Eur. J. Phys. **26** (2005) S79–S84 doi:10.1088/0143-0807/26/5/S09

Teaching computational bio-socio-econo-physics

D Stauffer

Institute for Theoretical Physics, Cologne University, D-50923, Germany

Received 23 May 2005
Published 8 July 2005
Online at stacks.iop.org/EJP/26/S79

Abstract
Some examples are given on how methods of computational statistical physics are applied outside physics.

1. Introduction

Statistical physics in this author's style is well known to be the pinnacle of human achievement, and computer simulation its proper tool. The foundations were laid in the 1950s with the Monte Carlo methods of Metropolis *et al*, and the molecular dynamics of Alder and Wainwright. Later, such methods were applied to fields outside of physics such as economy [1], biology [2] or sociology [3, 4]. One should be cautious suggesting such subjects for a thesis in physics, but some former physics students have made it this way into faculty positions outside physics. In Monte Carlo simulations, physicists let $\leqslant 10^{13}$ particles [5] change their status with a probability proportional to $\exp(-\Delta E/k_B T)$ where T is the absolute temperature, k_B is the Boltzmann constant and ΔE is the energy change. For molecular dynamics we solve Newton's law of motion, force = mass times acceleration, using $\leqslant 10^{10}$ atoms [6]. This microscopic emphasis on single particles, as opposed to macroscopic differential equations averaging over many particles, is often called 'independent agents' when applied outside physics. 'Emergence' in these fields is what physicists call self-organization which means systems of many simple particles showing complex behaviour (like freezing or evaporating) which is not evident from the single-particle properties.

An important difference between physics and applications outside physics is the thermodynamic limit. A glass of red wine has about 10^{25} water molecules, which is close enough to infinity for physicists. Economists, on the other hand, are less interested in stock markets with 10^{25} traders. Thus finite-size effects, which often are a nuisance in statistical physics simulations, may be just what we need outside of physics.

I have been teaching computational physics for two decades, but interdisciplinary applications separately only since 2001. Some students prefer the former and others the latter, thus separation may be better than a combined course. In this course and in the present paper, I select problems I am acquainted with since the whole field has grown too large for one semester or five pages. And these problems are those closest to statistical physics. An earlier and now outdated review of mine was published in [7].

2. Biological ageing

Darwinian evolution is similar to thermal physics in that two effects compete: Mother Nature wants to select the fittest and to minimize energy; but more or less random accidents (mutations in biology, thermal noise or entropy in statistical physics) lead to deviations from ideality, such as biological ageing or minimization of the free energy.

The probability to die within the coming year is very small below the age of 30 years but quite appreciable beyond 100 years. In between this, mortality approximately increases exponentially with age. There are many theories for this effect of ageing, and no consensus. Penna, then aged below 30, published the most widely simulated model [8]. The genome was represented there by a string of bits, using single-bit treatments known for decades as Ising spins. Now we present a newer model with worse results but conceptually simpler. Also, it is only a minor modification of the 120 year old Weismann theory that we die to make place for our children [9].

We are born, grow up, have children, become old and die. Let us assume that each baby has in its genes a minimum reproduction age R and a genetic death age D. This genetic death kills us if we are not killed before by hunger, infections or editors. Each adult aged beyond R gives birth at each iteration to one child, with probability p; this child inherits the parental R and D apart from random mutations by ± 1. Lack of food and space is modelled by a Verhulst death probability $N(t)/K$ where K determines the maximum population size. (We require $R > 0$ for the mutations.)

(Before Verhulst, Malthus predicted an exponential increase of human population $N(t)$ from $\mathrm{d}N/\mathrm{d}t = bN$ where b is the difference between birth and death rates. In 1844 Verhulst replaced this by the logistic equation $\mathrm{d}N/\mathrm{d}t = bN(1 - N/K)$ leading to a plateau in population. This additional death probability for each individual, equal to N/K, is not only realistic but also useful computationally to avoid memory overflow.)

In this version, the genetic death age D drifts to infinity since long-lived families produce more children and thus dominate in the population. To avoid this unrealistic longevity, the birth rate p is assumed to vary inversely proportional to $D - R$, which means D and R can vary under the constraint that the average number of children, $(D - R) * p$ is constant (typically chosen near 1.1 in this asexual model). Now D approaches a finite limit, and R remains very small. The mortality is roughly a linear function of age, instead of the desired exponential. The linear mortality function may be appropriate for mayflies (ephemerals), while the exponential increase with human age requires modifications [10]. A one-page computer program is listed in the appendix.

The most important practical application of this research was the explanation of the menopause (or its analogues) through the dependence of babies on their mother and the increased mortality when giving birth at an advanced age [11]; there was no need for a grandmother effect [12] or other human traits.

