

## CONTROL EN 2D DE UN ASISTENTE ROBÓTICO PARA CIRUGÍAS LAPAROSCÓPICAS

Jesús A. Febres, Carlos E. Prieto y Yomar González

*Instituto Nacional de Bioingeniería, Universidad Central de Venezuela, Ciudad Universitaria, Av. Los  
Estudios, Los Chaguaramos, Caracas-Venezuela, <http://www.inabio.edu.ve>*

**Palabras claves:** Aprendizaje Reforzado, redes neuronales, Adaptive Heuristic Critic, asistente robótico, cirugía laparoscópica.

**Resumen.** Se presenta un sistema de control para un manipulador robótico utilizado como asistente en cirugías laparoscópicas, basado en una red neuronal artificial que utiliza como algoritmo de aprendizaje el “Aprendizaje Reforzado”. El brazo robótico posee tres articulaciones, dos activas (controlables y medibles) y una pasiva (sólo medible). La red ve como una caja negra a todo el sistema manipulador-laparoscopio, el cual es modelado en un programa CAD. El problema de control es solucionado en dos dimensiones, utilizando el software desarrollado que contiene la implementación de la red. Para ello, se cuenta con una arquitectura de red particular conocida como Crítico Heurístico Adaptativo (Adaptive Heuristic Critic – AHC). La estructura original (A. Barto, *IEEE Trans. on Systems, Man and Cyb.*, 13(5):835-846 (1983)) ha sido modificada con el fin de mejorar su rendimiento ante el problema propuesto, permitiendo el manejo de variables continuas en la salida y mejorando la discretización del espacio de las variables de entrada a la red. Al final se obtiene un sistema de control con el que se planifica la trayectoria de la cámara, sin necesidad de realizar ningún cálculo analítico de la cinemática inversa ni de la geometría del robot.

## 1 INTRODUCCIÓN

La cinemática del robot estudia el movimiento del mismo con respecto a un sistema de referencia sin considerar las fuerzas que intervienen en él, por lo que la descripción analítica del movimiento espacial del robot en función del tiempo es de sumo interés, sobretodo, las relaciones entre la posición y orientación del extremo final del robot con los valores que toman sus coordenadas articulares (Barrientos et al., 2007).

Los estudios de modelos de manipuladores robóticos, mediante la cinemática inversa, se hacen más complejos a medida que los grados de libertad (GDL) se incrementan, a tal grado que resulta imposible la solución analítica. Para la solución de la cinemática inversa se han usado aproximaciones numéricas con buenos resultados, mediante la aplicación de lógica difusa y redes neuronales (Alsina & Gehlot, 1994).

En este trabajo se desarrolló e implementó una red neuronal entrenada con aprendizaje reforzado que proporciona la configuración en coordenadas articulares que debe adoptar un manipulador para posicionar su extremo de modo que el laparoscopio que sostiene, se ubique en una determinada posición.

### 1.1 El problema de la navegación laparoscópica

El problema de la navegación laparoscópica consiste en posicionar el extremo del laparoscopio en una posición determinada con el objetivo de apuntar la cámara hacia alguna zona en específico dentro de la cavidad abdominal, con el fin de poder mostrar al cirujano los órganos de interés. La muñeca emula los movimientos que debería realizar un cirujano asistente para realizar dicha navegación. En un caso general, el lugar geométrico que se describe con estos movimientos es una esfera, la cual tiene centro en el punto de inserción de la óptica. Para nuestro modelo en dos dimensiones esta esfera pasa a ser una circunferencia. Con una muñeca pasiva la localización del extremo de la óptica se determina tan sólo mediante la posición relativa de la cámara con respecto al fulcro (Funda et al., 1995).

