TIF 150: Information theory for complex systems

Time: March 8, 2012, 14.00 – 18.00.
Allowed material: Calculator (type approved according to Chalmers rules).
Teacher and examiner: Kristian Lindgren (772 3131)

All answers and derivations must be clear and well motivated.
Grade limits: 25p for 3, 34p for 4, and 42p for 5.
ECTS grades: 25p for E, 28p for D, 34p for C, 38p for B, 42p for A.
The results will be available on March 22.

1. Ball measurements.
Assume you have 3 balls, the first one (with an "A" etched into it) weighs between 10 g and 12 g, the second (with a "B" etched into it) between 10 g and 11g, and the third one (with a "C" etched into it) between 11 g and 12 g. You have at your disposal a balance measurement device.

(a) According to information theory, what should the a priori probability distributions of the individual weights for the three balls be?
(b) How much information do you gain if you weigh A against B and find that B is heavier?
(c) How much is the total expected gain of information if you weigh all pairs of balls, i.e., three measurements, using the balance.

(10 p)

2. An equilibrium spin system. Consider an infinite one-dimensional lattice system where each site can be in one of 4 states (A, B, C, D).

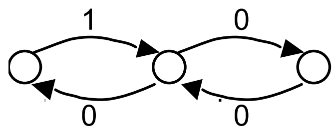
. . .	A	C	B	A	C	D	C	A	C	B	D	C	B	. . .

Neighbouring pairs of the same symbol are forbidden (e.g., AA etc). There is a local contribution to the energy, being $-J < 0$ when two neighbouring states are AC, CA, BD, and DB, and there is no interaction energy for the remaining neighbouring states. If the average energy is u, what is the equilibrium distribution? Or, in other words, what are the probabilities over sequences of symbols that characterize the system? (You should give the answer as a function of temperature instead of energy.) What is the entropy in the limit of zero temperature?

(10 p)

3. CA information.

Consider a one-dimensional cellular automaton given by elementary rule R50, i.e., neighbourhoods 100, 101and 001 map to 1, but all others to 0. Let the initial state be characterized by the following finite state automaton (where it is assumed that if two arcs leave the same node they have equal probabilities):



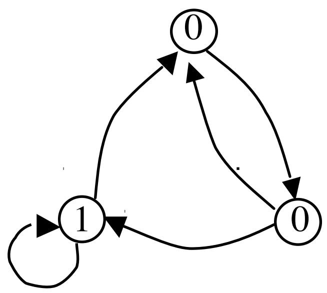
(a) What is the initial entropy (at $t = 0$)?

(b) Derive the finite state automaton that characterizes the CA state after one time step ($t = 1$). Describe why the entropy has decreased, and how this can be understood from the automaton? (You need not to calculate the entropy.)

(c) What does the entropy converge to in the infinite time limit?

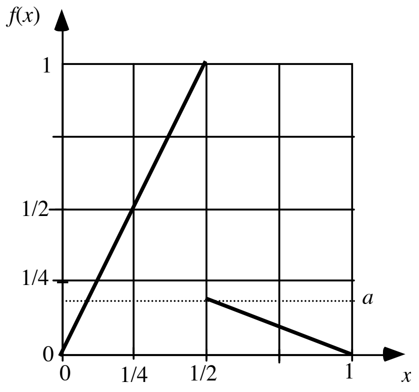
(10 p)

4. Correlation complexity. Determine the correlation complexity η for the process described by the hidden Markov model below. You need to argue in detail for the approach you use. (When two arcs leave a node it is assumed that they have the same probability.)

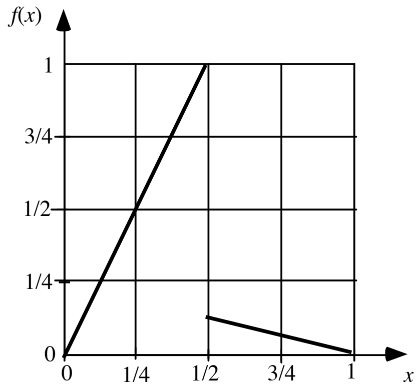


(8 p)

5. Chaos and information. Let a piecewise linear map $f(x)$ be defined by the figure below, where $0 < a < 1$, and where the mapping is determined by $f(0) = f(1) = 0, f(1/2^{+}) = a$ and $f(1/2^{-}) = 1$.



