Modelado basado en Agentes; propuesta de preservación y control aplicando la Teoría Evolutiva de Lamarck con caminatas aleatorias.

Mayra Angélica Bárcenas Castro[1*], Ramón Díaz de León-Zapata[1]

[1]Instituto Tecnológico de San Luis Potosí, Av. Tecnológico s/n, Soledad de Graciano Sánchez, San Luis Potosí, C.P. 78376, México.

*e-mail: mayrabarcenas0422@gmail.com

Resumen

En este trabajo se presenta una propuesta utilizando la perspectiva de la Teoría Evolutiva de Lamarck para entender el Aprendizaje de Refuerzo en sistemas estables. Se utiliza como arquetipo de estudio un modelo de cooperación con NetLogo y las caminatas aleatorias para entender la difusión de un mensaje, el cual puede preservarse o perderse a través del tiempo, incluso llegar a convertirse en un nuevo mensaje (llamado mutación) que quizá logre beneficiar o afectar a una población. La simulación evolutiva se realizó en dos partes con tres etapas. En la primera se encontró que a pesar de que en un principio había más agentes egoístas que cooperativos aunado a una muy baja probabilidad de cooperación durante la evolución, la cooperación se preserva en las tres etapas. En la segunda simulación se aumentó la distorsión del mensaje con baja probabilidad de cooperación y más individuos egoístas que cooperativos en una población; en la primera etapa predominó el egoísmo, en la segunda emerge la cooperación, mientras que la tercera prevaleció el egoísmo. Este trabajo permite reformular planteamientos en estudios de "Vida Artificial".

Palabras clave: Teoría Evolutiva de Lamarck, Caminatas Aleatorias, Modelo de cooperación, Aprendizaje de Refuerzo, Vida Artificial.

Abstract

In this paper a proposal is presented using the perspective of Lamarck's Evolutionary Theory to understand the Reinforcement Learning in stable systems. It is used as an archetype of study a model of cooperation with NetLogo and random walks to understand the diffusion of a message, which can be preserved or lost over time, even become a new message (called mutation) that may be able to benefit or affect a population. The evolutionary simulation was carried out in two parts with three stages. In the first it was found that although initially there were more selfish than cooperative agents combined with a very low probability of cooperation during evolution, cooperation is preserved in all three stages. The second simulation increased the distortion of the message with low likelihood of cooperation and more selfish individuals than cooperative in a population. In the first stage the selfishness prevailed, in the second emerges the cooperation, while the third dominated the selfishness. This work allows to reformulate approaches in studies of "Artificial Life".

Key Words: Lamarck Evolutionary Theory, Random Walks, Cooperation Model, Reinforcement Learning, Artificial Life.

1. Introducción

Actualmente se ha venido observando fenómenos caóticos que con la implementación de métodos heurísticos comparado con métodos matemáticos, se dificulta predecir comportamientos para la toma de decisiones en diferentes áreas del conocimiento.

Uno de los temas que se ha debatido en la comunidad científica es la estabilidad de sistemas a través del tiempo; por ejemplo en el área de electrónica se estudió el análisis de la evaluación de estabilidad transitoria de un sistema de potencia utilizando el concepto de

Centro de Inercia como referencia para elegir el ángulo específico en la definición de un modelo basado en tiempo real [1]; mientras que en geografía la noción de clima, se expresa como el resultado de condiciones habituales durante un largo periodo que como mínimo se suelen establecer en treinta años para que se consideren relativamente estables o en equilibrio general dentro de un sistema climático por lo menos a escala humana [2].

En este sentido estudiar la estabilidad de un sistema sea cualquiera su naturaleza, coadyuva a la preservación o destrucción de los escenarios satisfactorios o catastróficos.

Por otro lado el aprendizaje de refuerzo es un paradigma nuevo que se está desarrollando como una importante técnica de control en sistemas estables; algunas metodologías han combinado enfoques evolutivos para mejorar su rendimiento. Otros utilizan un conjunto de reglas difusas óptimas y aprendizaje de refuerzo genético para diseñar partes consecuentes de sistemas difusos [3].

En este contexto un tema bastante relacionado y que ha aterrizado con más frecuencia en la comunidad científica es la Teoría de Lamarck como arquetipo funcional en el aprendizaje de refuerzo evolutivo y/o genético. Esta Teoría desarrolla cuatro elementos analíticos, el primero la fase transitoria evolutiva en el tiempo, el segundo la interpretación de la naturaleza como un sistema material donde el cambio orgánico da sentido al proceso continuo dirigido a la conservación de la tierra, el tercero la definición de un principio genealógico sobre el origen de las especies que identifica el método natural y cuarto la reformulación del concepto especie utilizando argumentos de relatividad temporal e inestabilidad individual [4].

Aunque dicha teoría se descartó en gran medida cuando se aplicó a organismos biológicos se ha encontrado utilidad en la disciplina de la Vida Artificial, en [5] realizaron un investigación utilizando redes neuronales y algoritmos genéticos basándose en teorías evolutivas Darwinianas y Lamarckianas. Encontraron un potencial en la implementación de la herencia del conocimiento lamarckiano incluida en biología e informática, además de que este tipo de trabajos deben permitir fomentarse.