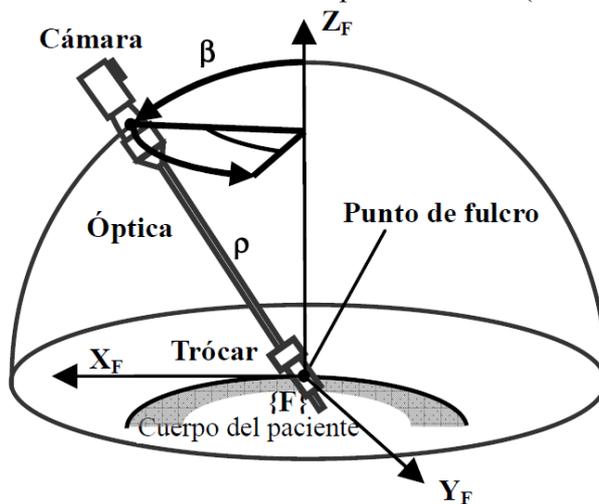


Figura 1: Navegación laparoscópica

En la Figura 1 se utiliza un sistema de referencia unido al punto de inserción, para poder describir la posición relativa de la cámara a través de las coordenadas esféricas  $\alpha$ ,  $\beta$  y  $\rho$ , esta última definida como la distancia que existe desde el centro de rotación de la muñeca pasiva hasta el punto de fulcro. Es fácil entender que tanto  $\alpha$  como  $\beta$  son valores que se

pueden calcular a partir de la lectura de los sensores internos del robot que se encuentran en la muñeca, sin embargo no ocurre lo mismo con el valor de  $\rho$ .

Para poder planificar la trayectoria, se deben conocer las localizaciones de los sistemas de coordenadas asignados a la cámara y al fulcro con respecto a la base del robot. La cámara se encuentra fija a la muñeca pasiva, y por tanto, su localización viene dada por la cinemática directa del robot. Sin embargo, el sistema de coordenadas del fulcro no se encuentra unido a ningún elemento del robot, sino que está situado en el trócar que se inserta en el cuerpo del paciente, por lo que su posición absoluta exacta resulta incierta. Esta situación ejerce una fuerte influencia en la precisión a la hora de alcanzar una coordenada esférica dada. Una localización errónea del sistema de coordenadas del fulcro, implica una imprecisión en el posicionamiento de la cámara, que conlleva a un error en los valores medidos de las coordenadas esféricas.

En (Muñoz et al , 2006) se propone una solución analítica para este problema que resulta complicado y conlleva a realizar estudios de los errores de precisión presentes, es por ello que el uso de un sistema de control basado en redes neuronales en donde el sistema a controlar se ve como una caja negra, juega un papel importante, evitándonos realizar los cálculos para determinar dichos valores. Por un lado, una vez entrenada la red neuronal, no se deben realizar los cálculos a tiempo real como en la solución analítica, lo que nos genera un ahorro en lo que a tiempo de computo se refiera, aumentando la velocidad de respuesta del sistema de control. Y por otro, se puede alcanzar la precisión deseada, ya que depende de qué tan largo sea el periodo de entrenamiento de la red.

## 1.2 Redes neuronales artificiales

Una red neuronal se parece al cerebro en dos ideas básicas: la red necesita una cantidad de información para entrenarse, y las conexiones entre las neuronas almacenan la información utilizada para tomar decisiones. El proceso utilizado para adaptar las conexiones o sinapsis se llama aprendizaje.

El modelo de neurona artificial tiene cuatro elementos básicos: un conjunto de sinapsis o enlaces caracterizados por el peso; un sumador que recibe las señales de las entradas de la neurona, el cual las pondera con los respectivos pesos sinápticos y luego las suma; una función de activación o transferencia para limitar la amplitud de la salida de la neurona y una entrada directa e independiente.

El entrenamiento de una red neuronal se asemeja a las técnicas de ajustes de curvas, donde las entradas y salidas de una función son conocidas, pero la función en sí misma no lo es. El proceso de entrenamiento de una red neuronal consiste en encontrar los valores de: los pesos sinápticos que conectan las entradas a la primera capa oculta, los pesos entre las capas ocultas, y los de la capa oculta previa a la capa de salida (Haykin, 2005).

## 1.3 Aprendizaje reforzado

Un aprendizaje mediante refuerzo trabaja con poca información del exterior. En él no existe un maestro que instruya sobre las salidas correctas ante determinadas entradas, pero sí existe un *crítico* que proporciona una información más evaluativa que instruccional. El *crítico* proporciona una evaluación, contenida en un *índice de resultados*, de las salidas calculadas por la red en función de los objetivos marcados, pero no informa sobre cuáles son las salidas correctas.

### 1.4 Crítico heurístico adaptativo.

Sutton desarrolló un algoritmo de aprendizaje por refuerzo que denominó crítico heurístico adaptativo, AHC (Adaptive Heuristic Critic). Su trabajo se enfocó principalmente en la construcción de un crítico adecuado para los problemas de aprendizaje mediante refuerzo (Sutton, 1984).

