

Simulación basada en agentes en Teoría de Juegos Evolutiva: ¿pueden las metanormas evitar el colapso de una norma social?

José Manuel Galán Ordax¹, Adolfo López Paredes², Ricardo del Olmo Martínez³

¹ Área de Organización de Empresas. Escuela Politécnica Superior. Universidad de Burgos. Avenida de Cantabria, s/n. Ed.A.; 09006 Burgos. jmgalan@ubu.es

² Área de Organización y Gestión de Empresas. Escuela Técnica Superior de Ingenieros Industriales. Universidad de Valladolid. Paseo del Cauce, s/n. 47011 Valladolid. adlo@eis.uva.es

³ Área de Organización de Empresas. Escuela Politécnica Superior. Universidad de Burgos. Edificio "La Milanera" C/ Villadiego s/n.; 09001 Burgos. rdelolmo@ubu.es

Resumen

El presente artículo demuestra la utilidad del modelado basado en agentes junto con la simulación computacional y el análisis matemático para la resolución de problemas de Teoría de Juegos Evolutiva planteados desde las ciencias sociales. En el trabajo se evalúa el mecanismo de metanormas como posible método para promover cooperación mediante normas sociales entre agentes egoístas a través de los modelos de Axelrod (1986). También se sugiere la generalización de este tipo de metodología para la evaluación de otros mecanismos de cooperación/competición utilizables desde la Ingeniería de Organización.

Palabras clave: Simulación, Sistemas Multiagente, Normas, Metanormas, Teoría de Juegos Evolutiva.*

1. Introducción

En los últimos años se ha experimentado un creciente interés por los mecanismos distribuidos de funcionamiento de sistemas de diversos campos científicos: desde el modelado de sistemas humanos para la predicción de demanda de agua (López-Paredes *et al.* 2005), el control de sistemas informáticos (Clearwater 1996) o la programación de tareas en una planta (Arauzo *et al.* 2003) por poner unos ejemplos. Este interés ha establecido un vínculo entre la teoría social y la ingeniería para buscar de forma compartida el correcto entendimiento de los mecanismos que subyacen a procesos generales como son la cooperación, la negociación o la competencia. El trabajo que se presenta nace de la revisión de los mecanismos de control que se sugieren desde la teoría social para la resolución de dilemas y que podrían ser de utilidad en muchos más ámbitos.

Un dilema social se caracteriza porque decisiones que parecen tener perfecto sentido desde el punto de vista del individuo pueden generar resultados que de forma agregada resultan desfavorables para cada uno de ellos. En su formalización más simple, los dilemas sociales se pueden modelar como juegos en los que los jugadores pueden

* Este trabajo se deriva de la participación de sus autores en el proyecto de investigación financiado por el MCYT con referencia DPI2004-06590 y por el proyecto de referencia VA034/04 financiado por la Junta de Castilla y León.

cooperar o ser desleales al grupo. El dilema surge del hecho de que cualquier individuo tiene incentivos para ser desleal al grupo, independientemente de las decisiones de los otros jugadores, sin embargo la cooperación universal resulta más beneficiosa para los jugadores que la deslealtad universal. Utilizando nomenclatura propia de Teoría de Juegos, en un juego dilema todos los jugadores tienen *estrategias estrictamente dominantes* que dan como resultado un *equilibrio deficiente* (Dawes 1980).

La metodología de modelado basado en agentes siempre ha tenido una especial fascinación hacia el estudio de los dilemas sociales y ha orientado un sustancial esfuerzo a identificar condiciones en las cuales es posible sostener la cooperación en esas problemáticas situaciones (Gotts *et al.* 2003). En particular, parte de ese trabajo se ha enfocado a investigar el papel de las normas sociales y de cómo pueden ser utilizadas para generar cooperación.

Siguiendo la definición de Axelrod, entendemos que “*una norma existe en un determinado contexto social si los individuos normalmente actúan de una determinada manera, y son castigados con frecuencia si se les ve no actuar de ese modo*”¹ (Axelrod 1986). Las normas proporcionan a los cooperadores en un dilema social una ventaja crucial: la opción de castigar selectivamente a aquellos que no cooperan (Boyd y Richerson 1992).

