4 Percolation – till 05.12.2023

Before you start working on this task, I recommend checking Introduction to Computational Physics (ICP). All of the algorithms listed below are described in ICP.

Write a program for the site percolation problem on the square lattice $L \times L$. Just to remind you, in the site percolation problem, each site of a lattice is occupied independently with probability p. Therefore a single Monte Carlo trial consists of initiating a lattice: for each site you choose a random number $r \sim U(0,1)$ and if $r < p$ then put $A[i,j] = 1$ at a given site, otherwise put $A[i,j] = 0$. After this step you are ready to:

1. check if the path connecting the first and the last row exists (use The Burning Method, check e.g. pages 24-25 in ICP)
2. the maximum cluster size s_{max}
3. distribution of clusters $n(s,p,L)$, where s is a size of a cluster (to find clusters use the Hoshen-Kopelman (HK) algorithm, check e.g. pages 28-31 in ICP)

Using Monte Carlo simulations, which should consist of T trials generate the output files for:

1. The probability P_{flow} that the path connecting the first and the last row exists as a function of p, where $p = p_0 : dp : p_k$.
2. The average size of the maximum cluster $\langle s_{max} \rangle$ as a function of p.
3. Distribution of clusters $n(s,p,L)$ for a given p.

The output file:

- Data described in points 1-2 should be saved in a single file for each set of input parameters with an automatically generated name according to the following scheme $'Ave\_L' + L + 'T' + T + '.txt'$ and should consists of 3 columns separated by double space in the following order: $p \;\; P_{flow} \;\; \langle s_{max} \rangle$, something like (artificial values are given here):

  0.1 0 7
  0.2 0 8
  0.3 0 7
  0.4 0 9
  0.5 0 10
  0.6 1 100

- The output file with the distribution of clusters $n(s,p,L)$ should have an automatically generated name according to the following scheme $'Dist\_p' + p + 'L' + L + 'T' + T + '.txt'$ and should consists of 2 columns separated by double space: $s \;\; n(s,p,L)$.

To collect output data you will need to introduce input data:

1. L: the linear size of the system,
2. T: number of trials,
3. p_0: minimum value of p in the loop,

4. p_k: maximum value of p in the loop,

5. dp: step with which p changes inside of the loop.

All this data should be saved in *perc_ini.txt* and used to run the program (program should load this data as an input). User changes data inside the file. This is much more convenient for the user than introducing each time data from the keyboard, especially if the user wants to change only one parameter. Example of such a file:

```
100     % L
1000    % T
0.01    % p_0
1       % p_k
0.01    % dp
```

4.1 Presentation of results

After collecting appropriate data you should present them properly.

1. Using output files create appropriate figures (in Matlab, Python, Gnuplot - you can choose yourself). Remember to label axis! Use different symbols to distinguish between series of data and add appropriate legend. Remember all figures should be readable also in Black&White that is why except of colors we use also different symbols. This is recommendation of all prestigious journals - ask yourself what is the reason for this. Here is the list of figures you should prepare. For points (b)-(d) use $T = 10^4$, if it is not possible use $T = 10^3$.

   (a) Visualize sample configurations for $L = 10$ and 3 values of $p = 0.4, 0.6, 0.8$ within two methods: (1) use the burning algorithm and describe each site by the number, as shown during the lecture, (2) use the HK algorithm and color each cluster with the different color.

   (b) Probability P_{flow} that the path connecting the first and the last row exists as a function of p for $L = 10, 50, 100$ (use legend).

   (c) The average size of the maximum cluster $\langle s_{max} \rangle$ as a function of p for $L = 10, 50, 100$ (use legend).

   (d) Distribution of clusters $n(s, p, L)$ for a given $p = 0.2, 0.3, 0.4, 0.5, p_c = 0.592746, 0.6, 0.7, 0.8$ (use legend). You can also prepare a figure with 3 subplots: for $p < p_c$, $p = p_c$ and $p > p_c$ (see Figure 2.10 in ICP).

2. (Extra) Analyze numerically cluster size distributions for 3 different p: 1) $p < p_c$ in which we have scaling law given by $n_{p<p_c}(s) = A_1 s^{-A_2} \exp(A_3 s)$; 2) $p \approx p_c$ for which we should have $n_{p \approx p_c}(s) = A_4 s^{-A_5}$; 3) $p > p_c$ where scaling law is $n_{p>p_c}(s) = A_6 \exp\left(-A_7 s^{\left(1-\frac{1}{A_8}\right)}\right)$. Find numerical fits for parameters A_1 to A_8.