La estructura está formada por dos componentes fundamentales, el Elemento Asociativo de Búsqueda, ASE (Associative Search Element), que se encarga de asociar a cada situación de entrada un vector de salida adecuado a la misma; y el Elemento Crítico Adaptativo, ACE (Adaptive Critic Element), que realiza la función de crítico adaptativo entregando el valor que sirve como señal de refuerzo interna para el ASE. La presencia del ACE nos permite disponer de un crítico que proporcione información suficiente sobre el *índice de resultados* de una tarea, incluso en los casos en los que la evaluación externa sea esporádica en el tiempo o esté corrompida por ruido.

## 2 METODOLOGÍA

### 2.1. Algoritmo de aprendizaje

Una síntesis tal algoritmo de aprendizaje lo podemos ver en la Figura 2. Cada iteración del ciclo de aprendizaje se comienza escogiendo aleatoriamente una posición deseada  $\vec{e}$  del extremo del laparoscopio. Luego se calcula el vector de actividad  $\vec{a}$  de dicha posición.

El ACE proporciona un valor de evaluación estimado  $p$  para calcular la recompensa interna  $\hat{r}$ . Este último valor es utilizado para calcular la varianza  $\sigma$ .

La salida en cada estado, es un vector de valores continuos que representan las coordenadas articulares del manipulador, estos son determinados estocásticamente con unidades SRV (Stochastic Real Valued units) (Gullapalli, 1990). Cada nodo de salida  $i$  del ASE calcula un valor recomendado  $g^R[i]$  para la salida como la suma ponderada de los pesos  $w$ . Los valores finales de la salida  $\vec{g}^F$  se calculan haciendo uso de una distribución normal centrada en los valores recomendados  $\vec{g}^R$  y con varianza  $\sigma$ .

Con el modelo realizado en el software CAD, se consigue la posición  $\vec{c}$  del laparoscopio correspondiente a las coordenadas articulares  $\vec{g}^F$  proporcionadas por la red. Luego se calcula la recompensa externa  $r_{ext}$  tomando en cuenta la distancia entre la posición deseada  $\vec{e}$  y la posición  $\vec{c}$  calculada con la red neuronal. La función utilizada garantiza que el valor de la recompensa externa se encuentra entre 0 y 1, alcanzando su máximo cuando la distancia es cero (0).

Tomando en cuenta la señal de evaluación  $\hat{r}$  provista por el ACE y la señal de recompensa externa  $r_{ext}$ , el ASE y el ACE se actualizan, para poder adoptar las variaciones que hayan dado resultados positivos y rechazar aquellas que los den negativos.

```

función aprendizajeAHC(numInt)
  entrada:    numInt, número de iteraciones para el ciclo de entrenamiento
  estático:  w, tabla de de valores de los pesos sinápticos
                E, conjunto de vectores de posición usados en el entrenamiento

  para k ← 1 hasta numInt hacer
     $\vec{e} \leftarrow \text{ALEATORIO}[E]$  //Se escoge aleatoriamente un vector de posición e del
                                //laparoscopio
     $\vec{a} \leftarrow \text{ACTIVACIÓN}[\vec{e}]$  //Se calcula el vector de actividad a correspondiente al
                                //estado e
     $p := \sum_j v[j] * a[j]$  //El ACE calcula el valor de evaluación estimado p
     $\hat{r} := \frac{1}{1+e^{-\alpha_1 * p}}$  //El ACE calcula la recompensa interna  $\hat{r}$ 
     $\sigma := 1 - \hat{r}$  //Se calcula la varianza  $\sigma$  usada por el ASE
    para i ← 1 hasta numSal hacer
       $g^R[i] := \sum_j w[i][j] * a[j]$  //El ASE calcula el valor
      recomendado
      //de las salidas
       $g^F[i] := N(g^R[i], \sigma)$  //Calculo estocástico de las salidas
      //(coordenadas articulares)
    fin para
     $\vec{c} \leftarrow \text{POSICION\_LAPAROSCOPIO}[\vec{g}^F]$  //Se calcula la posición  $\vec{c}$  del
                                                    laparoscopio //correspondiente a las
                                                    coordenadas articulares //  $\vec{g}^F$ 
     $r_{ext} := e^{-\alpha_2 * \|\vec{e} - \vec{c}\|}$  //Se calcula la recompensa externa en función de la
    distancia
    //entre la posición deseada  $\vec{e}$  y la posición calculada  $\vec{c}$ 
    para j ← 1 hasta dimension(v) hacer
       $v[j] += \beta * (r_{ext} - \hat{r}) * a[j]$  //Se actualiza el vector  $\vec{v}$ 
      para i ← 1 hasta numSal hacer
         $w[i][j] += \sigma * (r_{ext} - \hat{r}) * (g^F[i] - g^R[i]) * a[j]$  //Se
        actualizan
        //los pesos w
      fin para
    fin para
  fin para