En 1986, Axelrod escribió un pionero e influyente artículo² sobre el estudio de las normas en los dilemas sociales utilizando simulación computacional (Axelrod 1986). En ese artículo, se investigaba el papel de las metanormas (normas para seguir otras normas) como método para promover la cooperación en un modelo basado en agentes simple.

Su artículo estaba dividido en el estudio de dos modelos de teoría de juegos evolutiva implementados mediante simulación basada en agentes. En el presente trabajo se revisa los modelos propuestos por Axelrod y se restringe el ámbito de aplicación de sus conclusiones, fundamentalmente que el mecanismo de metanormas no puede garantizar el sostenimiento de una norma.

2. Los modelos de Axelrod

2.1. El juego de normas

El primer juego de Axelrod, el “Norms Game”, es una variación del Dilema del Prisionero (D.P.) para N-jugadores en su versión evolutiva. La variante consiste en que los jugadores, tras jugar un D.P. N-jugadores estándar, en caso de que algún jugador haya sido desleal, tienen la probabilidad de haberle visto y decidir si castigarle o no por ello. Este castigo tiene un doble efecto en el juego, por un lado inflige una puntuación muy negativa en el jugador desleal, por otro, el jugador que castiga también recibe un pequeño daño.

Más específicamente el juego de normas se caracteriza por la participación de 20 jugadores que deben tomar dos decisiones:

¹ Nótese el énfasis de la definición en la acción frente a la intención

² Reimpreso en *The Complexity of Cooperation: Agent based models of competition and collaboration* (Axelrod 1997)

1. Los jugadores tienen que decidir si cooperar o cometer defección al resto del grupo. Si un jugador comete defección, obtendría una puntuación para sí mismo de Tentación ($T=3$) e infligiría a cada uno de los otros jugadores una puntuación de Daño ($H= -1$). Si, por otro lado, el jugador coopera, los pagos a cada uno de los jugadores no se ven alterados.
2. Los jugadores tienen cierta probabilidad de, dada una defección efectuada por otro jugador, descubrirle. Para cada defección observada, el jugador observador ha de decidir si castiga al jugador desleal o no. Si se decide castigar, los castigadores incurrirían en un coste de cumplimiento (por hacer respetar la norma) ($E= -2$). A su vez los jugadores que cometieron defección, fueron vistos y son castigados, reciben un pago de castigo ($P= -9$) por cada jugador que después de descubrirlos decide castigarlos.

La estrategia de un jugador se define mediante su propensión a la defección (*Boldness*), y su propensión a castigar a los jugadores desleales descubiertos (*Vengefulness*)³. En el modelo de Axelrod, cada una de estas propensiones se implementa mediante una cadena de tres bits para denotar ocho niveles uniformemente distribuidos de probabilidad de 0 a 1 ($0/7, 1/7, \dots, 7/7$). Los valores que caracterizan las estrategias de cada uno de los agentes se determinan de forma aleatoria al principio de cada ejecución de una simulación.

Conviene precisar que una vez que se ha jugado el juego, cada agente tiene exactamente una única oportunidad para cometer defección y observar (y quizá castigar) cualquier deslealtad que haya tenido lugar. Una ronda se completa cuando se han jugado cuatro juegos. Después de cada ronda, se computan los pagos obtenidos por cada agente y entran en juego dos fuerzas evolucionistas:

1. Los agentes que han conseguido una puntuación por encima de la media en al menos una desviación típica se replican dos veces; los agentes que han tenido una puntuación por debajo de la media menos una desviación típica son eliminados, y el resto de los agentes pasan exactamente igual a participar directamente en la siguiente ronda (son replicado una única vez). El número de agentes se mantiene constante.
2. Cada vez que una cadena de bits se replica, cada bit tiene cierta probabilidad de ser modificado ($Mutation = 0.01$).

³ A partir de este momento utilizaremos en el documento los nombres originales en inglés de Axelrod para denominar las características de los agentes, *Boldness* y *Vengefulness*.