3. Sociology

3.1. Language evolution

The evolution of human languages has for a long time been compared to biological evolution [13]. Thus an algorithm similar to the Penna ageing program was used to simulate the evolution of languages (including sign languages, bird songs and alphabets) [14]. Each language consisted of eight bits; also 16 to 64 bits were simulated. Initially only one individual and one language exist. Then at each iteration each individual gets one child, which with probability p mutates the parental language by switching one randomly selected bit. A Verhulst death probability as in the above ageing model prevents the population increasing exponentially towards infinity. Speakers of small languages select with a high probability the language of a randomly chosen other person.

As a result, for low p one language dominates, together with a few variants differing by one bit only, like an alphabet today. For high p, many languages of roughly equal sizes remain, with their size distribution roughly log-normal but with a surplus of very small languages, as observed in reality [15]. One may also study the influence of one dominating language, or whether different languages merge into one by adopting bits of the other languages.

3.2. Ghetto formation

Schelling in the first issue of *Journal of Mathematical Sociology* re-invented an Ising-like model (with dilution and Kawasaki dynamics) to describe the formation of black ghettos in predominantly white USA [3]: Afro-Americans like to stay with other Afro-Americans, and analogous behaviour is assumed for people of European origin. As any Ising model simulator knows, at temperatures below the critical temperature, large domains of up-spins coexist between large-domains of down-spins, if no magnetic field is applied. This coexistence remains infinitely long for Kawasaki dynamics (conserved magnetization) and quite long for Glauber dynamics (fluctuating magnetization). Thus one can identify one spin orientation with white and the other with black, to study ghetto formation = phase separation. Actually, this identification of particles with humans is quite old: Jürgen Mimkes told me that more than two millennia ago Empedokles compared some groups of people with wine and water (they mix easily) and others with oil and water (they do not like to mix).

More recently, simulations were made to avoid ghetto formation, with two [16] or several [17] groups of people. For this purpose, the temperature, which measures the degree to which humans tolerate other ethnic groups as neighbours, was increased with time. If the increase was made fast enough, then the domains = ghettos could not become large.

4. Market fluctuations

In economical research, the idea of fully informed and completely rational traders is going out of fashion. Random decisions and herding instincts are now incorporated to simulate day-to-day fluctuations and also longer-lasting bubbles leading to market crashes such as the 1929 crash on Wall Street. And this is what statistical physics did since Boltzmann. Self-organization appears in fluids as well as in the 'invisible hand' of the market theory of Adam Smith, 230 years ago; conspiracy theories are out of fashion in statistical physics and perhaps soon in economics.

Not only the economy has its ups and downs, but also the opinions of business managers about the near future of the economy. We may take again an Ising model and identify an optimist with spin up, and a pessimist with spin down. Real opinion polls look like Ising magnetizations slightly below the critical temperature: the magnetization stays for a long time on one side, then tunnels through to the other side where again it stays for quite a long time [18]. To make the simulations more realistic with also a neutral opinion, one may use a spin-1 model with the states 1, 0 and -1, and a term favouring or disfavouring neutrality. This is the Blume–Capel model and was applied [19] to the swings in the business expectations of

Germany. Here and elsewhere in econophysics, the temperature measures the willingness to disagree with trends, or takes into account the impact of external information not explicitly included in the model.

And just as in physics, it is easier to explain events afterwards than to predict them reliably.

5. Discussion

The above examples were selected not according to their importance in the field but according to their similarity to traditional statistical physics, such as Ising models. In addition, I usually teach other things. If you want to know who will pay for your retirement, you combine bio, socio and econo aspects [20]. A monograph combining these aspects is [21] of which we plan to write a new version.

The Ising model serves here more as a nice example of how to do simulations of independent agents or to explain the emergence of order and broken symmetry; I am not aware of applications where the critical exponents in two dimensions, or the Curie temperature in five, play an important role.