Este trabajo pretende estudiar la estabilidad de un mensaje a través del tiempo (algunas generaciones), analizado con el modelado basado en agentes utilizando la formalidad matemática de trayectoria con caminatas aleatorias.

El modelado basado en agentes permite examinar de manera sencilla la complejidad, emergencia y no linealidad típica de fenómenos biológicos, sociales, de ingeniería entre otros, mostrando la visualización, desarrollo y evaluación con experimentos virtuales para predecir algunos escenarios [6].

Las caminatas aleatorias son un modelo de estudio para entender la difusión a nivel microscópico, por ejemplo en un estudio que se hizo con mejillones, examinaron si estos se agrupaban para sobrevivir a factores externos como las corrientes oceánicas o depredadores, de tal forma que si quedan solos serían presa fácil; supongamos que se agrupan, entonces estos compiten por los alimentos y si un mejillón tiene pocos vecinos se quedaría con la mayor parte del recurso. En este contexto en [7] propusieron un modelo basado en agentes en donde cada mejillón combina periodos de movimiento aleatorio con periodos de inmovilidad. Encontraron que con caminatas aleatorias llamadas de Lévy permiten maximizar la probabilidad de supervivencia de los mejillones y que además emerge por efectos de retroalimentación patrones colectivos de agrupamientos espacialmente no uniformes. En el caso de que se permitieran mutaciones aleatorias en el comportamiento, cualquier estado sub-óptimo se ve forzado a evolucionar en el tiempo hacia otro estado óptimo convirtiéndose en un atractor estable.

Como se puede observar para la visualización de dichos comportamientos se utilizan los algoritmos que se han convertido en un medio para dar sentido a la vida social. En [8] se ilustró un estudio en donde desafiaron a los agentes a caminar sin mapa o GPS, solo con un algoritmo desarrollado por una serie de reglas dinámicas con un código de ejecución incluida la continua reincidencia de reglas, observaciones, rigidez del procedimiento y la invocación selectiva del algoritmo como un ejemplo tangible. Se encontró que los algoritmos "no son de todo aleatorios", más bien son usados para dar sentido a las observaciones y que es un área de oportunidad de repensar problemas clave en la intersección de ciencias sociales y computacionales.

Ahora bien la caminata aleatoria en tiempo continuo es una forma de modelar los procesos con un formalismo matemático. En [9] propone un modelo que se puede usar en aplicaciones de física, química, biología, ciencias sociales entre otras.

Derivado de una caminata aleatoria en tiempo continuo para la distribución de probabilidad se tiene:

$$P(x, t + dt) = \frac{1}{2} P(x'dx, t) + \frac{1}{2} P(x + dx, t) \tag{1}$$

Donde x es el número de pasos y $t + dt$ es en el tiempo.

En este contexto para obtener el valor de un paso dada una caminata aleatoria se utiliza la expansión de la serie de Taylor en (1) con las siguientes ecuaciones:

$$P(xt) + dt\,\frac{\partial P}{\partial x} + \ldots n = \frac{1}{2}\left[P\,(x, t) - dx\frac{dP}{dx} + \frac{1}{2}dx^2\,\frac{\partial^2 P}{\partial x^2} + \cdots n\right] \tag{2}$$

$$P(xt) + dt\,\frac{\partial P}{\partial x} + \ldots n = \frac{1}{2}\left[P\,(x, t) + dx\frac{dP}{dx} + \frac{1}{2}dx^2\,\frac{\partial^2 P}{\partial x^2} + \cdots n\right] \tag{3}$$

Para caminatas aleatorias simétricas, donde la probabilidad de derecha a izquierda y viceversa es igual a la primera, los términos en orden cero se cancelan en (2) y (3), de modo que estas se consideran como partículas suspendidas en un líquido, es decir la facilidad con la que se mueve; en este contexto se utiliza la siguiente ecuación:

$$\frac{\partial P}{dt} = \frac{(dx)^2}{2\partial t}\frac{\partial^2 P}{\partial x^2} \quad \lim dx, dt \quad tal\ que\ \frac{\partial x^2}{2dx} = constante\ \equiv Coeficiente\ de\ difusión \quad (4)$$

Con la ecuación (4) se postula la evolución temporal de un sistema aislado con ciertas magnitudes un valor constante.

Dicha caminata aleatoria se mueve en una dimensión general con la siguiente ecuación:

$$\frac{\partial P}{\partial t} = L\nabla^2 P \quad (5)$$

Donde L es el Laplaciano de P.

Cabe mencionar que se tienen que tomar en cuenta dos criterios importantes; el primero consiste que solo en términos de análisis dimensional las unidades del coeficiente de difusión L por construcción, existen dos campos de longitud:

$$[L] = \frac{l^2}{t} \ ; \ l \text{ es la longitud} \quad (6)$$

El segundo es que P no tiene unidades y x^2 tiene unidades pero multiplicadas por la difusión.

$$[x^2] = [Lt] \quad (7)$$

Por lo que con cero se recupera el significado de desplazamiento de una caminata solo desde el análisis dimensional de la ecuación de difusión.