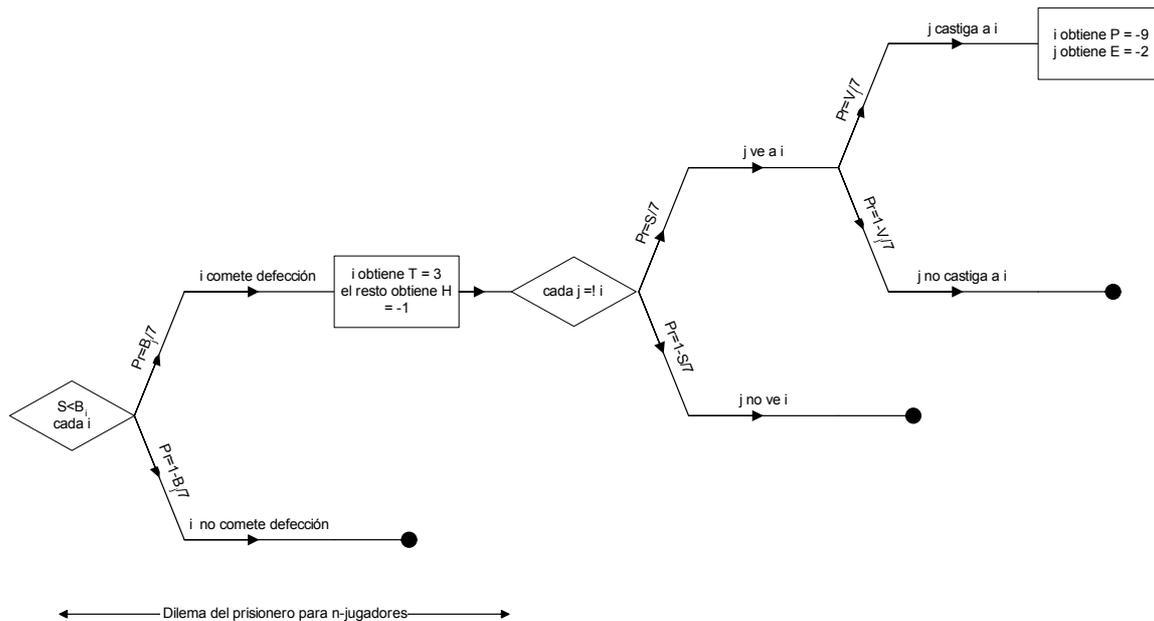


Figura 1. Estructura de decisión del juego de normas

Los resultados de este modelo no dejan lugar a la duda, la dinámica del proceso lleva indefectiblemente al colapso total de la norma y a la mutua deslealtad y la ausencia de castigo en casi la totalidad de los jugadores. La explicación a este hecho no es difícil; puesto que castigar la defeción es costoso y no proporciona a quien castiga ninguna ventaja diferencial, los castigadores están evolutivamente en desventaja (excepto cuando nadie tenga valores de *Boldness* mayores de 0, en cuyo caso el ser vengativo o no supone ninguna diferencia). Por tanto, cuando las fuerzas evolucionistas entran en acción en el sistema, deberemos esperar que la *Vengefulness* agregada tienda a decrecer. Cuando el nivel de *Vengefulness* es suficientemente bajo, los agentes con altos valores de *Boldness* tienden a conseguir mayores puntuaciones y el sistema tiende hacia el estado de colapso de la norma.

