

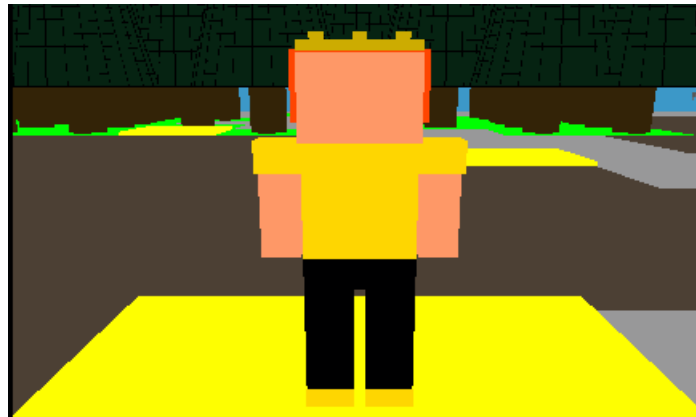
INTELIGENCIA ARTIFICIAL

E.T.S. de Ingenierías Informática y de Telecomunicación

Práctica 2

**Combinando comportamientos
reactivos y deliberativos**

(Los extraños mundos de BelKan II)



DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL

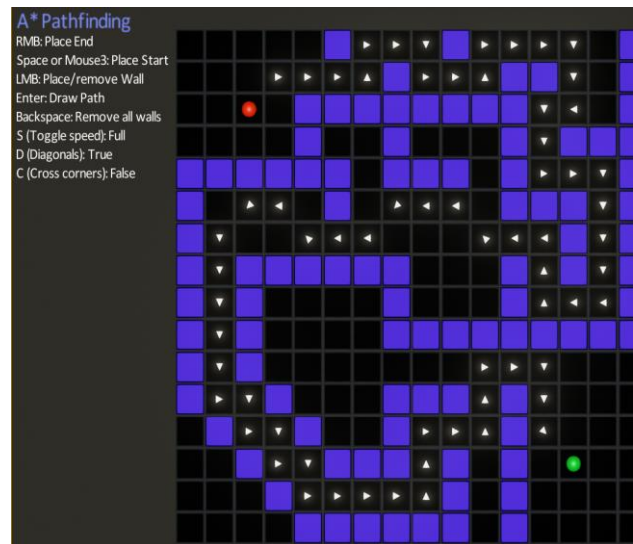
UNIVERSIDAD DE GRANADA

Curso 2016-2017

1. Introducción

1.1. Motivación

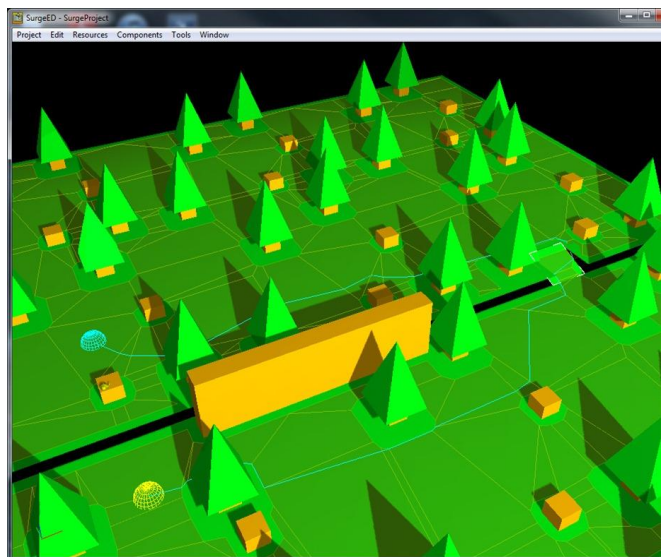
En la primera práctica de la asignatura *Inteligencia Artificial* se ha diseñado un agente puramente reactivo que es capaz de reconocer el terreno que le rodea, a partir de una información sensorial básicamente centrada en la visión (bueno, una representación a muy alto nivel de la información visual). En esta segunda práctica seguiremos trabajando con el mismo problema, y el objetivo será intentar mejorar el comportamiento del agente incluyendo procesos deliberativos. Para ello, serán de utilidad **las técnicas estudiadas en el tema 3 de la asignatura**, es concreto, las técnicas de búsqueda sin y con información.



