Solutions must be submitted via ATENEA by 23:59 on 11th October 2019.

One of the applications of Markov chains is the complexity analysis of randomised algorithms.

The *boolean satisfiability problem (SAT)* is the problem of determining whether there exists an assignation of boolean variables that satisfies a given boolean formula. We will let $x_1, \dots, x_k \in \{0, 1\}$ denote the boolean variables. A literal is either a variable (x_i) or a negation of it ($\overline{x_i}$). A clause c is a disjunction of literals. A *formula in conjunctive normal form (CNF)* ϕ is a conjunction of clauses (or a single one). For instance, consider the following CNF formula:

$$\phi = (x_1 \vee x_3 \vee \overline{x_4}) \wedge (\overline{x_2} \vee x_3) \wedge (\overline{x_3}) \wedge (x_2 \vee x_4) \ .$$

One can check that $\mathbf{x} = (x_1, x_2, x_3, x_4) = (1, 0, 0, 1)$ is a satisfying assignment. Obviously, not all the boolean formulas admit a satisfying assignment.

The 2-SAT problem is the SAT problem for CNF formulas where every clause has exactly 2 distinct variables. The previous example is not in 2-SAT, the next one is:

$$\phi = (x_1 \vee \overline{x_4}) \wedge (\overline{x_2} \vee x_3) \wedge (\overline{x_3} \vee \overline{x_1}) \wedge (x_2 \vee x_4) \ .$$

Consider the following randomised algorithm that given a satisfiable formula, finds a satisfying assignment:

Algorithm: Rand2SAT

Input: a satisfiable CNF formula $\phi = c_1 \wedge \dots \wedge c_m$ of 2-SAT with k variables.
Output: a satisfying assignment for ϕ.

(1) Let $\mathbf{x} = (0, 0, \dots, 0)$.

(2) While there is a violated clause in $\mathbf{x}$,

  (a) Choose c arbitrarily from the set of violated clauses.

  (b) Choose x uniformly at random from the two literals in c.

  (c) Switch the value of the variable x and update $\mathbf{x}$.

(3) Return $\mathbf{x}$.

a) Let X_n be the number of satisfied clauses after n iterations of the loop in (2) (we will refer to it as time n). Given the execution up to time $n - 1$, is it always true that $\mathbb{E}[X_n] \geq X_{n-1}$?

b) Since ϕ is satisfiable, let $\mathbf{x}^*$ be an arbitrary satisfying assignment. Let Y_n be the number of variables in $\mathbf{x}$ whose value coincides with the one in $\mathbf{x}^*$ at time n. Given the execution up to time $n-1$, is it always true that $\mathbb{E}[Y_n] \geq Y_{n-1}$?

c) Argue that if $Y_n = k$, then Rand2SAT terminates at time n. Is the converse true?

d) Is Y_n a Markov chain?

e) Design a Markov chain Z_n such that $Y_n \geq Z_n$. (Hint: modify the Gambler's ruin to design Z_n)

f) Use Z_n to prove that the expected running time of Rand2SAT is at most k^2.

*g) We modify Rand2SAT to stop in bounded time as follows. Let $\ell \in \mathbb{Z}$. If after $2\ell k^2$ iterations of the loop in (2) we have not halted, we break the loop and return the current assignment $\mathbf{x}$. Prove that the ouput of the modified Rand2SAT is a satisfying assignment with probability at least $1 - 2^{-\ell}$. (Hint: use Markov's inequality.)

Comment: In order to get an intuition, you can code the algorithm using your favourite language/mathematical software, and run simulations with different boolean formulas.