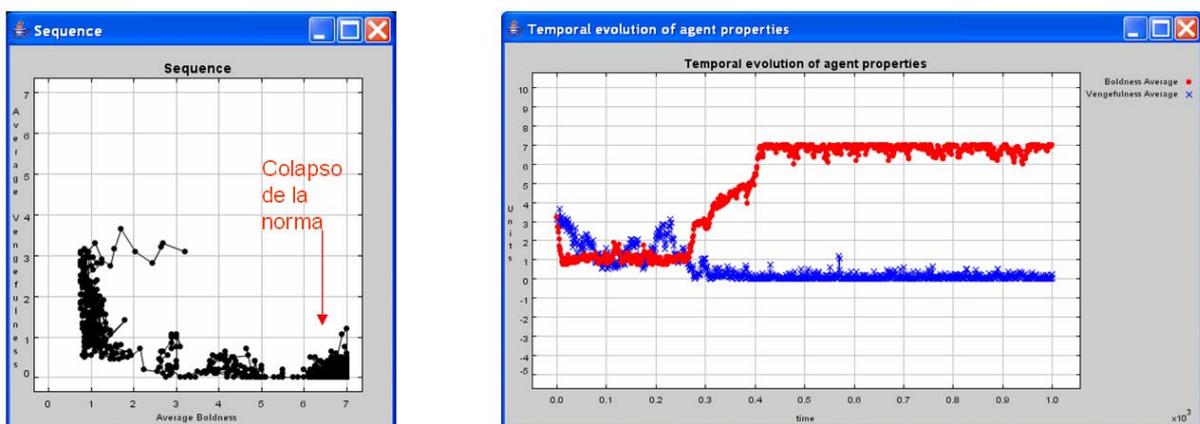


Figura 2. Dinámica del modelo de normas

2.2. El juego de metanormas

Habiendo concluido que la norma de cooperación colapsa en el modelo previo, Axelrod investiga el papel de las metanormas como método para hacer respetar las normas. La

metanorma dicta que uno debe castigar a aquellos que no siguen la norma (i.e. a aquellos que no castigan a los desleales descubiertos). En el Juego de Metanormas, la propensión a castigar a aquellos que no cumplen con la norma (meta-castigo) es la misma que la propensión a castigar a quien comete defeción⁴. Los Meta-castigadores incurren en un coste de *Meta-cumplimiento* ($ME = -2$) cada vez que ellos Meta-castigan ($MP = -9$) a alguien al que han visto que ha descubierto pero no castigado a un desleal. Un agente que observa una determinada defeción y no la castiga tiene una probabilidad de ser visto por cada uno de los otros 18 agentes.

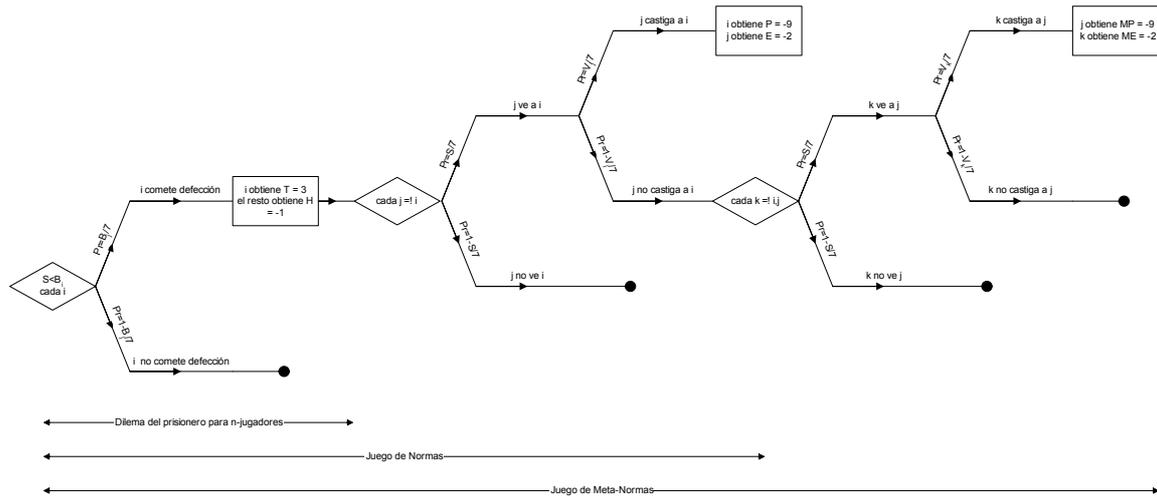


Figura 3. Estructura de decisión del juego de metanormas

Los estudios preliminares de Axelrod sugieren que las metanormas, esta nueva dimensión incluida en el segundo modelo de castigar a quien no castiga el incumplimiento de la norma, garantiza, si las condiciones iniciales son suficientemente adecuadas, que la norma social se establezca y no colapse.