```

Figura 2: Algoritmo de aprendizaje utilizado

## 2.2. Modelo del sistema

El brazo robótico presentado en la Figura 3 posee tres articulaciones, dos activas (rojo-azul y azul-amarillo) y una pasiva (amarillo-laparoscopio). Cabe aclarar que, aunque el diseño esta realizado en tres dimensiones, el mecanismo solo tiene movilidad en dos de ellas.

El pequeño cilindro rojo junto con la pieza naranja conforman otra articulación que representa el fulcro y es el punto de pivote para el laparoscopio representando el lugar de inserción del instrumento en el abdomen de un paciente.

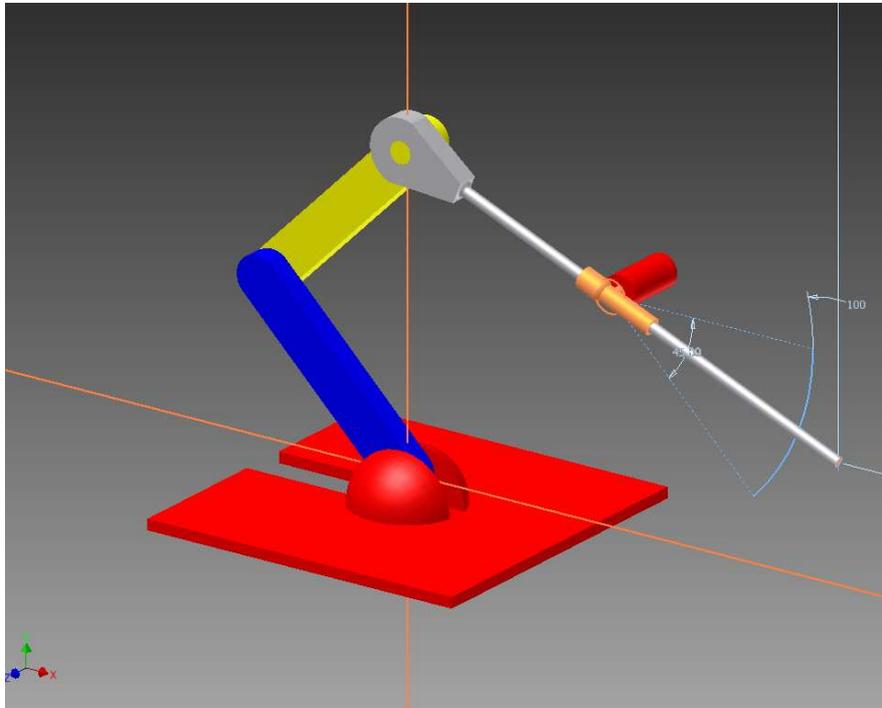


Figura 3: Modelo de brazo robótico con laparoscopio realizado en el software CAD

### 3 SIMULACIÓN Y RESULTADOS

Los valores de los parámetros utilizados en las simulaciones se muestran en la siguiente tabla:

$numInt$	$\alpha_1$	$\alpha_2$	$\beta$	$l_1(cm)$	$l_2(cm)$
1.600.000	0,1	0,2	0,8	100,0	100,0

Tabla 1 – Valores de los parámetros utilizados en la simulación. Donde  $l_1$  y  $l_2$  representan la longitud de la pieza azul y de la pieza roja respectivamente.

Una vez entrenada la red para una trayectoria determinada, se le introducen como entrada las coordenadas cartesianas de 29 puntos por donde el extremo del robot debe pasar. En las Tablas 2 se muestra el error relativo de cada una de las coordenadas de 5 de estos puntos. La trayectoria deseada y las posiciones calculadas con la red neuronal para seguir dicha trayectoria son mostradas en la Figura 4.

