

Modelado de la propagación del virus COVID-19

1. Descripción del sistema

El objetivo del sistema es simular las interacciones que se producen entre una población de individuos. En este caso, aunque se podría utilizar para diferentes aplicaciones, se va a simular la evolución y los efectos de un virus biológico sobre una población que se desplaza por un escenario compuesto por dos dimensiones.

Debido a que no es posible hacer una representación completamente realista de ningún sistema físico, por la cantidad de variables que habría que controlar, se utilizarán modelos que representan ciertas características de los componentes que se quiere modelar. Estos modelos están parametrizados, es decir, dependiendo de los valores con los que se inicialicen se comportarán de diferente forma. En particular en esta práctica se va a modelar el comportamiento del virus *COVID-19*.

2. Modelado de los componentes

El sistema está compuesto por tres elementos principales:

- Persona/Virus
- Población
- Escenario

Para cada uno de estos elementos habrá que construir un modelo el cual represente su comportamiento y las interacciones que pueda tener con otros elementos del sistema.

2.1. Persona/Virus

Debido a que un virus tiene que estar asociado a una persona para poder existir, el modelado de estos dos componentes se hará de manera conjunta. En particular, el modelado de cada persona se tiene que realizar en base a dos conjuntos de características. El primer conjunto sirve para modelar la infección:

- Edad: Número entre 0 y 100.
- Estado: Número entre 0 y 3: (0) si el individuo esta sano, (1) si está infectado pero no tiene síntomas, (2) si está infectado y tiene síntomas y (3) si se ha recuperado.
- Probabilidad de morir una vez que se ha infectado.

Por su parte, el segundo conjunto sirve para modelar la interacción entre los individuos y el escenario:

- Posición: Vector de coordenadas $p = \{p_x, p_y\}$ que representa la posición de un individuo en el escenario.
- Velocidad: Vector de velocidad $v = \{v_x, v_y\}$ que representa la dirección y la velocidad a la que el individuo se está moviendo.

2.2. Población

La población va a ser modelada utilizando un conjunto de personas cuyo tamaño puede variar en el tiempo. La interacción entre las personas se va a producir en un escenario 2D sobre el cual se pueden mover libremente atendiendo a una serie de parámetros con los que se inicializa el sistema. La población hay que representarla de manera dinámica, es decir, las personas que fallecen deben de ser eliminadas de la lista.

Por el momento, una población se va a definir utilizando los siguientes parámetros:

- **Tamaño de la población:** Indica el número máximo de individuos que tiene la población.
- **Media de edad de los individuos:** Este parámetro permite definir distintos tipos de poblaciones en función de la edad.
- **Radio de contagio:** Los individuos contagiados (presenten o no síntomas) pueden contagiar a otros que se encuentren en un radio menor o igual a este parámetro.
- **Probabilidad de contagio:** Los individuos que se encuentren dentro del radio de contagio pueden ser o no contagiados en función de este parámetro. Sirve para modelar si se toman medidas preventivas o no.
- **Periodo de incubación:** Tiempo que debe pasar desde que un individuo se contagia hasta que presenta síntomas.
- **Periodo de recuperación:** Tiempo que debe pasar desde que un individuo presenta síntomas hasta que se recupera. Durante este periodo de tiempo un individuo puede morir en función de la edad.
- **Probabilidad de cambio de velocidad y dirección:** Los individuos se mueven libremente por el escenario y se podrá cambiar su velocidad y dirección de manera aleatoria.

2.3. Escenario

El escenario representa el mundo sobre el que los individuos se mueven e interactúan. En una primera aproximación, vamos a representarlo como una malla (superficie) 2D de tamaño variable. Este tamaño es un parámetro del sistema y no se puede modificar durante la simulación. En este escenario, cada individuo será representado como un punto, es decir, si pensamos el escenario como si fuera una pantalla, cada individuo se representará con un pixel.

En la Figura 1 se puede ver una representación del estado del sistema en la que coexisten individuos sanos (verde), contagiados asintomáticos (naranja), contagiados sintomáticos (rojo) y recuperados (azul) en un escenario en dos dimensiones.