3. Revisión de los modelos de Axelrod

Para continuar con nuestro estudio conviene introducir la definición de un punto de estabilidad que llamaremos *Estado Evolutivamente Estable* (ESS) (Izquierdo y Galán 2004), y que está basado fundamentalmente en las ideas iniciales de Maynard Smith y Price (1973) y posteriormente desarrolladas por Weibull (1995) y Colman (1995). Un ESS es un estado (determinado por los niveles de Boldness y Vengefulness de cada agente) donde se cumplen simultáneamente tres condiciones:

- a) Cada agente en la población Θ recibe el mismo payoff (puntuación) esperado, (por lo que la presión evolutiva de la selección no alejará al sistema del estado),

$$\text{Exp}(\text{Payoff}_i) = \text{Exp}(\text{Payoff}_j) \quad \forall i, j \in \Theta \quad (1)$$

⁴ Yamagishi and Takahashi (1994) utilizan un modelo similar al de Axelrod, pero proponen un vínculo entre cooperación (bajo nivel de *Boldness*) y *Vengefulness*. Ese simple vínculo es suficiente para alcanzar un estado evolutivamente estable donde la norma queda establecida.

- b) Cualquier agente individual m (mutante) que cambie su estrategia (sea b_m su nuevo nivel de *Boldness* y v_m su nueva *Vengefulness*) obtiene estrictamente una puntuación esperada más baja que cualquier otro agente de la población incumbente I , (por lo que una única mutación individual en un agente no será capaz de invadir la población),

$$\text{Exp}(\text{Payoff}_m) < \text{Exp}(\text{Payoff}_i) \quad \forall m \in \Theta; \forall i \in I (m \notin I); \forall b_m, v_m \quad (2)$$

- c) Y tras la aparición de cualquier agente mutante individual m que modifique su estrategia, todos los agentes de la población incumbente I obtienen el mismo payoff esperado (por lo que ningún mutante simple puede distorsionar la composición de la población excepto quizá mediante por movimientos aleatorios),

$$\text{Exp}(\text{Payoff}_i) = \text{Exp}(\text{Payoff}_j) \quad \forall i, j \in I (m \notin I); \forall m \in \Theta; \forall b_m, v_m \quad (3)$$

Las tres condiciones proporcionan fuertes restricciones sobre la estabilidad de la dinámica del modelo, sin embargo ni siquiera son suficientes para garantizar con certeza que llegado a un punto de estabilidad el sistema no escapará de él. Conviene recordar que las restricciones están formuladas en términos de esperanzas matemáticas y que la componente estocástica del sistema provoca que, ni aunque se pudiese garantizar una sola mutación individual, la vuelta al estado de estabilidad esté asegurada.

Asumiendo continuidad en el modelo, se puede demostrar que en el caso del modelo de normas la función de pagos de un agente i en una ronda es

$$\text{Exp}(\text{Payoff}_i) = T \cdot b_i + H \cdot \sum_{\substack{j=1 \\ j \neq i}}^n b_j + E \frac{v_i}{2} \sum_{\substack{j=1 \\ j \neq i}}^n b_j^2 + P \frac{b_i^2}{2} \sum_{\substack{j=1 \\ j \neq i}}^n v_j \quad (4)$$

Y aplicando la siguiente condición necesaria de ESS

$$\left\{ \begin{array}{l} \frac{\partial \text{Exp}(\text{Payoff}_m)}{\partial b_m} = \frac{\partial \text{Exp}(\text{Payoff}_i)}{\partial b_m} \quad \forall i \in I \\ \text{OR} \left(b_m = 1 \text{ AND } \frac{\partial \text{Exp}(\text{Payoff}_m)}{\partial b_m} \geq \frac{\partial \text{Exp}(\text{Payoff}_i)}{\partial b_m} \quad \forall i \in I \right) \\ \text{OR} \left(b_m = 0 \text{ AND } \frac{\partial \text{Exp}(\text{Payoff}_m)}{\partial b_m} \leq \frac{\partial \text{Exp}(\text{Payoff}_i)}{\partial b_m} \quad \forall i \in I \right) \end{array} \right\} \forall m \in \Theta \quad (5)$$

Entonces el único punto ESS es el de colapso de la norma como cabía esperar.