Para la solución de la ecuación de difusión, en donde se pretende encontrar todos los componentes tomados a intervalos regulares de tiempo se utiliza la Transformada de Fourier como sigue:

$$f(k,t) = \int_{-\infty}^{+\infty} f(x,t)e^{-ikx}dx \quad (8)$$

La inversa de (8) es:

$$f(x,t) = \frac{1}{2}\pi \int_{-\infty}^{+\infty} f(k,t)e^{-ikx}dk \quad (9)$$

Y la Transformada de Laplace como sigue:

$$f(x,s) = \int_0^{+\infty} f(x,t)e^{-st}dt \tag{10}$$

La inversa de (10) es:

$$f(x,t) = \frac{1}{2\pi i}\int_c f(x,s)e^{st}ds \tag{11}$$

(Por el acceso imaginario negativo)

Utilizando la Transformada de Fourier y Laplace se obtiene:

$$\int_{-\infty}^{+\infty} \left[\frac{\partial P}{\partial t} = L\ \frac{\partial^2 P}{\partial x^2}\right] e^{ikx}dx \tag{12}$$

$$\int_0^{+\infty} \left[\frac{\partial P(k,t)}{\partial t} = -Lk^2\ P(k,t)\right] e^{-st}dt \tag{13}$$

Resolviendo (12) y (13) resulta:

$$SP(k,s) - P(k,t=0) = -Lk^2\ P(k,s) \tag{14}$$

Supongamos que la condición inicial de la caminata es cero, entonces el origen de la Transformada de Fourier de una función es $\Delta = 1$ por lo tanto la distribución de probabilidad en el dominio de la Transformada de Fourier en la Transformada de Laplace se representa como:

$$P(k,s) = \frac{1}{s+Lk^2} \tag{15}$$

Invirtiendo la ecuación (15) se tiene:

$$P(k,t) = \frac{1}{2\pi i}\int_c \frac{1}{s+Lk^2}\ e^{st}ds = e^{-Lk^2t} \leftrightarrow P(x,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-Lk^2t}e^{-ikx}\ dx \tag{16}$$

Una característica destacada de la Transformada de Fourier es que la Transformada Fourier-Gauss es otra gaussiana, de modo que esta sería más fácil de resolver y en este sentido L que es la constante de la ecuación de Difusión tomará lugar en la siguiente ecuación:

$$P(x,t) = \frac{1}{\sqrt{4\pi Lt}}e^{-x^2/4Lt}\ \text{, Donde el valor de L = 1} \tag{17}$$

Por lo tanto:

$P(r,t) = \frac{1}{(4\pi Lt)L/2} e^{-r^2/4Lt}$, Donde el valor de L es arbitrario (18)

La ecuación (18) corresponde a la distribución de probabilidad radial de la posición de la caminata aleatoria de manera general.

Finalmente las ecuaciones (17) y (18) son la representación de la distribución universal gaussiana para la posición de la distribución del paseo aleatorio en el tiempo continuo, que será el proceso que se usará en este trabajo.

2. Método y caso aplicativo

NetLogo 6.0.4 es un entorno de modelado de agentes múltiples desarrollado por Uri Wilnsky que impulsa su uso para investigaciones de sistemas complejos junto con Bill Rand. Dicho software es de uso gratuito y contiene varios ejemplos de casos prácticos de éxito. En este trabajo se utilizó el modelo de “Cooperación” basado en un ejemplo de biología evolutiva. En él los agentes (vacas) compiten por recursos naturales (pasto) de tal forma que aquellas vacas con más éxito se reproducen con más frecuencia y por lo tanto tienden a preservarse a través de las generaciones. Muestra como dos categorías de agentes (vacas cooperativas y vacas egoístas) compiten entre sí para evolucionar con el tiempo. Inspirados en este modelo de forma análoga se puede replantear como un conjunto de agentes (personas cooperativas y egoístas) compiten entre sí para la transmisión de un mensaje en donde las cooperativas son aquellas que tienen un mensaje completo y las egoístas son aquellas que toman el mensaje sin importar si está completo y transmitirlo a las siguientes generaciones. En este contexto el modelo propone reglas sencillas para generar la simulación del comportamiento a través del tiempo [10].

Para realizar el modelo se tiene lo siguiente; en primer lugar el escenario consta de establecer un número inicial de personas en un espacio dimensional, donde las personas son puestas aleatoriamente. Se propone realizar tres etapas de modelado, en la primera etapa consta de 4 personas cooperativas y 46 egoístas; en la segunda 3 personas cooperativas y 47 egoístas y la tercera 2 personas cooperativas y 48 egoístas. Estos agentes interactúan entre sí en donde cada iteración consta de un tick en NetLogo; cada tick representa un día y para la simulación completa constará de 10950 ticks que corresponden a 30 años. Tome en cuenta que para el ejemplo climático se determinan 30 años como periodo mínimo que se considera relativamente estable.

Las reglas que se utilizaran son:

1.- *La probabilidad de cooperación* que establece que un agente sea cooperativo.

