

**Máster universitario de Ciencia de Datos**

**Práctica Final – PRA1**

**Aprendizaje por refuerzo – Implementación de un agente para la guía autónoma.**

Autor:

Mario Ubierna San Mamés

|  |
| --- |
| Índice de Contenido |

[Índice de Contenido 3](#_Toc91333359)

[Índice de ilustraciones 4](#_Toc91333360)

[1. Entorno 5](#_Toc91333361)

[1.1. Ejercicio 1.1 (0.5 puntos) 5](#_Toc91333362)

[1.2. Ejercicio 1.2 (0.5 puntos) 5](#_Toc91333363)

[1.2.1. Espacios de observaciones 6](#_Toc91333364)

[1.2.2. Espacios de acciones 8](#_Toc91333365)

[2. Agente de referencia 11](#_Toc91333366)

[3. Propuesta de mejora 12](#_Toc91333367)

[4. Bibliografía 13](#_Toc91333368)

|  |
| --- |
| Índice de ilustraciones |

[Ilustración 1 - Espacio de observaciones KINEMATICS. 6](#_Toc91333369)

[Ilustración 2 – Espacio de observaciones GRAYSCALE IMAGE. 7](#_Toc91333370)

[Ilustración 3 - Espacio de observaciones OCCUPANCY GRID. 7](#_Toc91333371)

[Ilustración 4 – Espacio de observaciones TIME TO COLLISION. 8](#_Toc91333372)

|  |
| --- |
| Entorno |

*Estamos trabajando sobre el problema de guía autónoma y en particular queremos solucionar el caso de conducción por carretera.*

*Para ello, se elige highway-env como entorno simplificado. El entorno se puede encontrar en el siguiente enlace:* [*https://github.com/eleurent/highway-env*](https://github.com/eleurent/highway-env)*.*

*En particular, nos centramos en el entorno highway, donde se representa una carretera con múltiples carriles y con un cierto número de coches transitando.*

*El objetivo de nuestro agente es de controlar un vehículo de modo que pueda:*

* *evitar colisiones con los otros vehículos en carretera;*
* *tener la máxima velocidad posible;*
* *mantenerse en el carril derecho lo más posible.*

*El entorno se considera superado cuando en el tiempo de observación máximo establecido, se hayan verificado las tres condiciones anteriormente introducidas.*

*El entorno también proporciona diferentes tipos de observaciones que se pueden ver en la documentación adjunta:* [*https://highway-env.readthedocs.io/en/latest/observations/index.html#id1*](https://highway-env.readthedocs.io/en/latest/observations/index.html#id1)*.*

## Ejercicio 1.1 (0.5 puntos)

*Se pide explorar el entorno y representar una ejecución aleatoria.*

Resuelto en el fichero de *jupyter notebook*.

## Ejercicio 1.2 (0.5 puntos)

*Explicar los posibles espacios de observaciones y de acciones*.

*Highway-env* nos permite hacer uso de diferentes tipos de observaciones y de acciones para todos los entornos, es decir, que podemos configurar cómo son los estados y las acciones para cada entorno.

Cabe destacar que por defecto cada entorno tiene un tipo de observación y de acciones, pero éstos se pueden modificar.

### Espacios de observaciones

Respecto a los diferentes tipos de observaciones nos encontramos con cuatro espacios [1]:

* *Kinematics.*
* *Grayscale Image.*
* *Occupancy grid.*
* *Time to collision.*

#### KINEMATICS

Este espacio de observación se caracteriza porque es una matriz del tipo V x F, dónde V es el número de vehículos cercanos que hay y F el conjunto de características que definen a cada vehículo.

Cabe señalar que el número de vehículos es una constante, y por defecto toma el valor 50.

Un ejemplo de este espacio de observaciones sería el siguiente:

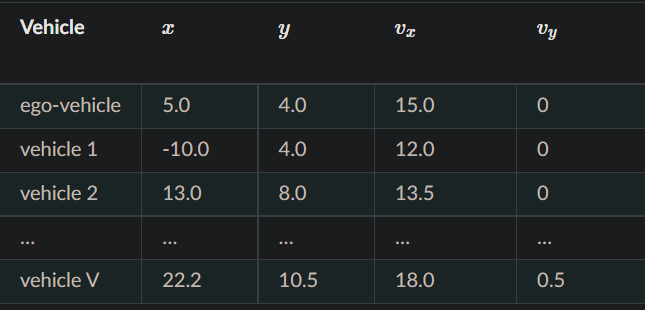


Ilustración 1 - Espacio de observaciones KINEMATICS.

Este es el espacio de observaciones más usado, y es el que vamos a utilizar para crear el agente de referencia.

#### GRAYSCALE IMAGE

La característica principal de este espacio de observaciones es que usa una imagen como tal para definir el estado, es decir, hay una matriz W x H donde W implica el ancho de la imagen y H la altura de la misma.

Un aspecto a tener en cuenta, es que la imagen está en una escala de grises y para transformar e interpretar dicha imagen a color se tiene que hacer una conversión de la misma, a partir de los pesos definidos en la configuración del espacio de observaciones.

Otra de las ventajas es que nos permite almacenar diferentes imágenes, por si se requieren a la hora del aprendizaje hacer uso de estados anteriores.