En el caso del juego de metanormas, aplicando de nuevo la condición necesaria de estabilidad evolutiva (5), pero ahora a su función característica de puntuación esperada (6), nos descubre que existen dos puntos susceptibles igualmente de ser considerados evolutivamente estables; ahora existe un punto de estabilidad en la zona de colapso de la norma y otro en la zona de establecimiento de la norma.

$$\begin{aligned} \text{Exp}(\text{Payoff}_i) = & T \cdot b_i + H \cdot \sum_{\substack{j=1 \\ j \neq i}}^n b_j + E \frac{v_i}{2} \sum_{\substack{j=1 \\ j \neq i}}^n b_j^2 + P \frac{b_i^2}{2} \sum_{\substack{j=1 \\ j \neq i}}^n v_j + \\ & + ME \frac{v_i}{4} \sum_{\substack{k=1 \\ k \neq i}}^n \sum_{\substack{j=1 \\ j \neq i, k}}^n b_k^3 (1 - v_j) + MP \frac{1 - v_i}{4} \sum_{\substack{k=1 \\ k \neq i}}^n \sum_{\substack{j=1 \\ j \neq i, k}}^n b_k^3 v_j \end{aligned} \quad (6)$$

Para resolver la dinámica del modelo recurrimos a la simulación computacional como método de verificación de los resultados obtenidos analíticamente.

Los resultados confirman tanto en el caso de las normas como en el de las metanormas lo esperado. Los resultados de Axelrod son consistentes con los nuestros en el caso de las normas (ver Figura 5), no así en el caso de las metanormas (ver Figura 6). Se puede ver que en el caso de las metanormas el sistema salta de un punto de estabilidad a otro sin que se pueda asegurar la permanencia en uno de ellos. Sí que es cierto, como muestra la figura, que normalmente en las primeras iteraciones la dinámica del modelo tiende inicialmente hacia el punto de establecimiento.