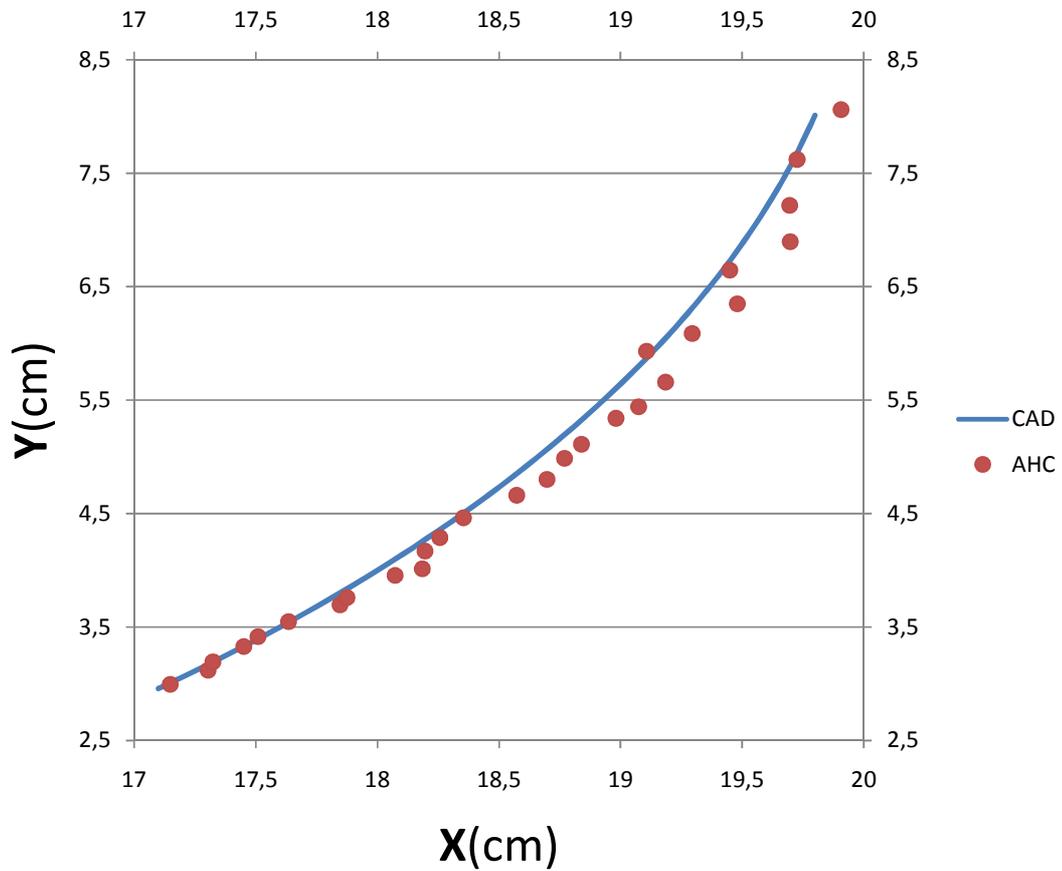


Figura 4: Gráfica de la curva teórica (CAD) y las posición calculada por la red (AHC)

<b>x_teorico(cm)</b>	<b>y_teorico(cm)</b>	<b>x_red(cm)</b>	<b>y_red(cm)</b>	<b>error_x</b>	<b>error_y</b>
17,5	3,39	17,92	3,89	0,024	0,149
18,0	4,00	18,14	4,19	0,008	0,047
18,5	4,73	18,54	4,79	0,002	0,012
19,0	5,64	19,07	5,80	0,004	0,027
19,5	6,88	19,80	8,00	0,015	0,163