(i)



(ii)

Consider the dynamical system

$$x_{t+1} = f(x_t) .$$

(i) Show that the system is chaotic for all values of a?

(ii) Suppose now that $a = 1/8$. Determine the invariant measure that characterizes the chaotic behaviour, and calculate the Lyapunov exponent λ. Find a partition that is generating, and calculate the measure entropy from the symbolic dynamics.

(12 p)

Information theory for complex systems – useful equations

Basic quantities

$$I(p) = \log\frac{1}{p} \qquad S[P] = \sum_{i=1}^{n} p_i \log\frac{1}{p_i} \qquad K\left[P^{(0)};P\right] = \sum_{i=1}^{n} p_i \log\frac{p_i}{p_i^{(0)}}$$

Max entropy formalism (with $k + 1$ constraints) using the Lagrangian L

$$L\left(p_1,\ldots,p_n,\lambda_1,\ldots,\lambda_r,\mu\right) = S[P] + \sum_{k=1}^{r} \lambda_k \left(F_k - \sum_{i=1}^{n} p_i f_k(i)\right) + \left(\mu - 1\right)\left(1 - \sum_{i=1}^{n} p_i\right),$$

$$p_j = \exp\left(-\mu - \sum_{k=1}^{r} \lambda_k f_k(j)\right), \quad \mu(\lambda) = \ln\sum_{j=1}^{n} \exp(-\sum_{k} \lambda_k f_k(j)), \quad \frac{\partial\mu(\lambda)}{\partial\lambda_k} = -F_k$$

Symbol sequences

$$s = \lim_{n\to\infty} \sum_{x_1\ldots x_{n-1}} p(x_1\ldots x_{n-1}) \sum_{x_n} p(x_n \mid x_1\ldots x_{n-1}) \log\frac{1}{p(x_n \mid x_1\ldots x_{n-1})} =$$

$$= \lim_{n\to\infty}\left(S_n - S_{n-1}\right) = \Delta S_\infty = \lim_{n\to\infty}\frac{1}{n} S_n$$

$$k_1 = \log\nu - S_1, \qquad k_n = -S_n + 2S_{n-1} - S_{n-2} = -\Delta S_n + \Delta S_{n-1} = -\Delta^2 S_n \quad (n = 2,3,\ldots)$$

$$S_{\text{tot}} = \log\nu = \sum_{m=1}^{\infty} k_m + s$$

$$\eta = \sum_{m=1}^{\infty} (m-1)\, k_m = \lim_{m\to\infty}\lim_{n\to\infty} \sum_{\sigma_m} p(\sigma_m) \sum_{\tau_n} p(\tau_n \mid \sigma_m) \log\frac{p(\tau_n \mid \sigma_m)}{p(\tau_n)} = \lim_{m\to\infty}(S_m - m\,s)$$

Geometric information theory

$$p(r;x) = \frac{1}{\sqrt{2\pi}\, r} \int_{-\infty}^{\infty} dw\, e^{-w^2/2r^2}\, p(x - w) \quad , \qquad \left(-r\frac{\partial}{\partial r} + r^2 \frac{d^2}{dx^2}\right) p(r;x) = 0$$

$$p_{\text{Gaussian}}(r;x) = \frac{1}{\sqrt{2\pi}\sqrt{b^2 + r^2}} \exp\left(-\frac{x^2}{2(b^2 + r^2)}\right)$$

$$K[p_0;p] = \int dx\, p(x) \ln\frac{p(x)}{p_0(x)} = \int_0^{\infty} \frac{dr}{r} \int dx\, k(r,x), \qquad k(r,x) = r^2 p(r;x)\left(\frac{d}{dx}\ln p(r;x)\right)^2$$

$$d(r) = D_{\text{E}} - r\frac{\partial}{\partial r} \int d\mathbf{x}\, p(r;\mathbf{x}) \ln\frac{1}{p(r;\mathbf{x})}$$