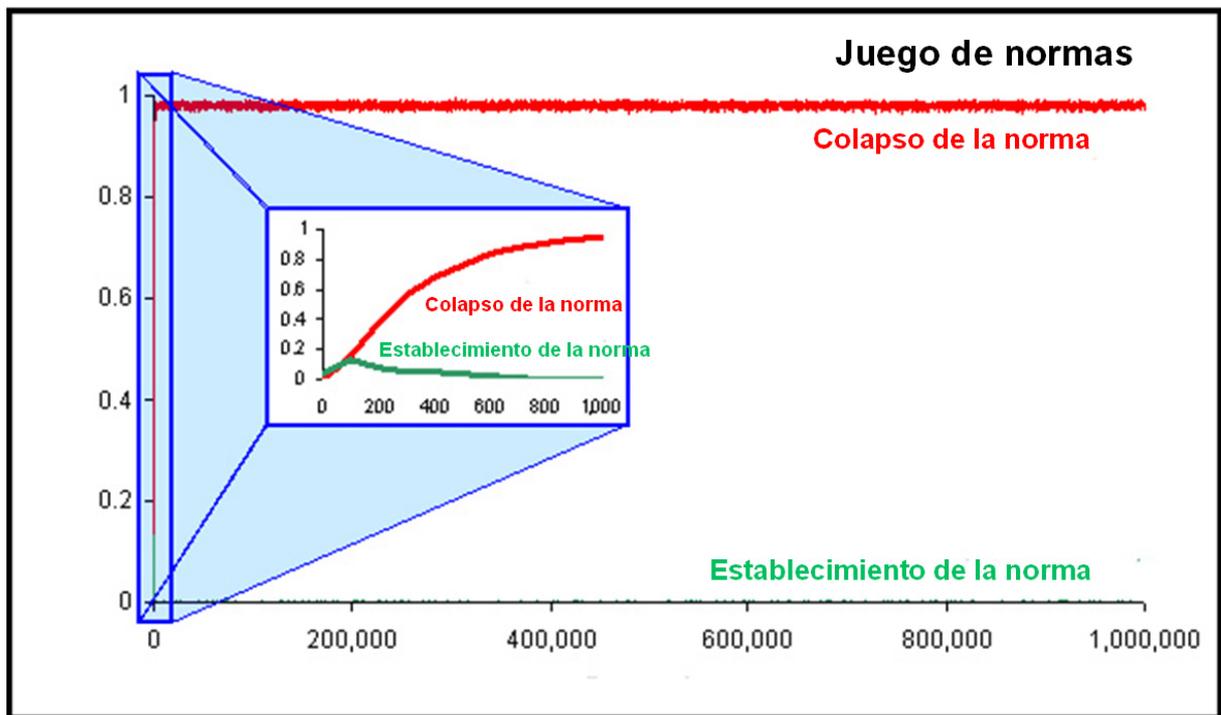


Figura 4. Porcentaje de tiempo de permanencia media del modelo de normas en las zonas de establecimiento y de colapso de la norma después de ejecutar el modelo 1000 veces

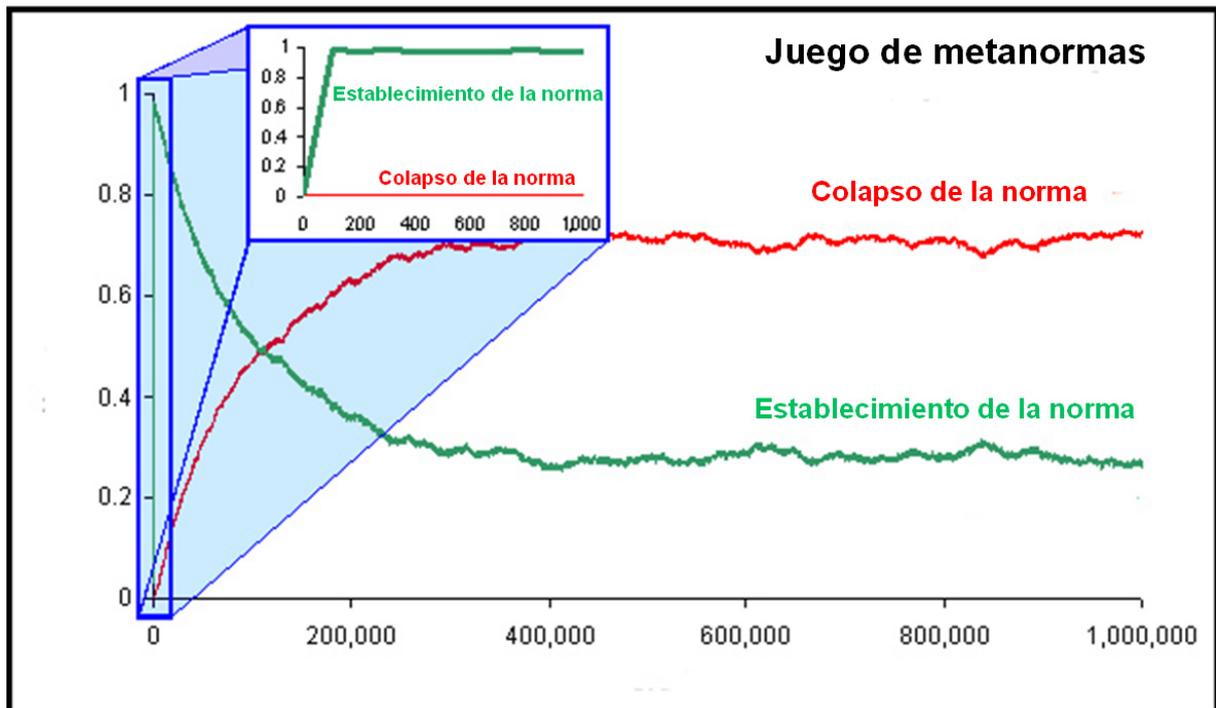


Figura 5. Porcentaje de tiempo de permanencia media del modelo de metanormas en las zonas de establecimiento y de colapso de la norma después de ejecutar el modelo 1000 veces