2.- *Longitud de paso* que es el valor que determina el movimiento de los agentes, además será tipo caminata aleatoria continua de forma radial.

3.- *Valor Energético*, cada agente adquiere la información de esa celda que ocupa actualmente y aumentará su energía por el valor del control deslizante.

4.- *Metabolismo*, en un principio un agente obtiene un 50% de energía, con el paso del tiempo el agente pierde energía establecida por el control deslizante (Valor descendiente del ciclo de vida). Si el agente pierde completamente su energía muere.

5.- *Umbral de Reproducción*, si el agente alcanza el valor de este control deslizante se reproduce. Este valor representa el éxito de recolección de energía.

6.- *Costo de Reproducción*, cada vez que un agente se reproduce pierde cierta cantidad de energía que está establecida por el control deslizante.

7.- *Oportunidad de Crecimiento bajo*. Este valor es el porcentaje de posibilidad de que la energía por debajo del umbral de crecimiento vuelva a aumentar. Cuanto mayor es el valor, menor es la discrepancia entre los comportamientos cooperativos y egoístas.

8.- *Oportunidad de Crecimiento alto*. Este valor es el porcentaje de la posibilidad de que la energía por encima del umbral vuelva a aumentar. Cuanto menor es este valor mayor es la discrepancia entre ambos comportamientos cooperativos y egoístas.

9.- *Máximo aumento de energía*. Este valor establece la longitud de energía de la celda.

10.- *Mínima energía*. Por encima de este valor la energía vuelve a crecer con una

Oportunidad de aumentar.

11.- *Alta oportunidad de aumento de energía*. Se interpreta como el valor de **aprendizaje de refuerzo** para la estabilidad.

12.- *Baja oportunidad de aumento de energía.* Es el valor probable de que se distorsione la información.

3. Resultados

A continuación se muestran los resultados modelando en NetLogo 6.0.4 la cooperación con caminatas aleatorias continuas focalizando el aprendizaje reforzado evolutivo como control y la distorsión del mismo.