Un ejemplo del espacio de observaciones sería:

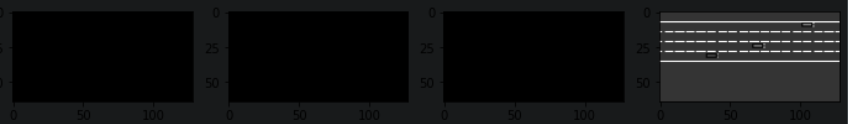


Ilustración 2 – Espacio de observaciones GRAYSCALE IMAGE.

#### OCCUPANCY GRID

Este espacio de observaciones “fusiona” los dos anteriores, es decir, tenemos una matriz de tamaño W x H, pero ahora estos valores no se corresponden con el tamaño de una imagen, sino que representan el espacio que hay alrededor de nuestro vehículo a partir de celdas rectangulares uniformes.

Otro componente de este espacio de observaciones es F, dando lugar a un array W x H x F, es decir, cada celda se describe a partir de sus características. Por ejemplo, para la característica *“presence*” podríamos tener la siguiente matriz:



Ilustración 3 - Espacio de observaciones OCCUPANCY GRID.

Lo que se interpreta de la anterior imagen es que hay dos vehículos cercanos al nuestro, es decir, hay dos “1”; uno de ellos está un poco más al norte que nosotros y el otro está más al este.

#### TIME TO COLLISION

Es el último espacio de observaciones en el que la idea es completamente diferente a las demás, en este caso se busca calcular y representar el tiempo de colisión que hay entre nuestro vehículo y los vehículos que están en la misma carretera.

Para ello se hace uso de una matriz de matrices del tipo V x L x H, V son las diferentes velocidades a las que puede ir nuestro vehículo, por lo tanto, si hay tres velocidades vamos a tener tres matrices. L representa los carriles contiguos a nuestro carril actual y H representa el tiempo necesarios (pasos) para que se produzca una colisión.

Un ejemplo de este espacio de observaciones sería:

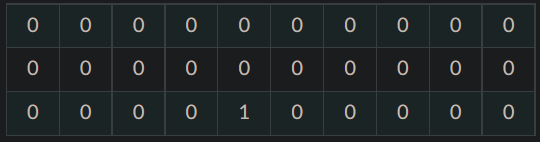


Ilustración 4 – Espacio de observaciones TIME TO COLLISION.

En la anterior imagen vemos que solo tenemos una matriz, por lo que V = 1. Dentro de la matriz hay tres filas o carriles (L = 3) y finalmente H es igual a 10, es decir, se predice la colisión en los próximos 10 pasos. Tal y como podemos ver en la imagen, en el carril de la derecha circula un vehículo con el que nos podemos chocar en 5 segundos (pasos).

### Espacios de acciones

Respecto a los diferentes espacios de acciones tenemos [2]:

* *Continuous Actions.*
* *Discrete Actions.*
* *Discrete Meta-Actions.*
* *Manual control.*

#### CONTINUOUS ACTIONS

Se caracteriza porque el agente es capaz de controlar el vehículo pero a un bajo nivel, es decir, se puede controlar la aceleración y el ángulo de dirección necesarios para alcanzar el objetivo del problema pero accediendo directamente a los parámetros del vehículo.

#### DISCRETE ACTIONS

Este espacio de acciones es igual que el anterior pero hace una cuantificación uniforme, es decir, que la cuantificación es la misma para todos los niveles.

#### DISCRETE META-ACTIONS

En este caso, se añade una capa para que el control del vehículo por parte del agente sea más fácil. De esta forma se consigue que el agente sea capaz de llevar el vehículo a la velocidad que se desea.

Estas *meta-actions* se basan en definir los cambios de carril y la velocidad, y estos valores son usados por los controladores de bajo nivel.

Estas acciones son básicamente cinco:

* 0: se corresponde con cambiar al carril de la izquierda.
* 1: mantenernos en el mismo carril y a la misma velocidad.
* 2: cambiarnos al carril de la derecha.
* 3: aumentar la velocidad.
* 4: disminuir la velocidad.

Cabe destacar que este tipo de espacio de acciones va a ser el usado en la práctica.

En resumen, todos los espacios de acciones comentados se basan en lo mismo, pero éste nos permite controlar el vehículo a un alto nivel, es decir, no tenemos que tener conocimiento sobre los diferentes parámetros usados en el comportamiento del vehículo, esto es lo que diferencia este espacio de acciones respecto a los dos anteriores.

#### MANUAL CONTROL

En este caso el vehículo no aprende de forma automática a moverse, sino que somos nosotros los que con las flechas del teclado movemos el vehículo, es decir, nosotros nos convertimos en el agente del problema.

|  |
| --- |
| Agente de referencia |

*Para este ejercicio considera la base de datos descrita en el documento “Diseño de una base de datos para una app de mensajería instantánea” que se encuentra en los materiales del curso.*

|  |
| --- |
| Propuesta de mejora |

*Para la realización de este ejercicio se seguirán las instrucciones del caso de estudio ubicado en la siguiente URL*

|  |
| --- |
| Bibliografía |

[1] «Observations — highway-env documentation». https://highway-env.readthedocs.io/en/latest/observations/index.html (accedido dic. 25, 2021).

[2] «Actions — highway-env documentation». https://highway-env.readthedocs.io/en/latest/actions/index.html (accedido dic. 25, 2021).