2019

## Jericho McLeodCSI-873Homework 5

**Exercise 6.3, Page 198** Consider the concept learning algorithm FindG, which outputs a maximally general consistent hypothesis (e.g., some maximally general member of the version space).  
a) Give a distribution for  $P(h)$  and  $P(D|h)$  under which FindG is guaranteed to output a MAP hypothesis. b) Give a distribution for  $P(h)$  and  $P(D|h)$  under which FindG is not guaranteed to output a MAP hypothesis. c) Give a distribution for  $P(h)$  and  $P(D|h)$  under which FindG is guaranteed to output a ML hypothesis but not a MAP hypothesis.

Solution a) Any distribution that contains no errors (noise), or  $P(D|h) = 1$  if  $D$  and  $h$  are consistent, and 0 otherwise, and either contains a uniform probability distribution over  $H$  or has distribution  $H$  where  $P(h_1) \geq P(h_2)$  if  $h_1$  is more specific than  $h_2$ , will output a MAP hypothesis. For example:

Attribute 1	Attribute 2	Outcome
+	-	+
-	+	-

Solution b) Inconsistent data will cause a FindG to not output a MAP hypothesis.

Attribute 1	Attribute 2	Outcome
+	-	+
-	+	+

Solution c) A dataset that contains noise can output a Maximum Likelihood Estimator, but not a Maximum A Posteriori hypothesis. As an example, consider the noisy dataset:

x	f(x)
1.2	2.1
1.9	4
0.9	2.2