1.2. Personajes virtuales en juegos de ordenador

En el enunciado de la primera práctica se hablaba de los personajes virtuales en los juegos de ordenador y en especial de aquellos casos en los que los comportamientos reactivos se descuidan, dando mayor relevancia a los deliberativos. Esta falta de atención a los reactivos tiene una explicación. En muchos de estos juegos, la parte deliberativa tiene un papel más importante, en especial, en los juegos de estrategia dónde la inclusión de la inteligencia conlleva una exploración en el tiempo de los resultados de la aplicación de una determinada política, para determinar si se debe cambiar.

De entre estos comportamientos deliberativos, quizás el más utilizado en videojuegos es el que permite encontrar un camino (“path finding”). Por ejemplo, en un juego como “Age of Empire” (o cualquier juego de este tipo) cuando se selecciona uno de los personajes y se marca un lugar en el mapa (indicando que se quiere que ese personaje vaya hasta esa posición) se desencadena un comportamiento deliberativo



para construir un camino que permita llevarlo a esa posición, eludiendo los accidentes geográficos del terreno. Dicho proceso deliberativo es un proceso de búsqueda que construye una secuencia de acciones (en este caso, exclusivamente relacionadas con el movimiento) devolviendo el plan que debe seguir el personaje.

El caso anterior, es el más simple de uso. Ahora supongamos que en lugar de marcar sobre un punto del terreno, marcamos sobre un personaje de un adversario con la intención de atacar. Aunque el procedimiento de búsqueda es semejante al anterior, las condiciones cambian, ya que si ese personaje puede moverse, las coordenadas del estado objetivo cambian. Si además, el movimiento de ese personaje es impredecible, la situación aún se complica más.

Existen distintas técnicas para abordar este problema. La simple sería replanificar en cada iteración del juego. Es la más simple, y también la más costosa en tiempo. Técnicas alternativas tratan de arreglar los planes, detectando que partes del plan original se mantienen útiles, y cuales deben replanificarse. Una situación como la anterior podría darse en nuestro problema: “nuestro agente decide ir a un determinado punto del mapa y en el camino se cruza un aldeano”. Es por tanto, interesante, por un lado, controlar que el plan se está ejecutando correctamente y por otro lado, si se detecta que el plan ha fallado, encontrar una forma eficiente en tiempo para reparar el plan.

Hablando en general de los sistemas que combinan procesos reactivos y deliberativos, un error común, como ya se comentó en la práctica anterior, es considerar los comportamientos deliberativos de forma independiente a los reactivos. En concreto, se comentó, que en el caso del juego “Age of Empire” se decía: *“... El segundo ejemplo, y más importante a la hora de jugar, tiene que ver con el desplazamiento de las unidades guerreras por el mapa. Lo habitual es tomar un grupo de soldados e indicarles que vayan a un punto concreto del mapa. Bien, el grupo emprende el camino y va a donde le has dicho, y nada les distrae de esta acción, así que es posible que durante el trayecto un grupo de soldados enemigos te ataquen sin que ellos se defiendan. De hecho, una estrategia muy simple para ganar al jugador no humano es situarse en un punto intermedio de esa trayectoria y atacarles ahí. Los rivales no se defenderán. Claramente, no es un comportamiento inteligente y debería mejorarse incluyendo comportamientos reactivos que implementen la lógica más básica: <<si quiero llegar a un sitio, tengo que mantenerme vivo para llegar>>”*. En un comportamiento inteligente, normalmente las acciones desencadenadas por procesos reactivos tienen prioridad sobre los deliberativos y por consiguiente, darse cuenta de esta prioridad, es clave para definir un buen comportamiento.