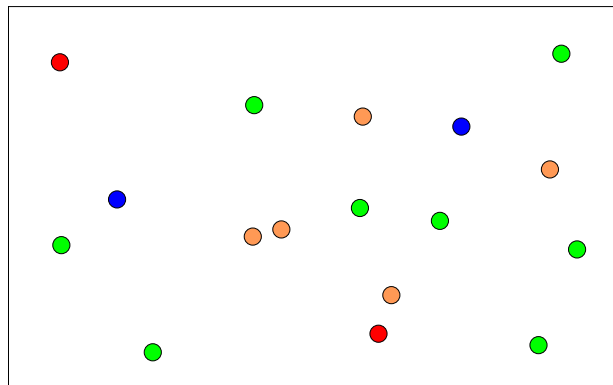


Figura 1: Representación del escenario 2D en el que interactúan los individuos.

3. Simulación del sistema

La simulación del sistema será de tipo *time-driven*. Ésto quiere decir que en cada unidad de tiempo hay que calcular todas las interacciones que se producen entre los individuos de la población, es decir, hay que actualizar todos los componentes del sistema y recoger y almacenar las métricas necesarias..

3.1. Inicialización del sistema

En esta etapa hay que inicializar todos los parámetros de la simulación. En particular hay que definir el tamaño del escenario 2D, el número de individuos de la simulación y los parámetros asociados con cada individuo. Así mismo, hay que definir la duración de la simulación en unidades de tiempo. Todas las simulaciones comenzarán eligiendo de manera aleatoria al paciente 0.

3.2. Cálculo de las interacciones

Las interacciones que se producen entre los diferentes individuos y el escenario se pueden clasificar en tres tipos:

- Estado de los individuos: Hay que controlar para cada individuo el estado en el que se encuentra teniendo en cuenta el tiempo transcurrido entre los diferentes estados. En particular, hay que controlar si un individuo está infectado pero asintomático, cuando se convierte en sintomático, si un individuo está infectado y sintomático, cuando se recupera o, en su caso, cuando muere (usando la probabilidad de muerte por edad).
- Propagación: Para cada individuo hay que controlar su estado y el radio de contagio. En caso de que esté contagiado y haya individuos sanos dentro de su radio de acción, hay que determinar si estos individuos se contagiarán usando la probabilidad de contagio.
- Movimiento: Hay que calcular la nueva posición de todos los individuos en función del vector posición y el vector velocidad. La dirección y la velocidad en la que se mueven los individuos puede ser alterada de manera aleatoria. Hay que tener en cuenta que las dimensiones de nuestro escenario 2D son finitas por lo que habrá que controlar que los individuos no se salgan del escenario.

En la Figura 2 se puede ver un ejemplo de los dos primeros puntos. En concreto, se pueden ver tres ejemplos del primer punto los cuáles están marcados con flechas grises: dos individuos que estaban contagiados pero eran asintomáticos (naranjas) han pasado a ser sintomáticos (rojos) y un individuo que estaba contagiado (rojo) se ha recuperado (azul). Respecto al segundo punto, se puede ver cómo se evalúa el radio de acción de los individuos contagiados. En este caso dos individuos sanos se encuentran dentro del radio pero sólo uno de ellos resulta contagiado (flechas rosas). Esta decisión se toma utilizando la probabilidad de contagio de los individuos contagiados.