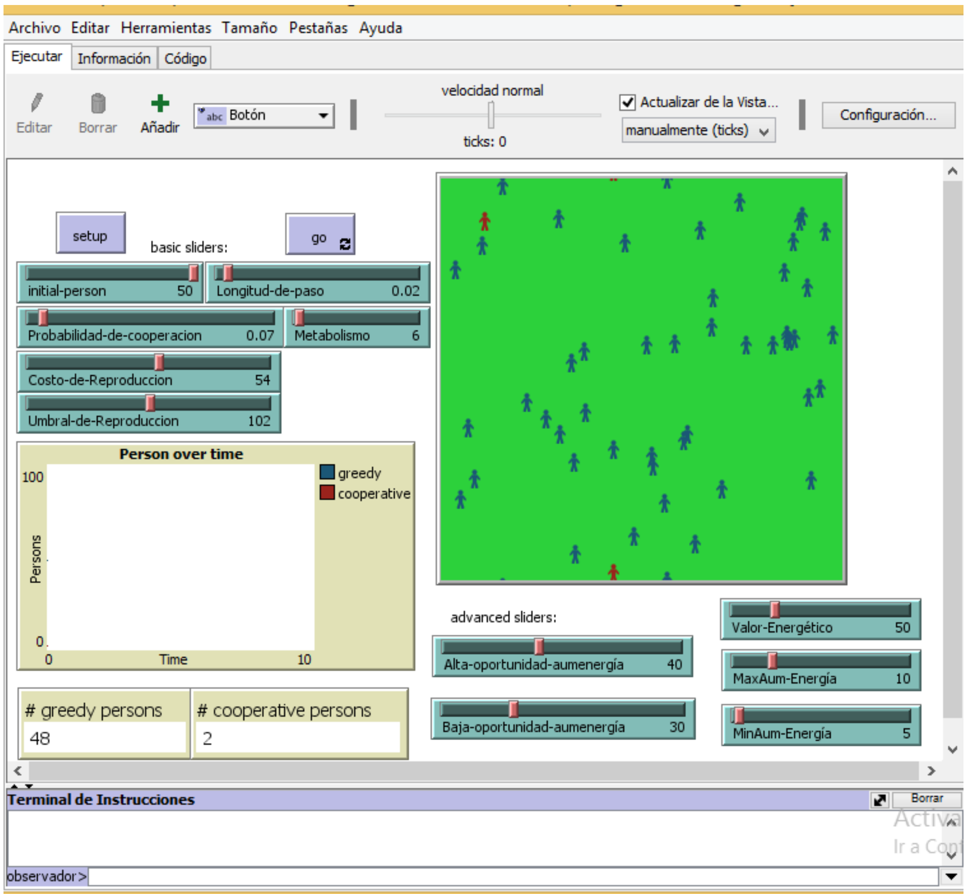
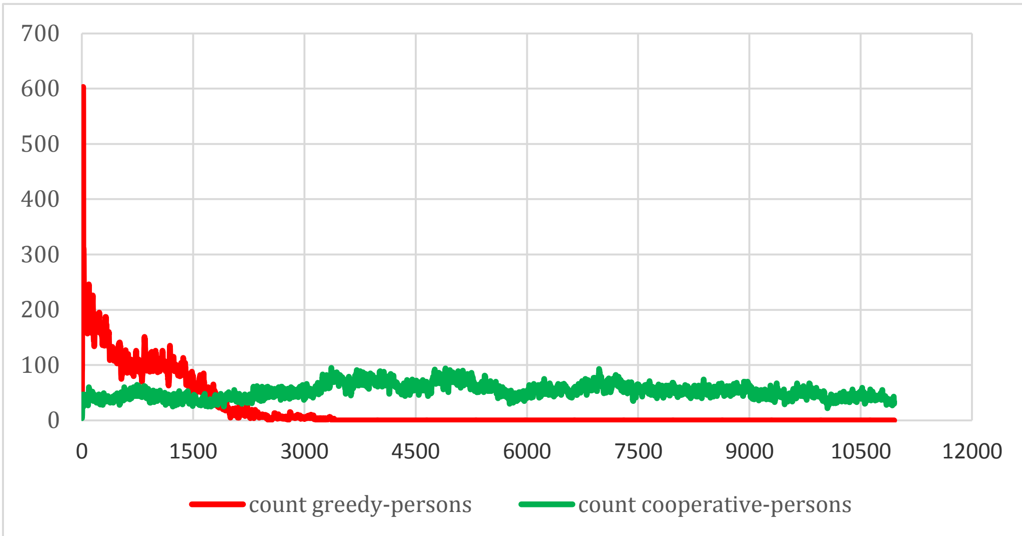
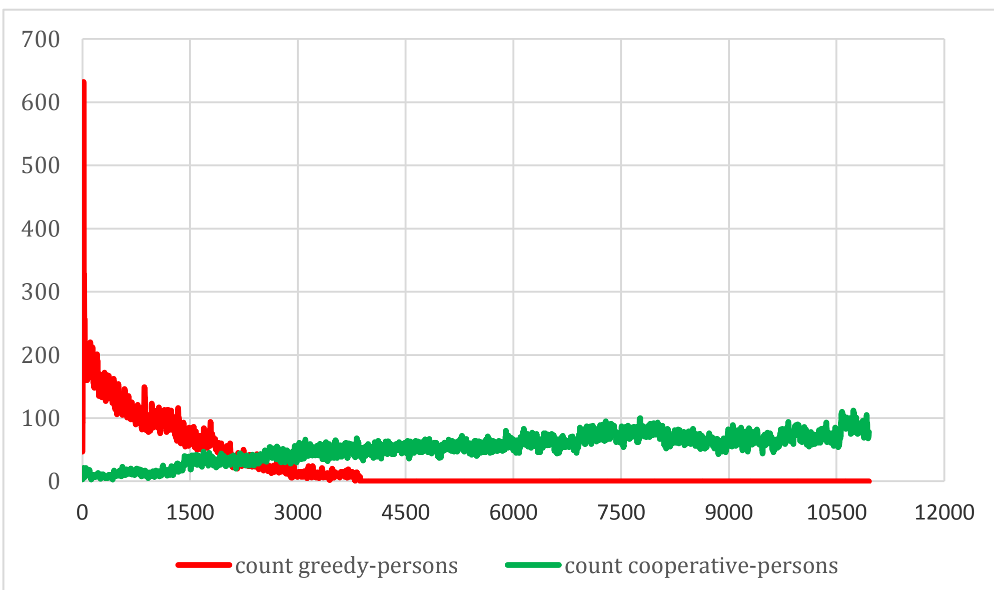


__Figura 1__. Ejemplo de los valores que se utilizaron para realizar la modelación con 2, 3 y 4 agentes cooperativos y el resto egoístas. El total de población es de 50 personas.

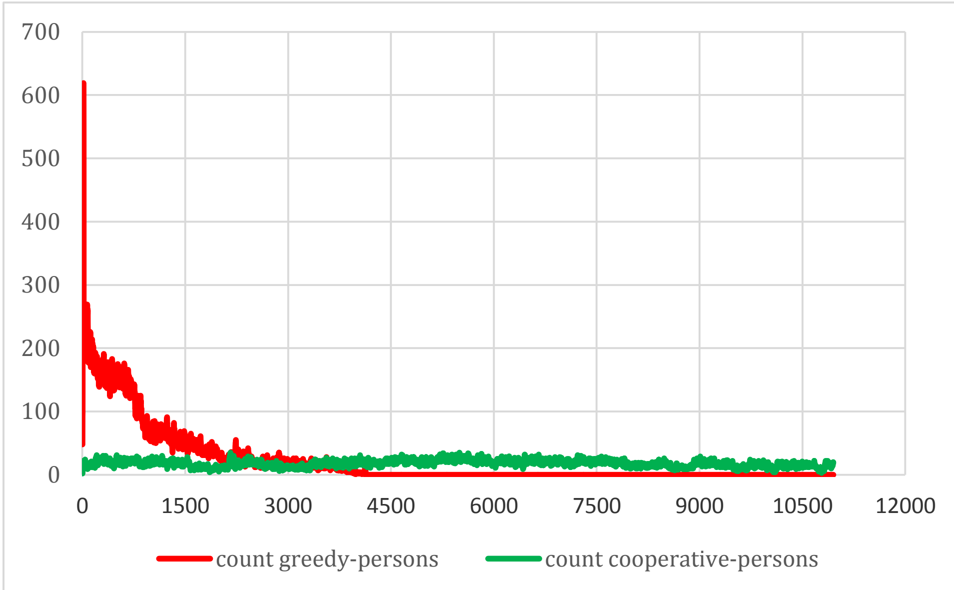
Para este ejemplo se puede observar que el aprendizaje reforzado es con una probabilidad de 40% y 30 % de distorsión.



