Chapter 9

Bayesian approach and MDL

9.1 Bayesian induction

$$P(H_i|E) = \frac{P(H_i)P(E|H_i)}{\sum_{i=1}^{n} P(H_i)P(E|H_i)}$$

9.2 Occams razor

$E^+ = \{0, 000, 00000, 000000000\}$,
$E^- = \{\varepsilon, 00, 0000, 000000\}$.

$G_1 : S \to 0|000|00000|000000000$,
$G_2 : S \to 00S|0$,

9.3 Minimum Description Length (MDL) principle

$$-\log_2 P(H_i|E) = -\log_2 P(H_i) - \log_2 P(E|H_i) + C,$$

$$L(H|E) = L(H) + L(E|H),$$

$$L(E) > L(H) + L(E|H).$$

9.4 Evaluating propositional hypotheses

$$L(R_i) = -\log_2 \frac{1}{\binom{ts}{k_i}} = \log_2 \binom{ts}{k_i}.$$

9.4.1 Encoding exceptions

$$L(E|H) = \log_2 \binom{tp + fp}{fp} + \log_2 \binom{tn + fn}{fn}.$$

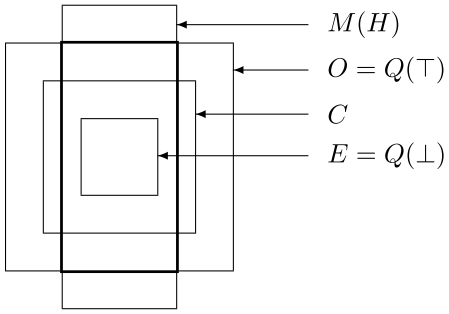

Figure 9.1:

9.4.2 Encoding entropy

$$e_i = -\frac{p_i}{n_i} * \log_2 \frac{p_i}{n_i} - \frac{n_i - p_i}{n_i} * \log_2 \frac{n_i - p_i}{n_i},$$

9.5 Evaluating relational hyporheses

9.5.1 Complexity of logic programs

$$L_{PC}(E|H) = \sum_{A \in E} L_{PC}(A|H).$$

$$L_{PC}(E|\top) = \sum_{A \in E} \log_2 c^n = |E| * n * \log_2 c,$$

$$L_{PC}(E|\bot) = \sum_{A \in E} L_{PC}(A|\bot) = |E| * \log_2 |E|.$$

9.5.2 Learning from positive only examples