HW2 MPRI 2.11.1 — Molecular Programming

You are asked to complete the exercise marked with a [★] and to send me your solutions to:
nicolas.schabanel@ens-lyon.fr
as a PDF file named **HW2-Lastname.pdf** on **Thu. 20/01 before 08:45**.

!! Pick *one of the two* exercises marked with a $[\bigstar]$!! You can work on the origami in team of 2-4 people

■ **Exercise 1 (Exponential random variables & kTAM implementation).** Recall that an *exponential random variable* X with parameter $\lambda > 0$ is defined by: $(\forall x \geqslant 0)\ \mathsf{Pr}\{X \geqslant x\} = e^{-\lambda x}$.

▶ **Question 1.1)** *Compute* $\mathbb{E}[X]$.
▷ *Hint. Recall that if X is a non-negative random variable, then $\mathbb{E}[X] = \int_0^\infty \mathsf{Pr}\{X \geqslant x\}dx$.*

▶ **Question 1.2)** *Show that the exponential distribution is memoryless, i.e. if X is exponentially distributed with parameter λ, then $(\forall t, u \geqslant 0)\ \mathsf{Pr}\{X \geqslant t + u \mid X \geqslant t\} = \mathsf{Pr}\{X \geqslant u\}$.*

Let X and Y be two independent exponentially distributed random variables with respective parameters λ and μ.

▶ **Question 1.3)** *Show that* $\min(X, Y)$ *is also exponentially distributed. What is its parameter?*

▶ **Question 1.4)** *What is the probability that* $\min(X, Y) = X$?

▶ **Question 1.5)** *Same questions as the two above for n indenpendent exponentially dsitributed variables $X_1, \ldots, X_n$ with parameters $\lambda_1, \ldots, \lambda_n$.*

▶ **Question 1.6)** *Assume that a non-negative random variable X is given by its tail distribution $F(x) = \mathsf{Pr}\{X \geqslant x\}$. Show that X is identically distributed as $F^{-1}(U)$ where U is a uniform random variable in $[0, 1]$.*
*Describe how to sample an exponential random variable of rate λ.*

▶ **Question 1.7)** *Propose an algorithm together with a data structure to implement the kTAM model with attachement rate $r_f = k_f[Strand] = k_f\, e^{-G_{mc}}$ and detachment rate $r_{s,b} = k_f\, e^{-b \cdot G_{se}}$ where b is the number of bonds made by the strand with the current agregate.*
Use parameters $k_f = 10^6/M/sec$, $G_{mc} = 12.9$ and $G_{se} = 6.5$ for the algorithmic phase.

■ **Exercise 2 (Tileset for simulating cellular automata).** A cellular automaton consists of a finite set of *states* Q, a function $f : Q^3 \to Q$, called the *rule*, and an initial configuration $c^0 \in Q^*$. The configuration at time $t + 1$ is obtained from the configuration at time t as follows: $c_i^{t+1} = f(c_i^t, c_{i+1}^t, c_{i+2}^t)$ for $0 \leqslant i < |c^t| - 2$. The calculation stops at the first time T such that $|c^T| < 3$ and the result of the computation is c_0^T. A classic visualization of the computation

of a cellular automaton consists of a pyramid where the bottom line is the initial configuration and time goes upwards. Here is an example:
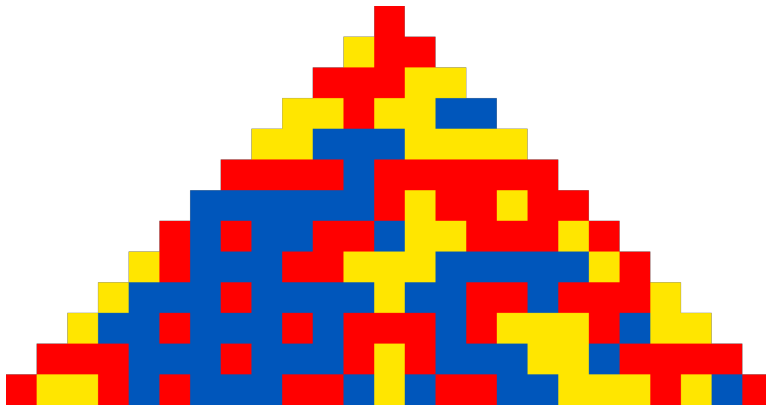

$$
\text{for the rule } f(x, y, z) = \begin{cases} \square & \text{if } \{x, y, z\} = \{\blacksquare\} \text{ or } \{\square, \blacksquare\} \\ \blacksquare & \text{if } \{x, y, z\} = \{\square\} \text{ or } \{\blacksquare, \blacksquare\} \\ \blacksquare & \text{if } \{x, y, z\} = \{\blacksquare\} \text{ or } \{\square, \blacksquare\} \text{ or } \{\square, \blacksquare, \blacksquare\} \end{cases}
$$