2. Los extraños mundos de BelKan II

En esta práctica, nos volvemos a situar en el mundo de las aventuras gráficas de los juegos de ordenador para intentar construir sobre él personajes virtuales que manifiesten comportamientos propios e inteligentes dentro del juego. Intentamos situarnos en un problema habitual en el desarrollo de juegos para ordenador y vamos a añadir comportamientos deliberativos al comportamiento reactivo previamente definido en la práctica anterior.

A continuación pasamos a describir brevemente los distintos componentes del juego. Esta descripción se puede completar con la lectura del guión de la primera práctica.

2.1. El escenario de juego

Este juego se desarrolla sobre el mismo tipo de mapa bidimensional discreto que usamos en la práctica 1. La descripción del terreno es exactamente la misma. Por otra parte, como ya se comentó sobre esta superficie existen elementos que tienen la capacidad de moverse por sí mismos o bien de ser trasladados sobre dicho mapa. Podemos clasificar a dichos elementos en dos tipos, personajes del juego y objetos. Los personajes del juego son los mismos que ya existían en la práctica 1 (aldeanos y lobos) a los que añadimos un nuevo personaje, en concreto:

- Reyes: Habitantes inmóviles del juego que están deseosos de recibir todos los regalos que el jugador tenga a bien darles. Su posición es inicialmente desconocida por el jugador, pero una vez ubicados, esta posición nunca cambia. Su número depende del tamaño del mapa. Se codifican con el símbolo 'r'.

En relación con los objetos, además de los que ya se tenían en la práctica 1, aparece un tipo nuevo, los regalos. Los cinco objetos que aparecen en el juego por lo tanto son los siguientes:

- '0': Hueso
- '1': *Bikini*
- '2': *Zapatillas*
- '3': *Llave*
- '4': *Regalo*

El hueso es útil para los lobos, el biquini para el agua, las zapatillas para los bosques, la llave para las puertas y el regalo para los reyes. La forma de uso de cada objeto debe ser descubierta por el estudiante.

2.2. Nuestro Personaje

Obviamente, nuestro personaje es el protagonista de la historia y debe enfrentarse a las adversidades que le proponga el mundo y sus habitantes.

El objetivo de esta práctica es mejorar el comportamiento del agente definido en la primera práctica incorporando procesos deliberativos con la intención de descubrir en el tanto por ciento más alto posible todos los elementos inmóviles del mapa y completar el mayor número de misiones.

Por lo tanto aparece un **nuevo objetivo**: completar el máximo número de misiones. En este juego llamaremos misión al hecho de entregar un “regalo” (objeto codificado con '4') a un “rey” (el nuevo personaje del juego).

Para cumplir los objetivos tened en cuenta que la descripción del mapa, la orientación inicial, la capacidad sensorial del personaje, las acciones disponibles, las unidades de vida y el uso de la

mochila, coincide exactamente con la descripción realizada de todos esos elementos en la primera práctica.

Un aspecto nuevo que se incorpora en la práctica 2 es que aparecen **dos nuevos sensores** para nuestro personaje. El primero de ellos es una lista de las coordenadas que ocupan los objetos etiquetados como “regalo” en el mapa de juego. El segundo denominado “tiempo” devuelve la cantidad acumulada de tiempo consumido por el agente para realizar la toma de decisiones expresada en segundos. Teniendo en cuenta que el agente no puede superar la cantidad de 300 segundos en las 20,000 iteraciones de la simulación, es relevante que en el comportamiento se considere este sensor. Los 300 segundos de tiempo máximo hacen referencia a la suma de los tiempos de cada operación de “pensar” del proceso completo. Por ejemplo, utilizando el mapa30 de la primera práctica y el modelo básico definido en el tutorial de la primera práctica, el tiempo utilizado sería de 0.04 segundos aproximadamente. El tiempo con el mapa100 sería de 0.05 segundo aproximadamente.