***Gráfica 1**.- Resultado de modelar 4 agentes cooperativos y 46 egoístas por 30 años (10950 ticks) con los parámetros de la Figura 1.*



***Gráfica 2**.- Resultado de modelar 3 agentes cooperativos y 47 egoístas por 30 años (10950 ticks) con los parámetros de la Figura 1.*



***Gráfica 3.**- Resultado de modelar 2 agentes cooperativos y 48 egoístas por 30 años (10950 ticks) con los parámetros de la Figura 1.*

Se puede observar que en las tres gráficas se preserva la cooperación y que entre 5 y 10 años existe un punto de equilibrio entre los cooperadores y egoístas.

Sin embargo se observó que si se aumentamos la distorsión que casi llegue a ser similar con el aprendizaje reforzado se encontró lo siguiente:

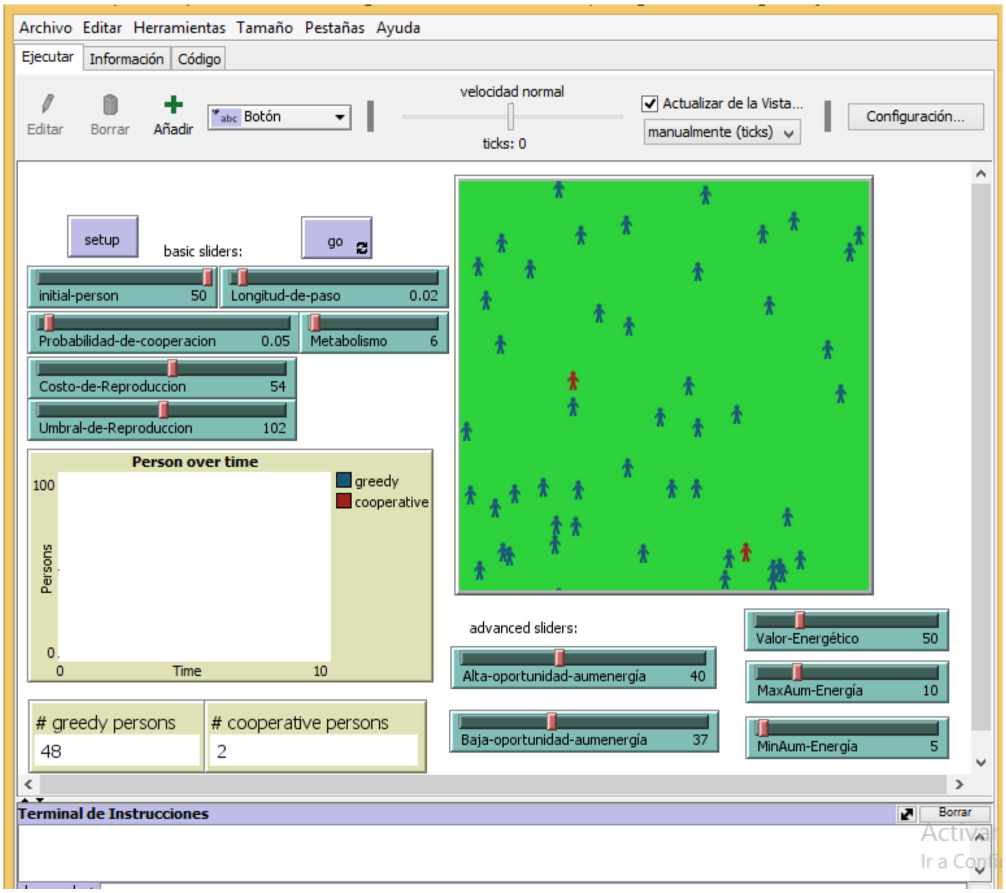
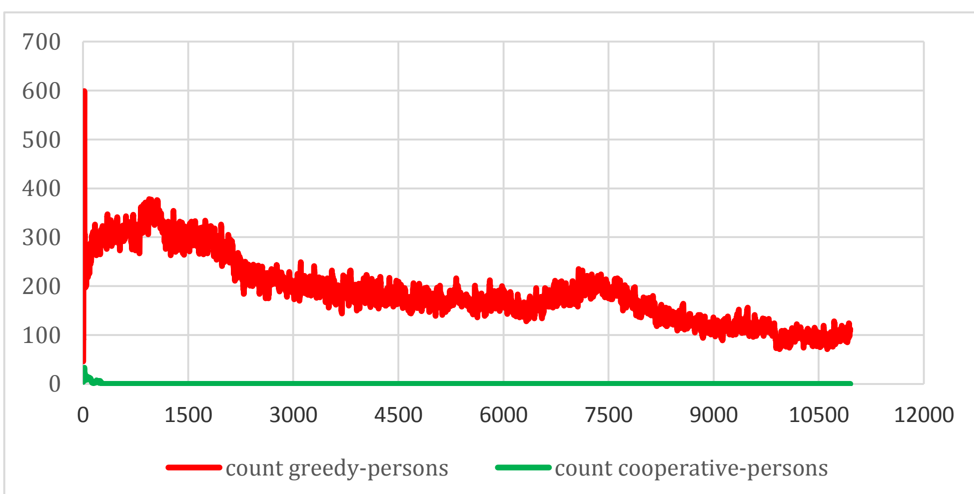
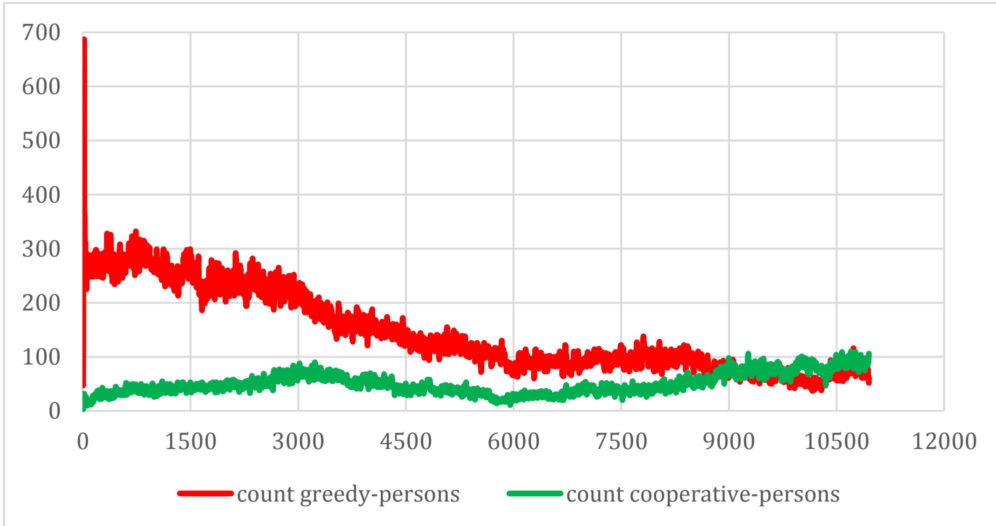


