Vera Koponen
Department of Mathematics
Uppsala University

Exercises

All references to sections, propositions etc refer to *Notes on logic and probability.*

1. Let $\mathbf{P} = \{p, q, r, s\}$ be a set of propositional variables and let

$$\mathbf{E} = \{(1,1,0,1),(1,0,1,0),(0,1,1,0),(0,1,0,1),(1,0,1,1)\},$$
$$\mathbf{N} = \{(0,0,0,0),(1,1,1,1),(1,1,1,0),(0,1,0,0),(1,0,0,1)\}.$$

Recall that for $i, j, k, l \in \{0, 1\}$, the tuple (i, j, k, l) represents the truth assignment which assigns p the value true if $i = 1$ and false otherwise, q the value true if $j = 1$ and false otherwise, and so on.

(a) Construct a maximally general CNF φ such that φ covers $\mathbf{E}$ and avoids $\mathbf{N}$. Use the ideas from Proposition 3.2.6 and Algorithm 1 in Section 3.3.

(b) Modify your CNF φ to a CNF ψ which is as specific as you can make it, but ψ must still cover $\mathbf{E}$ and avoid $\mathbf{N}$. The idea is that some clauses of φ can perhaps be simplified and some clauses can perhaps even be removed. See Remark 3.3.2.

(c) Suppose that μ is a probability distribution on $\{0,1\}^4$ such that $\mu(\mathbf{E}) \geq 0.65$ and $\mu(\mathbf{N}) \geq 0.1$. Give upper and lower bounds (using only this information) for the probabilities $\mu(\varphi)$ and $\mu(\psi)$.

(d) Assuming that $\mu(i, j, k, l) > 0$ for every $(i, j, k, l) \in \{0,1\}^4$, do we have $\mu(\varphi) > \mu(\psi)$, $\mu(\varphi) < \mu(\psi)$ or $\mu(\varphi) = \mu(\psi)$?

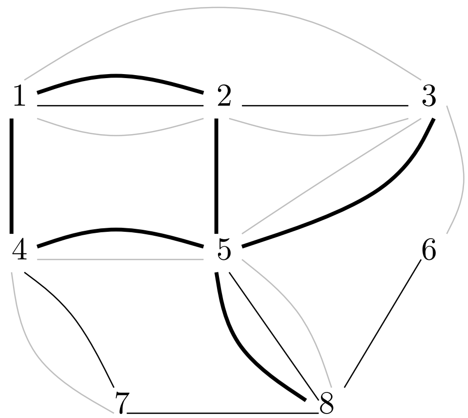
(e) For each of the following sets, use ψ as a logic program to determine which of p, q, r, s, $\neg p, \neg q, \neg r, \neg s$ is a consequence of the set (if a set is inconsistent say so):

$$\{\psi, p, q\}, \quad \{\psi, p, \neg q\}, \quad \{\psi, \neg p, \neg q\}, \quad \{\psi, p, r\}.$$

2. Suppose that the first-order langauge L has three 2-ary (binary) relation symbols called P, Q and R. Consider the L-structure $\mathcal{A}$ with domain $A = \{1,2,3,4,5,6,7,8\}$ where the interpretations of P, Q and R are illustrated by the edges/lines in the diagram below, where black narrow edges represent $P^{\mathcal{A}}$ (the interpretation of P in $\mathcal{A}$), grey edges represent $Q^{\mathcal{A}}$ and black wide edges represent $R^{\mathcal{A}}$. (We can imagine that the members of A are objects in a database and that $P^{\mathcal{A}}, Q^{\mathcal{A}}$ and $R^{\mathcal{A}}$ are relationships that pairs of objects in the database may satisfy or not.) For example, since there is a black narrow edge between 1 and 2 it means that $(1,2),(2,1) \in P^{\mathcal{A}}$. All three relation symbols are interpreted as symmetric relations (otherwise I would have used arrows instead of edges/lines). Let μ be the uniform probability distribution on A (so μ gives all members of A the same probability).

Construct, using the algorithm in Section 5.1, a Bayesian network for the formulas $P(x, y), Q(x, y)$ and $R(x, y)$ viewed as binary random variables. Do it by treating the binary random variables in the order $P(x, y), Q(x, y)$ and $R(x, y)$ (then we can compare our results). Moreover, in this exercise we consider two random variables $X : A^k \to \{0,1\}$ and $Y : A^k \to \{0,1\}$ to be independent (or independent over a third random variable

$Z : A^k \to \{0,1\}$) if for all $i \in \{0,1\}$ (and $j \in \{0,1\}$), $\left|\mu^k(X=1 \mid Y=i) - \mu^k(X=1)\right| \le 0.1$ ($\left|\mu^k(X=1 \mid Y=i, Z=j) - \mu^k(X=1 \mid Z=j)\right| \le 0.1$) whenever both terms are defined.



Partial answers/solutions

1. (a) To get a maximally general CNF, construct a clause for each truth assignment in **N** in the way explained (for example) on the lectures (or see the notes). Then one gets

$$(p \vee q \vee r \vee s) \wedge (\neg p \vee \neg q \vee \neg r \vee \neg s) \wedge (\neg p \vee \neg q \vee \neg r \vee s) \wedge (p \vee \neg q \vee r \vee s) \wedge (\neg p \vee q \vee r \vee \neg s).$$

(b) The clauses can be simplified, for example, in the following way (still retaining a formula which covers **E** and avoids **N**):

$$(p \vee q) \wedge (\neg p \vee \neg q \vee \neg r) \wedge (\neg p \vee \neg q \vee \neg r \vee s) \wedge (r \vee s) \wedge (q \vee r).$$

(c) Let φ and ψ denote the CNF's from (a) and (b), respectively. Since every truth assignment which satisfies ψ also satisfies φ true (in symbols $\psi \models \varphi$) it follows $\mu(\psi) \leq \mu(\varphi)$. As all truth assignments in **E** satisfy ψ it follows that $0.65 \leq \mu(\mathbf{E}) \leq \mu(\psi)$. Let **X** be the set of truth assignments which satisfy φ. Since every truth assignment in **N** does *not* satisfy φ it follows that $\mathbf{X} \subseteq \{0,1\}^4 \setminus \mathbf{N}$ and hence $\mu(\varphi) = \mu(\mathbf{X}) \leq \mu(\{0,1\}^4 \setminus \mathbf{N}) = 1 - \mu(\mathbf{N}) \leq 1 - 0.1 = 0.9$. Hence, $0.65 \leq \mu(\mathbf{E}) \leq \mu(\psi) \leq 0.9$.

(d) We have $\mu(\psi) \leq \mu(\varphi)$. There is a truth assignment which satisfies φ but not ψ. (Exercise to find one.) The assumption that every truth assignment has positive probability now implies that $\mu(\psi) < \mu(\varphi)$.

(e)

$\{\psi, p, q\} \models \neg r, s$,
$\{\psi, p, \neg q\} \models r$ (we get no information about s),
$\{\psi, \neg p, \neg q\}$ The set is inconsistent, that is, no truth assignment satisfies all formulas in it.
$\{\psi, p, r\} \models \neg q, s$.

2. The DAG associated with the Bayesian network (given the instructions of the assignment) is:

$$P(x,y) \longrightarrow Q(x,y) \longrightarrow R(x,y)$$

(Of course the complete Bayesian network also needs to associate (conditional) probabilities to each formula/random variable.)