Tabla 2: Valores cinemáticos teóricos y calculados con la red neuronal

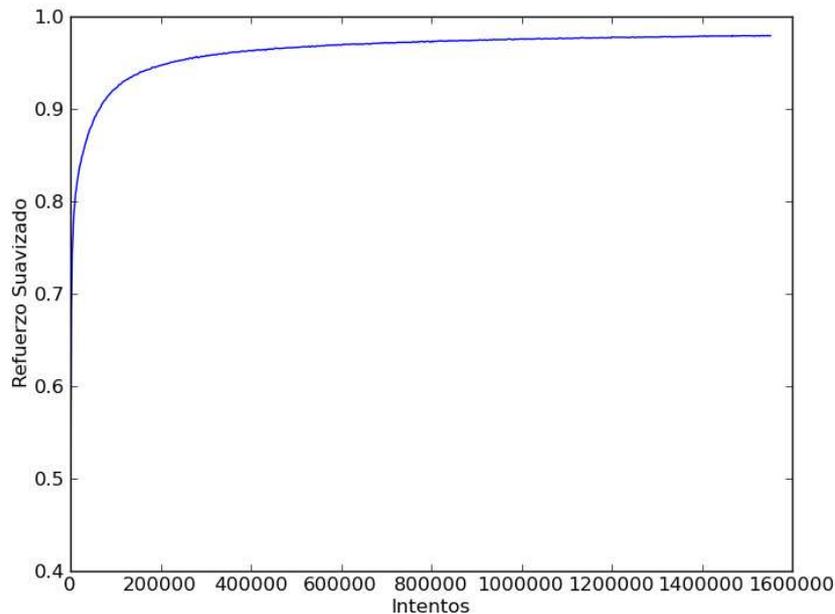


Figura 5: Índice de resultados

La eficiencia de la red fue medida considerando el valor del refuerzo suavizado. Dicho esfuerzo se calculó en cada paso de entrenamiento, como el promedio del valor de la recompensa externa en las últimas 1000 iteraciones.

En la Figura 5 se puede observar que el índice de resultados converge a 1, a medida que aumentan los intentos realizados por la red.

## 2. CONCLUSIONES Y RECOMENDACIONES

De los resultados del proceso de entrenamiento y la prueba de la red neuronal en el seguimiento de la trayectoria se encontró que la red neuronal AHC presentada, constituye una solución eficiente al problema no lineal de la cinemática inversa, consiguiendo valores del índice de resultados mayores 0,98 de manera rápida y sin necesidad de recurrir a la solución analítica. Por otro lado se debe destacar que la ubicación del punto de pivote es totalmente desconocida, sin embargo, la red logra realizar el control con suficiente precisión, sin necesidad de realizar ningún cálculo de dicha ubicación.

Así mismo, la convergencia a la solución, lo que equivale al aumento de la precisión en los cálculos de la posición, podría ser mayor si se aumentan los ciclos de entrenamientos y se considera un espacio de entrada con una discretización más fina. Sin embargo, esto acarrearía costos computacionales elevados.

Por último, se puede decir que la red entrenada puede ser utilizada para generar el muestreo de la trayectoria en coordenadas articulares que sirvan como referencias en el control dinámico, en donde se consideran fuerzas y torques.

En los posteriores trabajos se pretende modelar el fulcro en movimiento, para simular los cambios de nivel que sufre el abdomen durante una cirugía real, para luego pasar a un modelo en tres dimensiones.

**REFERENCIAS**

- Alsina, P., & Gehlot, N., Direct and inverse kinematics of robotic manipulator based on modular neural networks. *International Conference on Control, Automation, Robotics and Vision*, 3:1743-1749, 1994.
- Barrientos, A., Peñín, L., Balaguer, C. & Aracil, R., *Fundamentos de Robótica*. McGraw-Hill, 2007
- Barto, A., Sutton, R., & Anderson, C., Neuronlike elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-13(5): 835-846, 1983.
- Funda, J., Gruben, K., Eldridge, B., Gomory, S., Taylor, R., Control and Evaluation of a 7-axis Surgical Robot for Laparoscopy. *Proc. of IEEE 1995 International Conference on Robotics and Automation*, 1477-1484, 1995.
- Gullapalli, V., A stochastic reinforcement learning algorithm for learning real-valued functions. *Neural networks*, 3(6): 671-692, 1990.
- Haykin, S., *Neural Networks, A Comprehensive Foundation*. Prentice Hall, 2005
- Muñoz, V., García, I., Pérez del Pulgar, C., Gómez de Gabriel, J., Fernández, J., Control Cartesiano de un Asistente Robótico para Cirugía Laparoscopica. *Revista Iberoamericana de Automática e Informática Industria*, 3(4): 1477-1484, 2006
- Sutton, R., *Temporal Credit Assignment in Reinforcement Learning*. PhD thesis, University of Massachusetts, Amherst, 1984.