# Final Exam *Algorithmical Modelling* Spring 2005 — Solutions

1. (5 points) "In 2020, 90 % of the humans on earth will have internet access." Formalize the message of this sentence by describing in words a suitable event space $\Omega$ and a measure space $E$, by specifying a random variable $X$, and finally, by writing down a formula that represents the contents of that sentence. (Altogether two lines).

**Solution.** One possibility:
$\Omega$: The set of humans on earth in 2020. $E$: $\{0,1\}$. $X$: $\Omega \to E$, $X(\omega) = 1$ iff $\omega$ has internet access. The formula: $P(X = 1) = 0.9$.

2. (25 points) Show that for two random variables $X$, $Y$, each with values in $\{0,1\}$, uncorrelatedness implies independence. It may happen that this proof leads you to a four-case distinction. Treat only one of the cases; the other are similar, don't bother about them.

**Solution.**
Let $P(X = 0, Y = 0) = a$, $P(X = 0, Y = 1) = b$, $P(X = 1, Y = 0) = c$, $P(X = 1, Y = 1) = d$. Uncorrelatedness means $E(X\,Y) = E(X)\,E(Y)$, that is $d = (c + d)(b + d)$. Using that $a + b + c + d = 1$, a few elementary transformations reveal that this is equivalent to $ad = bc$.

Now consider independence. This means that $P(X = x, Y = y) = P(X = x)\,P(Y = y)$, where $x, y = 0, 1$. We consider the case $P(X = 0, Y = 0) = P(X = 0)\,P(Y = 0)$. This is equivalent to $a = (a + b)(a + c)$. The same transformations as used for uncorrelatedness show that this is equivalent to $ad = bc$, which is just uncorrelatedness. The other three choices for $x$, $y$ run similarly.

3. (20 points) How might one use sampling to obtain approximate values of integrals of the kind $\int_{\Re^n} f(\mathbf{x})d\mathbf{x}$, where $f$ is non-negative and the integral has a finite value? Assume that a function $g(\mathbf{x}) \geq f(\mathbf{x})$ with known integral value $C = \int_{\Re^n} g(\mathbf{x})d\mathbf{x}$ is provided, and that there is a way known of how to sample from the pdf $g/C$, giving you a sequence of samples $S = \mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N$ distributed according to $g/C$. Describe in pseudocode how $S$ can be used to obtain an estimate $E$ of $\int_{\Re^n} f(\mathbf{x})d\mathbf{x}$.

**Solution.**

```
accumulator = 0;
for n = 1 to N
        accumulator = accumulator + f(x_n) / g(x_n);
end
E = C * accumulator / N.
```

4. (30 points) A study in modelling population dynamics. Assume that Australia is covered by a square grid with grid cells the size of a rabbit. At a given time $n$, each grid cell $c_{ij}$ is either inhabited by a rabbit or not. Rabbits like to have friends, so they tend to flock.

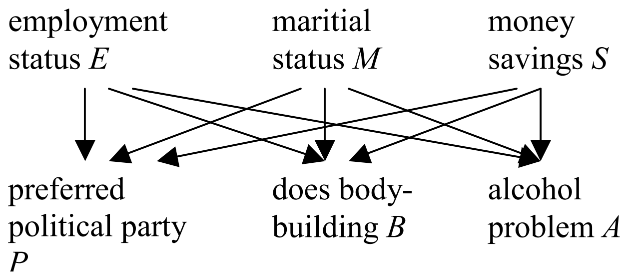
However, if flocks become too large, they devastate the land around them such that the rabbits of the flock are forced to disperse, seeking new feeding grounds. Model this (in an admittedly simplistic way) by microstates (specify what a microstate is), provide a reasonable energy function, try an intuitive explanation of what socio-psychological phenomenon a "temperature" parameter might describe (take this part of the assignment with a lighthearted smile, you will not be judged for bio-socio-psychological veracity). Describe in words a biologically plausible proposal distribution of an MCMC sampler, which should respect biological constraints (rabbits don't materialize out of thin air but keep living as individuals most of the time...). Describe in words the microstate distribution at low temperature.

**Solution.** A microstate is a function $s$: $C \rightarrow \{0,1\}$, where $C$ is the set of grid cells of Australia, with a value of 1 indicating that the cell is inhabited. A plausible energy function would have to account for the two conflicting tendencies of flock-building by two terms, for instance

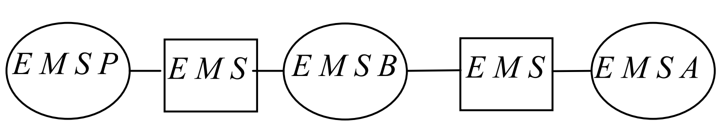
$$E(s) = -\sum_{\substack{|i-i'|<100 \\ |j-j'|<100}} F\, s(c_{ij})s(c_{i'j'}) + \sum_{\substack{1000<|i-i'|<10000 \\ 1000<|j-j'|<10000}} D\, s(c_{ij})s(c_{i'j'})$$

where the first term models the sociability flocking desire and the second the overcrowding dispersion tendency. At high temperatures the spatial rabbit distribution becomes on average more and more "white noise". Socio-psycho-dynamically, this means that the rabbits lose their social habits (social flocking becomes unimportant) and that there is an abundance of food (no necessity to prevent overgrazing) [think about this connection: are social habits connected to scarcity of living conditions?]. A simple proposal distribution would only have to account for a spatial move of a single rabbit into an adjacent, unoccupied cell. This could be enriched by allowing an infrequent duplication or death of a rabbit. At low temperatures, we would expect a large variety of microstates (due to the system's frustrated nature), each of which is typically well-organized into flocks separated by empty space. The size of flocks and empty space would depend on $F$ and $D$.

5. (15 points) Give a join tree for the BN from Fig 5.10 of the script, rendered again here for convenience:




**Solution.**




6. (15 points) Show that when the Gödel implication is used, the tautology $(\psi \wedge \neg\psi) \rightarrow \varphi$ always evaluates to 1 even in FL. Concretely, show that $\mu_{(A \cap A^c) \rightarrow B}(x, y) \equiv 1$ in the case where `A` and `B` share the same universe of discourse, and if you use the min operator for

the "fuzzy and", the max operator for the "fuzzy or", and the Gödel implication for the "→". Note: the membership function of the binary fuzzy relation ≤ that figures in the Gödel implication is given by

$$\mu_{A \le B}(x, y) = \begin{cases} 1, \text{if } x \le y \\ 0, \text{else} \end{cases}$$

**Solution.** Let $\mu_A(x), \mu_B(y)$ be two unary membership functions. Then

$$\begin{aligned} \mu_{(A \cap A^c) \to B}(x, y) &= \mu_{((A \cap A^c) \le B) \cup B}(x, y) = \max(\mu_{(A \cap A^c) \le B}(x, y), \mu_B(y)) \\ &= \max([\min(\mu_A(x), \mu_{A^c}(x)) \le \mu_B(y)], \mu_B(y)) = \\ &= \max([\min(\mu_A(x), 1 - \mu_A(x)) \le \mu_B(y)], \mu_B(y)) = \max(1, \mu_B(y)) \\ &= 1 \end{aligned}$$

for any *x*, *y*. The [] operator in this equation denotes the Boolean 0-1 truth value of its argment.