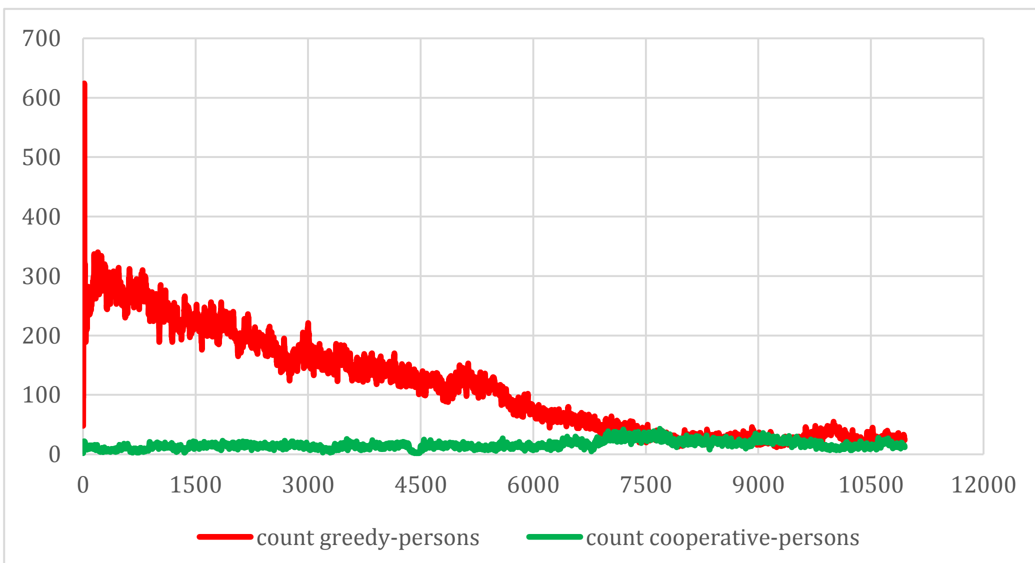
Figura 2. *Ejemplo de los valores que se utilizaron para realizar la modelación con 2, 3 y 4 agentes cooperativos y el resto egoístas. El total de población es de 50 personas.*



Gráfica 4.- *Resultado de modelar 4 agentes cooperativos y 46 egoístas por 30 años (10950 ticks) con los parámetros de la Figura 2.*