4. Conclusiones

En este trabajo se demuestra como la combinación del análisis matemático de modelos simplificados unidos a la simulación computacional mediante mecanismos evolutivos pueden dar respuesta a problemas propuestos desde la teoría social e inicialmente planteados como modelos de Teoría de Juegos. En este caso específico se confirma la robustez teórica de los resultados obtenidos por Hardin (1968) en la famosa *Tragedia de los Comunes*; incluso a pesar de dotar a los jugadores de la posibilidad de castigar, el sistema tiende indefectiblemente a la mutua deslealtad en el caso del juego de normas. E incluso aunque mecanismos como el de metanormas pueden llevar inicialmente⁵ al sistema a puntos de mutua colaboración no pueden garantizar completamente su permanencia.

Desde este trabajo también se sugiere la utilización extensiva de esta metodología para el análisis de mecanismos de cooperación y competición que puedan ser utilizados desde ámbitos más amplios de la Ingeniería de Organización y el diseño institucional.

Agradecimientos

Los autores agradecen la inestimable ayuda de Luis R. Izquierdo del Macauley Intitute de Aberdeen (Escocia) para la elaboración del presente estudio, sin cuyo trabajo habría sido imposible realizar.

⁵ Quizá en determinadas aplicaciones este resultado pueda ser suficiente para ser utilizado

Referencias

- Arauzo, J.A.; Benito, J.J.; Del Olmo, R. (2003) El Control De Planta: Una Solución Basada En Agentes. En Benito J J, Redondo A, (Editores). *Actas Del Congreso De Ingeniería De Organización 2003* / V.S. Merino.
- Axelrod R.M. (1986) An Evolutionary Approach To Norms. *American Political Science Review*, Vol. 80 No 4, Pp. 1095-1111
- Axelrod R.M. (1997) *The Complexity Of Cooperation. Agent-Based Models Of Competition And Collaboration* . Princeton, N.J / Princeton University Press.
- Boyd, R.; Richerson P.J. (1992) Punishment Allows The Evolution Of Cooperation (Or Anything Else) In Sizable Groups. *Ethology And Sociobiology*, Vol. 13, Pp. 171-195
- Clearwater, S. (1996) *Market-Based Control: A Paradigm For Distributed Resource Allocation* / World Scientific.
- Colman, A.M. (1995) *Game Theory And Its Applications In The Social And Biological Sciences* , 2nd Edition Ed. Oxford, UK./ Butterworth-Heinemann.
- Dawes, R.M. (1980) Social Dilemmas. *Annual Review of Psychology*, Vol. 31, Pp. 161-193
- Gotts, N.M.; Polhill, J.G.; Law A.N.R. (2003) Agent-Based Simulation In The Study Of Social Dilemmas. *Artificial Intelligence Review*, Vol. 19 No 1, Pp. 3-92
- Hardin, G. (1968) The Tragedy Of The Commons. *Science*, Vol. 162, Pp. 1243-1248
- Izquierdo, L.R.; Galán, J.M. (2004) Appearances Can Be Deceiving: Lessons Learned Re-Implementing Axelrod's "Evolutionary Approach To Norms". *Proceedings of the 2nd European Social Simulation Association Conference & 2nd Model To Model Workshop*.
- López-Paredes, A.; Saurí, D.; Galán, J.M. (2005) Urban Water Management With Artificial Societies Of Agents. The FIRMABAR Simulator. *Simulation*, Vol. 11 No 3
- Maynard Smith, J.; Price, G. (1973) The Logic Of Animal Conflict. *Nature*, Vol. 246 No 2, Pp. 15-18
- Weibull, J.W. (1995) *Evolutionary Game Theory*. Cambridge, MA: MIT Press.