Exercise 1 (Exponential random variables & kTAM implementation). Recall that an exponential random variable $X$ with parameter $\lambda > 0$ is defined by:

$(\forall x \geq 0) \Pr\{X \geq x\} = e^{-\lambda x}$.

Question 1.1) Compute $E[X]$.

Answer: Indeed, by Fubini’s theorem:

$$E[X] = \int_0^\infty \int_0^\infty \Pr\{X \in [x, x + dx]\} dy = \int_0^\infty \Pr\{X > y\} dy = \int_0^\infty e^{-\lambda x} dx = \frac{1}{\lambda}.$$ 

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

Answer: Let $M = \min\{X, Y\}$, then:

$$e^{-\lambda(t+u)+\mu} = e^{-\mu} = \Pr\{X \geq u\}.$$ 

Let $X$ and $Y$ be two independent exponentially distributed random variables with respective parameters $\lambda$ and $\mu$.

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

Answer: Let $M = \min\{X, Y\}$, then:

$$\Pr\{M \geq t\} = \Pr\{X \geq t \wedge Y \geq t\} = \Pr\{X \geq t\} \cdot \Pr\{Y \geq t\} = e^{-(\lambda+\mu)t}.$$ 

$M$ is thus exponentially distributed with parameter $\lambda + \mu$.

Question 1.4) What is the probability that $\min\{X, Y\} = X$?
Answer. ▶ Let \( M = \min(X, Y) \), then:

\[
\Pr\{M = X\} = \int_0^\infty \Pr\{Y > t \land X \in [t, t + dt]\} = \int_0^\infty \Pr\{Y > t\} \cdot \Pr\{X \in [t, t + dt]\}, \quad \text{by independence of } X \text{ and } Y
\]

\[
= \int_0^\infty e^{-t} \cdot \lambda e^{-\lambda dt} = \frac{\lambda}{\lambda + \mu}.
\]

▶ Question 1.5) Same questions as the two above for \( n \) independent exponentially distributed variables \( X_1, \ldots, X_n \) with parameters \( \lambda_1, \ldots, \lambda_n \).

Answer. ▶ By immediate recurrence, let \( Z = \min(X_1, \ldots, X_n) \) and \( \Lambda = \lambda_1 + \cdots + \lambda_n \):

\[
Z \sim \Exp(\Lambda), \quad \mathbb{E}[Z] = \frac{1}{\Lambda}, \quad \text{and } \Pr\{Z = X_i\} = \frac{\lambda_i}{\Lambda}.
\]

▶ Question 1.6) Assume that a non-negative random variable \( X \) is given by its tail distribution \( F(x) = \Pr\{X \geq x\} \). 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 \( \lambda \).

Answer. ▶ First, note that \( F \) is a non-increasing function going from 1 at \( x = 0 \), to 0 at \( \infty \), its inverse \( F^{-1} : [0, 1] \to \mathbb{R}_+ \) is thus defined almost everywhere. Now, for all \( x \geq 0 \),

\[
\Pr\{F^{-1}(U) \geq x\} = \Pr\{U \leq F(x)\} = F(x) = \Pr\{X \geq x\}
\]

as \( F \) is non-increasing. Thus \( X \) and \( F^{-1}(U) \) are identically distributed.

Now, if \( X \sim \Exp(\lambda) \), we have \( F(x) = \Pr\{X \geq x\} = e^{-\lambda x} \). As \( F^{-1}(u) = -\ln(u)/\lambda \), we can sample \( X \) by picking a uniform random value \( u \) in \([0, 1]\) and outputting \( x = -\ln(u)/\lambda \). ▶

▶ Question 1.7) Propose an algorithm together with a data structure to implement the kTAM model with attachment rate \( r_f = k_f [\text{Strand}] = k_f e^{-G_{mc}} \) and detachment rate \( r_{s,b} = k_f e^{-G_{se}} \) where \( b \) is the number of bonds made by the strand with the current aggregate.

Use parameters \( k_f = 10^6/\text{M/sec}, G_{mc} = 12.9 \) and \( G_{se} = 6.5 \) for the algorithmic phase.

