1. *Reasoning Under Uncertainty using Bayes' Theorem*

problem: the patient shows the symptoms $S_1, ..., S_m$ and no other symptoms — what is the probability that the patient has the disease D_i?

available data:

$P(D_i)$:= the a priori probability that the patient sufferes from disease D_i; the probability that the patient has the disease *before* any symptoms have been observed.

$\lambda_{ij} = \frac{P(S_j|D_i)}{P(S_j)}$:= Estimates the relationship between the occurrence of the symptom S_j and the occurrence of the disease D_i. For example: $\lambda_{ij} = 8$ expresses that the symptom occurs 8 times more frequent if the patient has the disease (in this case the symptom provides some positive evidence for the disease). On the other hand, $\lambda_{ij} = 0.125$ expresses that the symptom occurs 8 times less frequent together with the disease(in this case the symptom provides some negative evidence for the occurrence of the disease). Finally, $\lambda_{ij} \approx 1$ expresses that the symptom S_j does not provide any evidence at all for the occurrence of the disease D_i.

solution: Under the conditional independence assumption (concerning the symptoms involved in the reasoning process and concerning the symptoms assuming the disease is present) the probability of having the disease D_i when showing the symptoms $S_1, ..., S_m$ can be calculated as follows:

$$P(D_i|S_1, ..., S_m) = P(D_i) \times \prod_{k=1}^{m} \lambda_{ik}$$

Another formulation that facilitates calculations is the following:

$$log_2(P(D_i|S_1, ..., S_m)) = log_2(P(D_i)) + \sum_{k=1}^{m} log_2(\lambda_{ik})$$

remark: in most systems λ_{ij} is called *the "new evidence multiplier"* — warning: do not mix up with $\frac{P(S_j|D_i)}{P(S_j|\neg D_i)}$ the *odds-multiplier* for rules in PROSPECTOR.

3. The Conditional Independence Assumption

Applying Bayes' theorem the a posterori probability of a disease D_j can be computed as follows:

$$P(D_j|S_1 \wedge ...S_n) = \frac{P(D_j) * P(S_1 \wedge ... \wedge S_n|D_j)}{P(S_1 \wedge ... \wedge S_n)}$$

Assuming that we have a diagnostic problem that involves 50 diseases and 500 symptoms, more than

$50 * 2^{500} \approx 10^{150}$

conditional probabilities are needed for a dignostic expert system of this size. However, frequently, the following simplified formula is used instead, which relies on the *conditional independence assumption*:

$$P(D_j|S_1 \wedge ...S_n) \approx \frac{P(D_j) * P(S_1|D_j) * ... * P(S_n|D_j)}{P(S_1) * ... * P(S_n)}$$

With the second formula only appoximately 25000 conditional probabilities are needed; it sacrifices precision in order to reduce the enourmous knowledge acquisition costs in diagnostic expert systems. In general, if the conditional independence assumption is not valid, computation errors will occur, especially if small probabilities are involved.

Final Remarks:

- *In summary, we frequently have to tolerate imprecision by making assumptions (such as the conditional independence assumption) in the design of a system in order to make it feasible to get such a system running, only spending a limited amount of time and money.*
- Bayesian approaches that rely on the conditional independence assumption are nowadays called *naive Bayesian approaches* in the literature. More complicated approaches, such as *Bayesian networks* (that were introduced by Judea Pearl) and *influence networks*, that do not rely on the condition independence assumption have been introduced in the last decade in the literature.