Jericho McLeod
CSI-873
Homework 5

Exercise 6.3, Page 198

Consider the concept learning algorithm FindG, which outputs a maximally general consistent hypotehsis (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.

MAP Hypothesis: is an estimate of an unknown quantity, that equals the mode of the posterior distribution; the most probable hypothesis given the data.

$$h_{MAP} \equiv argmax_{h \in H} P(h|D)$$

Using Bayes Theorem:

$$h_{MAP} \equiv argmax_{h \in H} P(D|h)P(h)$$

which drops the denominator value $P(D)$ as it is a constant that is independent of h.

ML Hypothesis: assume every hypothesis is equally probable a priori, thus the $P(h)$ term need not be considered, giving:

$$h_{ML} \equiv argmax_{h \in H} P(D|h)$$

Solution a) Any prior probability distribution which favors more specific hypotheses will return a MAP hypothesis in FindG.

Solution b) Any prior probability distribution that does not favor more specific hypotheses will not return a MAP hypothesis in FindG.

Solution c) Any prior probability that is uniform will return a maximum likelihood estimator for FindG, but not a maximum a priori hypothesis for FindG.