Appendix. An ageing program

This program in Fortran 66 (mostly) denotes R and D as `imin` and `idea` and uses 64-bit integers as random numbers. The function `isign` gives the sign of the second argument, multiplied with the absolute value (here: 1) of the first argument. Questions should be sent to stauffer@thp.uni-koeln.de.

```
   parameter(nmax=6000000,init=nmax/10, nmem=init)
   dimension ibirth(127), idhist(127), numage(127), ibhist(127)
   byte imin(nmem), iage(nmem), idea(nmem)
   integer*8 numage, ibhist, idhist, ibm, ibirth, iverh
   real*8 avdea, avmin
   data iseed/1/, max/ 50000/, ibirth,idhist,numage,ibhist/508*0/
   print *, nmax, init, nmem, iseed, max, ' 10%, 64 bit rng'
   ibm=2*iseed-1
   do 2 i=1,init
     iage(i)=1
     imin(i)=1
2    idea(i)=16
   ibirth(1)=2147483648.0d0*2*2147483648.0d0
   do 9 j=2,127
9    ibirth(j)=2147483648.0d0*(4.40d0/j-2.00d0)*2147483648.0d0
   npop=init
   do 3 itime=1,max
     iverh=2147483648.0d0*(npop*4.0d0/nmax-2.0d0)*2147483648.0d0
     n0=npop
```

```
c     first comes the loop over all Verhulst and genetic deaths; no
      ageing do 1 i=n0,1,-1
        ibm=ibm*16807
        if(ibm.gt.iverh.and.iage(i).lt.idea(i)) goto 1
        imin(i)=imin(npop)
        idea(i)=idea(npop)
        iage(i)=iage(npop)
        npop=npop-1
1    continue
     if(npop.le.0) stop 9
c    deaths are finished, now come births
     n0=npop
     do 7 i=1,n0
       im=imin(i)
       if(iage(i).le.im) goto 7
       ibm=ibm*65539
       if(ibm.gt.ibirth(idea(i)-im)) goto 7
       npop=npop+1
       if(npop.gt.nmem) stop 8
       ibm=ibm*16807
       imin(npop)=max0(1,im+isign(1,ibm))
       ibm=ibm*16807
       idea(npop)=idea(i)+isign(1,ibm)
       iage(npop)=0
7      continue
     avmin=0.0d0
     avdea=0.0d0
     do 8 i=1,npop
       iage(i)=iage(i)+1
       avmin=avmin+imin(i)
8      avdea=avdea+idea(i)
     if(itime.le.max/2) goto 3
     do 10 i=1,npop
       numage(iage(i))=numage(iage(i))+1
       ibhist(imin(i))=ibhist(imin(i))+1
10     idhist(idea(i))=idhist(idea(i))+1
3    if(mod(itime,1000).eq.0)print*,itime,npop,avmin/npop,avdea/npop
    do 11 i=1,126
11   if(numage(i+1).ne.0) print *, i, numage(i),ibhist(i)
1     ,idhist(i),-alog(float(numage(i+1))/float(numage(i)))
    stop
    end
```

Random odd integers are easily and efficiently produced by multiplying them by 16 807 (or by 13^{13} if 64-bit integers are available), where most computers then throw away all leading bits and keep the last 32 (or 64) bits only. Thus if something is to be done with a fixed probability p, you do this if and only if your random integer is below $i_p = (2p - 1) * 2^{31}$ since the random integers are between -2^{31} and $+2^{31}$ in Fortran. (C has unsigned long integers between 0 and 2^{32}). If instead you call a built-in random number generator RAND, RANF or RND, it may do the same or something similar but will normalize each time the resulting integer to the interval between 0 and 1, which is slower and less transparent than calculating i_p once.

References

[1] Stigler G J 1964 *J. Bus.* **37** 117
[2] Kauffman S A 1969 *J. Theor. Biol.* **22** 437
[3] Schelling T C 1971 *J. Math. Sociol.* **1** 143
[4] Zhang J 2004 *J. Math. Sociol.* **28** 147
[5] Tiggemann D 2004 *Int. J. Mod. Phys.* C **15** 1069
[6] Kadau K, Germann T C and Lomdahl P S 2004 *Int. J. Mod. Phys.* C **15** 193
[7] Stauffer D 1999 *Am. J. Phys.* **67** 1207
[8] Penna T J P 1995 *J. Stat. Phys.* **68** 1629
[9] Stauffer D 2002 *Biological Evolution and Statistical Physics* ed M Lässig and A Valleriani (Berlin: Springer) p 258
[10] Makowiec D *et al* 2001 *Int. J. Mod. Phys.* C **12** 1067
[11] Moss de Oliveira S, Bernardes A T and Sa Martins J S 1999 *Eur. Phys. J.* B **7** 501
Sousa A O 2003 *Physica* A **326** 233
[12] Hawkes K 2004 *Nature* **428** 128
Lahdenperä M *et al* 2004 *Nature* **429** 178
[13] 2004 Evolution of language *Science* **303** 1315–35
[14] Schulze C and Stauffer D 2005 *Int. J. Mod. Phys.* C at press
Schulze C and Stauffer D *Phys. Life Rev.* **2** 89
Mira J and Paredes A 2005 *Europhys. Lett.* **69** 1031
Kosmidis K, Halley J M and Argyrakis P 2005 *Physica* A **353** 595
Schwammle V 2005 *Int. J. Mod. Phys.* C at press
de Oliveira V M, Gomes M A F and Tsang I R 2005 *Physica* A at press
[15] Sutherland W J 2003 *Nature* **423** 276
[16] Meyer-Ortmanns H 2003 *Int. J. Mod. Phys.* C **14** 351
[17] Schulze C 2005 *Int. J. Mod. Phys.* C **16**
[18] Jansen K *et al* 1988 *Phys. Lett.* **213** B 203
[19] Hohnisch M, Pittnauer S, Solomon S and Stauffer D 2005 *Physica* A **345** 646
[20] Stauffer D 2002 *Exp. Gerontol.* **37** 1131
Sà Martins J S and Stauffer D 2004 *Ingenierias* **7** 35
[21] Moss de Oliveira S, de Oliveira P M C and Stauffer D 1999 *Evolution, Money, War and Computers* (Leipzig: Teubner)
Stauffer D, Moss de Oliveira S, de Oliveira P M C and Sa Martins J S *Biology, Sociology and Geology by Computational Physicists* (Amsterdam: Elsevier) in preparation