Gráfica 5.- *Resultado de modelar 3 agentes cooperativos y 47 egoístas por 30 años (10950 ticks) con los parámetros de la Figura 2.*



Gráfica 6.- *Resultado de modelar 2 agentes cooperativos y 48 egoístas por 30 años (10950 ticks) con los parámetros de la Figura 2.*

En las gráficas 4 y 6 se puede visualizar que el egoísmo predomina no así para la gráfica 5 en donde se alcanza a preservar la cooperación. También se nota que a menor cantidad de agentes cooperativos con el tiempo también el egoísmo es menor; de forma inversa a mayor agentes cooperativos mayor agentes egoístas. En este sentido se puede afirmar como en [11] en el texto "*La voz del pasado. Historia oral*" donde se argumenta que existe una problemática en el almacenamiento y conservación de evidencia oral. En el año 1988 la grabación de cinta magnética era una técnica reciente, sin embargo no quedaba claro la medida de seguridad de que la cinta fuera susceptible a desintegrarse. En este hilo conductor [11] propone dos reglas básicas para esta problemática; la primera la calidad de la cinta a utilizar para el almacenamiento debe ser cuidadosamente seleccionada; análogamente a este trabajo se puede interpretar como aquellos agentes que contienen la información del mensaje completo y la importancia de las estrategias que coadyuven a traspasar la información porque de esta forma su conservación a través del tiempo será estable. En segundo lugar se ha de tener en cuenta el lugar de almacenamiento, de modo que tratándose de un mensaje en un lugar donde no se toma en cuenta variaciones del exterior podría tener sentido si en el lugar el idioma es particular del lugar. En el caso de que llegase un externo si éste no comprende del todo el lenguaje materno entonces se distorsionará pero en medida que el nuevo agente tenga conocimiento o el nacimiento de un nuevo concepto del mensaje se le llamaría mutación.

4. Conclusión

Las caminatas aleatorias son trayectorias del formalismo matemático que permite estudiar fenómenos evolutivos. En este trabajo se presentó una propuesta utilizando la perspectiva de la Teoría Evolutiva de Lamarck para entender el Aprendizaje de Refuerzo en sistemas estables. Por otro lado este trabajo se inspiró comprender cómo la distorsión de un mensaje puede preservarse o perderse a través del tiempo, incluso llegar a convertirse en un nuevo mensaje (llamado mutación) que quizá logre beneficiar o afectar a una población, que fue el ejemplo que se utilizó para dicho procedimiento. Se encontró que a pesar de que en un principio había más agentes egoístas que cooperativos aunado a una muy baja probabilidad de cooperación durante la evolución, la cooperación se preserva. Por otro lado si aumentamos la distorsión del mensaje con baja probabilidad de cooperación y más individuos egoístas que cooperativos en una población en la primera y última etapa predominó el egoísmo, no así en la segunda donde emerge la cooperación. En este sentido se puede decir que a menor cantidad de agentes cooperativos con el tiempo también el egoísmo es menor; de forma inversa a mayor agentes cooperativos mayor agentes egoístas. Este trabajo pretende reformular estudios en informática, biología y ciencias sociales, particularmente estudios de "Vida Artificial". De este modo se invita al lector a reproducir esta metodología en fenómenos similares.

Referencias

[1] Echeverría, D. E., & Cepeda, J. C. (2018). Evaluación de Estabilidad Transitoria de Sistemas de Potencia utilizando el concepto de Centro de Inercia. Revista Técnica Energía.

[2] Pérez, M. E. (2018). Fluctuaciones climáticas y variabilidad temporal del clima en el norte argentino–1931/2005. Geográfica digital, 3(6), 1-27.

[3] Peysakhovich, A., & Lerer, A. (2018). Maintaining cooperation in complex social dilemmas using deep reinforcement learning.

[4] Galera, A. (2009). Lamarck y la conservación adaptativa de la vida. Asclepio, 61(2), 129-140.

[5] Betkher, Y., Nabais, N., & Santos, V. (2017). Impact of alife simulation of Darwinian and Lamarckian evolutionary theories. In Intelligent and Evolutionary Systems (pp. 31-43). Springer, Cham.

[6] Garcia-Valdecasas Medina, J. I. (2011). Agent-based Modelling: A New Way of Exploring Social Phenomena. REVISTA ESPANOLA DE INVESTIGACIONES SOCIOLOGICAS, (136), 91-109.

[7] de Jager, M., Weissing, F. J., Herman, P. M., Nolet, B. A., & van de Koppel, J. (2011). Lévy walks evolve through interaction between movement and environmental complexity. Science, 332(6037), 1551-1553.

[8] Ziewitz, M. (2017). A not quite random walk: Experimenting with the ethnomethods of the algorithm. Big Data & Society, 4(2), 2053951717738105.

[9] Grebenkov, D. S., & Tupikina, L. (2018). Heterogeneous continuous-time random walks. Physical Review E, 97(1), 012148.
[10] Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.
[11] Thompson, P. (1988). La voz del pasado. *Trans. J. Domingo). Valencia: Edicions Alfons El Magnànim-Institució Valenciana D'Estudis i Investigació*.