3. Objetivo de la práctica

La práctica tiene como objetivo diseñar e implementar un comportamiento reactivo/deliberativo a nuestro personaje. Dicho comportamiento debe mejorar el comportamiento puramente reactivo definido en la práctica 1 para descubrir el mapa original. Además, se incluye un objetivo adicional que consiste en maximizar el número de misiones. En el contexto de este juego, como ya se ha dicho una misión consiste en llevar el objeto etiquetado como “regalo” a uno de los “reyes” que hay en el juego.

Es importante tener en cuenta que hay dos restricciones en el juego:

1. El proceso anterior tiene como duración 20,000 pasos de simulación.
2. Para esos 20,000 pasos, nuestro agente no puede superar los 300 segundos de consumo de tiempo, según la idea de tiempo explicada anteriormente. Este consumo se puede consultar en el sensor “tiempo”.

Se puede partir de la versión entregada por cada estudiante en la primera práctica (ficheros jugador.cpp y jugador.hpp) con la idea de mejorarla y ampliarla.

4. El Software

Para la realización de la práctica se proporciona al alumno una implementación tanto del entorno simulado del mundo en donde se mueve nuestro personaje como de la estructura básica del agente.

Se proporcionan versiones del software para el sistema operativo Linux (Ubuntu) y Windows. Dicho software se puede encontrar en el apartado de “Material de la Asignatura” dentro de la plataforma docente de DECSAI, <http://decsai.ugr.es>.

El proceso de instalación y uso del software es similar al de la práctica 1.

Los sensores que usaremos en la práctica son:

terreno	vector <unsigned char>
superficie	vector <unsigned char>
colision	bool
mochila	unsigned char
reset	bool
vida	int
objetoActivo	unsigned char
mensajeF	int
mensajeC	int
tiempo	double
regalos	vector < pair<int,int> >

5. Método de evaluación y entrega de prácticas

5.1. Método de evaluación

A partir del comportamiento entregado por el alumno en los ficheros '**jugador.cpp**' y '**jugador.hpp**' se procederá a hacer una simulación sobre un conjunto de nuevos mapas, donde cada agente compite solo (tenéis que considerar que los otros personajes definidos en el juego pueden moverse) sobre dicho mundo. La simulación constará de 20.000 pasos, el agente podrá consumir como máximo 300 segundos y las condiciones iniciales serán las mismas para los comportamientos definidos por cada estudiante.

El agente será testado sobre un conjunto de mapas. Sobre cada mapa se obtendrán dos medidas: (a) porcentaje de coincidencia con el mapa original y (b) número de misiones completadas. Aplicado sobre el conjunto de mapas, para cada estudiante se obtendrá una media para ambas medidas. La primera de ellas supondrá el 60% de la nota inicial, mientras que la medida (b) será el restante 40%.

La nota final que cada estudiante obtendrá será la combinación de las dos puntuaciones anteriores. Esta nota será matizada tanto al alza como a la baja, atendiendo a la complejidad del comportamiento definido por el estudiante.

Obviamente, las prácticas que se detecten copiadas implicarán un suspenso para el alumno en la asignatura, y la no presentación al proceso de defensa, un cero en la práctica.

5.2. Entrega de prácticas

Se pide desarrollar un programa (modificando el código de los ficheros del simulador '**jugador.cpp**' y '**jugador.hpp**') con el comportamiento requerido para el agente. Estos ficheros deberán entregarse mediante la plataforma web de la asignatura, en un fichero ZIP que no contenga carpetas. El archivo ZIP deberá contener sólo el código fuente de estos dos ficheros con la solución del alumno y un archivo de documentación en formato PDF describiendo (en lenguaje natural o mediante gráficos) el comportamiento definido en su agente, así como el algoritmo de búsqueda utilizado con un máximo de 5 páginas.

No se evaluarán aquellas prácticas que contengan ficheros ejecutables o virus.

5.3. Fechas Importantes

La fecha fijada para la entrega de las prácticas será el 17 de abril hasta las 23:00 horas para todos los grupos. Y por supuesto la entrega se podrá hacer antes de esa fecha.