Answer. ▶ We consider cells of \( \mathbb{Z}^2 \) where the seed is placed forever (the seed must be marked as undetachable). We determine the date of the next attachment (resp. detachment) event in an empty (resp. full) cell of \( \mathbb{Z}^2 \), using an independent exponential random variable \( \Exp(r_f) \) (resp. \( \Exp(r_{s,b}) \)) so that these events happen every \( 1/r_f \) (resp. \( 1/r_{s,b} \)) on average.

Each time a cell is updated, we draw new random dates for the upcoming events of its neighbors (recall that exponential variable are memoryless). We store all the upcoming events in a priority queue (a heap) from which we pop the next event (the one with the minimum date).

Recall to delete obsolete upcoming events each time one of its neighbors is updated to avoid flooding the memory. Recall also to avoid flooding the heap with attachment events in an isolated cell.

Note that an attachment event is determined as follows: Pick the event time as \( t \sim \Exp(k_f \cdot \sum_j [\text{Strand}_j]) \) and select \( \text{Strand}_i \) with probability \( [\text{Strand}_i]/\sum_j [\text{Strand}_j] \) as it is distributed as the first time of a strand among \( \text{Strand}_1, \ldots, \text{Strand}_n \) attaches (recall the answers to questions above). ▶
Exercise 2 (Triangle). We consider the family of shapes $S_{p}$, $p \geq 1$, of (discretized) right triangles whose horizontal and vertical sizes are respectively $np$ and $n$, for some $n \geq 1$. The figure below shows the shape in $S_3$ for which $n = 4$.

Question 2.1) Give a tile set $T_p$ whose final productions are exactly the shapes in $S_p$ at temperature $T^0 = 2$. How many tiles does it use? Give a generic assembly and indicate the glues, the seed tile and the assembly order on it. No proof asked.

Answer. We use $p + 2$ tiles plus the seed tile to build the hypothenuse by repeating the pattern $1, \ldots, p + 1$ except for the last one which uses an extra special tile $p + 2$ that non-deterministically stops the growth. The triangle is filled using a last tile type $p + 3$ attaching to the aggregate by the north and west sides. Here is the tile set containing $p + 4$ tiles ($\star$ is the seed tile):

for $1 \leq i < p$

There are $p$ types of red glues which are used to count the $p$ steps to the right before going up using the green glue. The blue glue is used to fill the triangle. The assembly is illustrated below together with the assembly order.

A simple recurrence on the shapes of partial assemblies shows that the final productions are exactly the shapes in $S_p$.

Exercise 3 (Tileset for simulating cellular automata). A cellular automaton consists of a finite set of states $Q$, a function $f : Q^3 \rightarrow Q$, 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: $c_{i+1} = f(c_i, c_{i+1}, c_{i+2})$ for $0 \leq i < |c| - 2$. The calculation stops at the first time $T$ such that $|c_T| < 3$ and the result of the computation is $c_T$. 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:
for the rule \( f(x, y, z) = \) 
\[
\begin{align*}
\text{if } \{x, y, z\} = \{\square, \square\} & \quad \text{or } \{\square, \square\} \\
\text{if } \{x, y, z\} = \{\square, \square\} & \quad \text{or } \{\square, \square\} \\
\text{if } \{x, y, z\} = \{\square, \square\} & \quad \text{or } \{\square, \square\} \text{ or } \{\square, \square\}
\end{align*}
\]

**Question 3.1** Propose a finite tileset whose self-assembly 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.

Answer. As we need to know the states of both neighbors to compute the next step, we choose to slant the pyramid by shifting every row by 2 positions to the right. The first two tiles of each row will collect the first two states of the row and the following tiles on that row will compute the new state and propagate the values of the two new preceding states to the right.

Here is the tileset. The tiles marked with a star are used to encode the initial configuration. There are \( O(|Q|^3) \) tiles in total.

| \( 0 \) tiles | \( |Q| \) tiles | \( |Q|^2 \) tiles | \( |Q|^3 \) tiles | \( |Q|^3 \) tiles |
|----------------|----------------|----------------|----------------|----------------|
| \( \star \) tiles | \( \star \) tiles | \( \star \) tiles | \( \star \) tiles | \( \star \) tiles |

Here is a generic assembly:
Exercise 4 (Probabilistic simulation Turing Machine at $T^\circ = 1$ 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:

![Only the 1-probe goes thru!](image)

**Question 4.1** 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 $\varepsilon$ for each symbol read in 2D at temperature $T^\circ = 1$. 