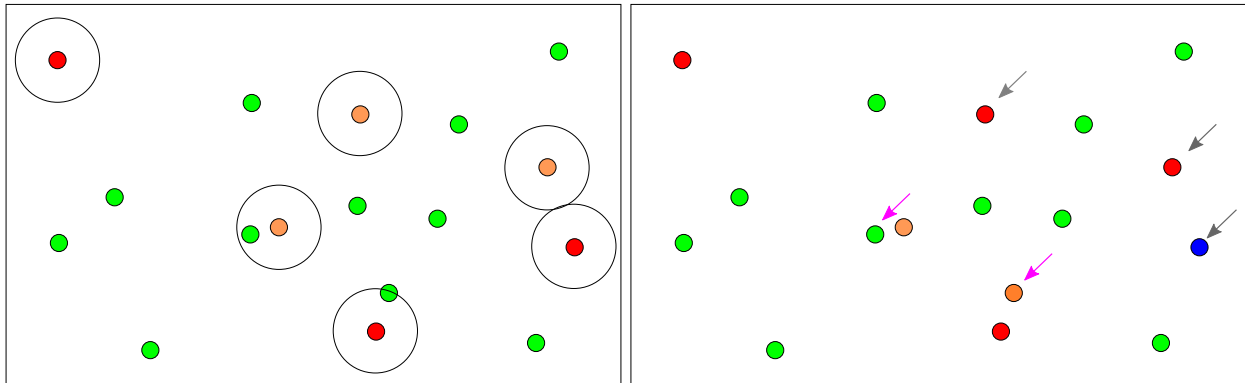


Figura 2: Representación de cómo se calculan los primeros dos puntos en el cálculo de interacciones.

Respecto al tercer punto (no mostrado en la Figura), un ejemplo de movimiento sería por ejemplo un individuo que se encuentra en la posición $p = (2,3)$ y que tiene un vector de velocidad $v = (2,4)$. En este caso la nueva posición será $p' = (4,7)$. Es posible que los individuos cambien de dirección y velocidad en base a las probabilidades cambio de dirección y velocidad definidas para el sistema, es decir, estas probabilidades son globales para todos los individuos. Con ellas podremos simular medidas de confinamiento.

3.3. Recogida de métricas

Dentro de la información que el sistema debe guardar se encuentran las métricas que utilizaremos para registrar la evolución del sistema. El objetivo de la simulación es medir la evolución del número de personas sanas, contagiadas (asintomáticas y sintomáticas) y recuperadas. Así mismo, nos interesa conocer la posición en el escenario 2D de cada persona, así como una métrica utilizada en epidemiología, llamada número reproductivo básico (R_0), la cual indica la capacidad de contagio, es decir, el número de personas que es capaz de contagiar un paciente infectado.

Debido a que la cantidad de información que se requiere registrar en simulaciones de larga duración puede ser demasiado grande, se deberá realizar una simulación por *batches*, es decir, se seleccionarán periodos de tiempo y sólo se guardarán las medias de los valores de las métricas en esos intervalos. Si el valor del *batch* es 1, se guardará toda la información.

4. Tareas a realizar / Plazo

- Diseñar una versión serie que simule el comportamiento descrito. **23/04/2020**
- Implementar el diseño realizado en el punto anterior. **23/04/2020**
- Documentar el rendimiento de la versión serie, en particular, la memoria requerida por el sistema en función del tamaño del escenario y del número de individuos de la población, así como el tiempo requerido para realizar simulaciones de distinta duración. **23/04/2020**
- Propuesta de paralelización para la implementación serie. Recordad que es muy importante diseñar la versión serie teniendo en cuenta como planeáis realizar la versión paralela. **14/05/2020**
- Implementación de la versión paralela y documentación del rendimiento obtenido utilizando los mismos experimentos que los realizados para la versión serie. **14/05/2020**

Appendices

A. Uso de distribuciones de probabilidad

La definición de varias características del sistema depende de números aleatorios lo que nos va a permitir modelar diferentes escenarios. El modelado de estos valores se realiza utilizando distribuciones de probabilidad. Normalmente para generar números aleatorios utilizamos la distribución uniforme (*rand()*) pero en este caso, para modelar algunos escenarios mas complejos, deberíamos utilizar otras. Por ejemplo, para modelar la probabilidad de muerte de un individuo podríamos implementar una función que calculara la probabilidad en base a los siguientes datos:

- < 50 años: 0.004
- ≥ 50 y < 60 años: 0.013
- ≥ 60 y < 70 años: 0.036
- ≥ 70 y < 80 años: 0.08
- ≥ 80 años: 0.148

En este ejemplo hemos utilizado una distribución *ad-hoc* pero podríamos utilizar otras como por ejemplo una distribución normal para modelar la edad de la población. Otro ejemplo podría ser el modelado del período de incubación del virus que varía entre 1 y 14 días siendo la media 5 o 6.