Chemical systems and information flow

$$E = k_{\text{B}} T_0 \frac{N}{V} K, \qquad K = \int_V d\mathbf{x}\, K[c_0;c(\mathbf{x})] = \int_V d\mathbf{x} \sum_{i=1}^{M} c_i(\mathbf{x}) \ln\frac{c_i(\mathbf{x})}{c_{i0}}, \quad \sum_{i=1}^{M} c_i(\mathbf{x},t) = 1$$

$$K = V\sum_{i=1}^{M}\bar{c}_i \ln\frac{\bar{c}_i}{c_{i0}} + \int_V d\mathbf{x}\sum_{i=1}^{M} c_i(\mathbf{x})\ln\frac{c_i(\mathbf{x})}{\bar{c}_i} = K_{\text{chem}} + K_{\text{spatial}}$$

$$K_{\text{spatial}} = \int_0^{\infty}\frac{\partial r}{r}\int d\mathbf{x}\, k(r,\mathbf{x})$$

$$k(r,\mathbf{x}) = r^2\sum_i \tilde{c}_i(r,\mathbf{x})\left(\nabla\ln\frac{\tilde{c}_i(r,\mathbf{x})}{c_{i0}}\right)^2 = r^2\sum_i\frac{(\nabla\tilde{c}_i(r,\mathbf{x}))^2}{\tilde{c}_i(r,\mathbf{x})} = \left(-r\frac{\partial}{\partial r} + r^2\nabla^2\right)\sum_i \tilde{c}_i(r,\mathbf{x})\ln\frac{\tilde{c}_i(r,\mathbf{x})}{c_{i0}}$$

$$\dot{c}_i(\mathbf{x},t) = \frac{d}{dt}c_i(\mathbf{x},t) = D_i\nabla^2 c_i(\mathbf{x},t) + F_i(\mathbf{c}(\mathbf{x},t)) + b_i(c_{i,\text{res}} - c_i(\mathbf{x},t))$$

$$\sigma(\mathbf{x},t) = \sum_i\left(D_i\frac{(\nabla c_i(\mathbf{x},t))^2}{c_i(\mathbf{x},t)} - F_i(\mathbf{c}(\mathbf{x},t))\ln\frac{c_i(\mathbf{x},t)}{c_{i0}}\right)$$

$$j_r(r,\mathbf{x},t) = \sum_i\left(D_i\frac{(\nabla\tilde{c}_i(r,\mathbf{x},t))^2}{\tilde{c}_i(r,\mathbf{x},t)} - \tilde{F}_i(\mathbf{c}(\mathbf{x},t))\ln\frac{\tilde{c}_i(r,\mathbf{x},t)}{c_{i0}}\right),\ \tilde{F}_i(\mathbf{c}(\mathbf{x},t)) = \exp(\tfrac{r^2}{2}\nabla^2)F_i(\mathbf{c}(\mathbf{x},t))$$

$$\mathbf{j}(r,\mathbf{x},t) = -r^2\nabla\sum_i\tilde{F}_i(\mathbf{c}(\mathbf{x},t))\ln\frac{\tilde{c}_i(r,\mathbf{x},t)}{c_{i0}},\quad J(r,x,t) = -\sum_i b_i(\tilde{c}_i + c_{i,\text{res}})[r\nabla\ln\tilde{c}_i]^2$$

$$\dot{k}(r,\mathbf{x},t) = r\frac{\partial}{\partial r}j_r(r,\mathbf{x},t) - \nabla\cdot\mathbf{j}(r,\mathbf{x},t) + J(r,x,t)$$

<u>Chaos and information</u>

$$\int dx\mu(x)\varphi(x) = \lim_{T\to\infty}\frac{1}{T}\sum_{k=0}^{T-1}\varphi(f^k(x(0))),$$

$$\lambda = \lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t-1}\ln|f'(x(k))| = \int dx\mu(x)\ln|f'(x)|$$

$$h(\mu,\mathbf{A}) = \lim_{n\to\infty}\frac{1}{n}H(B^{(n)}) = \lim_{n\to\infty}\left(H(B^{(n+1)}) - H(B^{(n)})\right),\quad s_\mu = \lim_{diam(\mathbf{A})\to 0}h(\mu,\mathbf{A}),\quad s_\mu = \lambda$$