▶ **Question 2.1)** *Propose a finite tileset whose self-assembly at temperature $T° = 2$ simulates the computation of any Q-state cellular automata from any initial configuration and whose size is independent of the initial configuration length. Give a generic example of the execution of your assembly for generic computation steps. Give the number of variants of each tile type as a function of $|Q|$. Provide the procedure which selects the tiles used to simulate a given Q-state cellular automaton.*

▷ *Hint. Do you need upscaling? Consider reshaping the pyramid to simplify your design.*

■ **Exercise 3 (Probabilistic simulation Turing Machine at $T° = 1$ in 2D).** Recall that in 3D, for any single-tape binary-alphabet Turing machine M, there is a tile set which simulates M using a clever trick to encode 0s and 1s. These are encoded with bridges and read using two probes where only one go through the bridge:
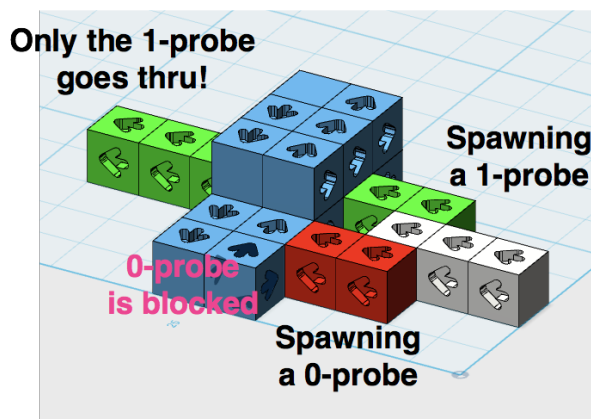
▶ **Question 3.1)** *By adjusting the concentrations (and thus the rate at which the different tiles attached), describe a tile set together with concentrations for each tile type, that simulates a given single-tape binary-alphabet Turing machine M with an arbitrary small error ε for each symbol read in 2D at temperature $T° = 1$.*

[★] **Exercise 4 (Scale the wall).** Recall that a tile assembly system $\mathcal{T} = (T, \sigma, \tau)$ consists of a tile set T, a seed tile $\sigma \in T$ and a temperature $\tau \in \mathbb{N}$. Consider the situation in Fig. 1 consisting of a wall of height h.
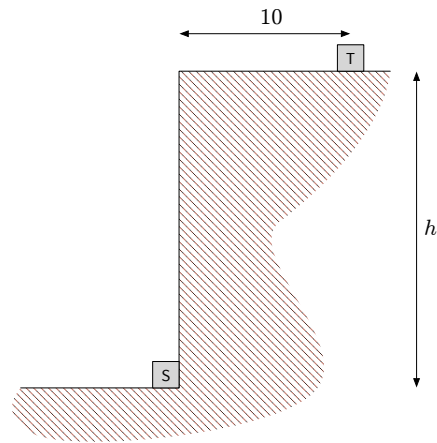
Figure 1: A wall of height h.

▶ **Question 4.1)**   *Can you find a tile assembly system $\mathcal{T}$ for the abstract Tile Assembly Model (aTAM) where the rules are as follows?*

- *The seed tile is placed at position $S = (0,0)$*

- *For all $h \in \mathbb{N}$, every terminal assembly of $\mathcal{T}$ should place a tile at the target position $T = (10, h)$ and be of finite size*

- *$\mathcal{T}$ may not place tiles to the right and below the cut of the plane shown in Figure 1.*

- *You may give an infinite sequence of glues such that the h-prefix of that sequence will appear on the wall, to help the tiles 'climb up'.*

[★] **Exercise 5 (DNA origami).**   Use ENSnano to design a DNA origami.

Getting the software.   First, download ENSnano (either compiled or source) from:

http://www.ens-lyon.fr/ensnano/

You will also find useful tutorials on ENSnano youtube channel:

https://www.youtube.com/channel/UC-5XMNJouMPsB9pW5Q9UVcA

▶ **Question 5.1)**   *1. Pick a 3D shape to design it as a DNA Origami. You might take inspiration from:*

http://www.dna.caltech.edu/Papers/DNAorigami-nature.pdf
or https://www.dietzlab.org/research/

2. Design your 3D DNA origami using ENSnano
   (the scaffold must be at most 7249 nucleotides long)
   Compute its expected size in nanometers. Recall that parallel double helices are about 2.65 nm away when next to each other; two nucleotides are about 0.332 nm distant along the helix axis; a complete turns correspond to ~ 10.44 nucleotides; the angle between the two helices in the narrow groove is $\sim \frac{12}{34} \cdot 2\pi$.
*3. email me the **.ens** file and a one-page report presenting your origami.*