UACM
Universidad Autónoma de la Ciudad de México
Nada humano me es ajeno

COLEGIO DE CIENCIAS Y HUMANIDADES

MAESTRÍA EN DINÁMICA NO LINEAL Y SISTEMAS COMPLEJOS

Toma de decisión en un modelo de votación: un enfoque de sistemas complejos

TESIS QUE PARA OBTENER EL GRADO DE MAESTRO EN DINÁMICA NO LINEAL Y SISTEMAS COMPLEJOS

P R E S E N T A
JOSÉ JOAQUÍN URBINA ALONSO

D I R E C T O R

Dr. Ricardo Lino Mansilla Corona

Ciudad de México, junio de 2018.

SISTEMA BIBLIOTECARIO DE INFORMACIÓN Y DOCUMENTACIÓN

UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO
COORDINACIÓN ACADÉMICA

RESTRICCIONES DE USO PARA LAS TESIS DIGITALES

Director

Dr. Ricardo Lino Masilla Corona

Lectores

Dr. Felipe Humberto Contreras Alcalá

Dr. Lester Alfonso Augusto Díaz

Dr. Edgar Acatitla Romero

Dr. Jorge Zaragoza Badillo

19.junio.2018

*«...es preciso aprovechar la dinámica compleja del cambio social en nuestro tiempo
y tender los puentes necesarios entre la ciencia y el humanismo
porque la separación de ambos campos no sólo es artificial
sino esencialmente errónea y engañosa.
Construyamos los puentes para educarnos en una nueva cultura, aprendamos
a ser científicos conscientes de nuestra responsabilidad social y
hagamos de la práctica de la ciencia una forma de ser humanistas.»*

— Germinal Cocho, Ciencia • Humanismo • Sociedad (**Cocho *et al.* , 2017**)

Agradecimientos

A el colegio de profesores del Posgrado en Ciencias de la Complejidad de la Universidad Autónoma de la Ciudad de México por las facilidades y el apoyo para llevar a buen término el presente trabajo.

Un agradecimiento especial a Germinal Cocho, por sus enseñanzas en el paradigma de los Sistemas Complejos y su labor científico - humanista.

Mi más sincero agradecimiento a mi director de tesis Ricardo Mansilla Corona por su respaldo en la realización de este proyecto.

A mis lectores Felipe Humberto Contreras Alcalá, Lester Alfonso Augusto Díaz, Edgar Acatitla Romero y Jorge Zaragoza Badillo por sus valiosas observaciones y comentarios en la revisión de la tesis.

A mis padres y a todos mis familiares por su apoyo a lo largo de este proceso.

Finalmente agradezco a mis amigos Edgar Acatitla, Jorge Zaragoza, Gustavo Carreón y Raymundo Vite del Seminario de Complejidad y Economía por la retroalimentación y las sesiones interdisciplinarias que contribuyeron a la mejora del presente trabajo.

Dedicatoria

Para tres personas indispensables:

Nicté-Há, Inda-Jani y Gabriela,

parámetros esenciales de mi existencia.

Índice general

Índice de figuras

Introducción

Estudiar la toma de decisiones ha significado uno de los grandes retos para la ciencias, siendo este un problema que subyace a la naturaleza de las personas y los colectivos, se han buscado explicaciones diversas desde campos de estudio como la economía, la sociología , la política , la sicofísica, etc. Estudiar la toma de decisiones nos lleva de manera natural a investigar de forma interdisciplinaria sus orígenes.

El presente trabajo consiste de una propuesta de modelación empleando el enfoque de los sistemas complejos para un caso particular del fenómeno referido como la toma de decisiones electorales.

Nuestro estudio integra un análisis de las preferencias de voto, así como un modelo computacional que puede simular la interacciones de un grupo de votantes y observar sus preferencias electorales a lo largo de un año en un escenario con tres opciones de voto.

La cualidad de estudiar tres opciones de voto en el modelo es relevante pues como lo mencionaremos más adelante, las propuestas hasta ahora desarrolladas involucran modelos para dos opciones, las cuales son adecuadas para sociedades (por ejemplo anglosajonas) donde por lo general las contiendas electorales se disputan entre republicanos o demócratas; sin embargo esta situación ha sido diferente en el caso de México.

Las preguntas de investigación que surgen son:

¿En un escenario con tres opciones de voto, se ordenan las preferencias de los votantes?, si es así ¿Podemos vislumbrar un ganador del proceso electoral?.

Nuestra hipótesis es: dado que las interacciones entre los votantes influyen en el proceso de toma de decisión, no hay posibilidad de que se ordenen las preferencias y se logre predecir a un ganador.

Existe abundante bibliografía sobre el tema de la toma de decisiones. Como antecedentes del problema tenemos el estudio de los procesos de votación, la fundamentación sobre racionalidad completa y su contra-parte, la racionalidad limitada; la segunda parte de la tesis la abocamos a presentar los conceptos sobre los sistemas complejos y finalmente desarrollamos nuestra propuesta de modelación junto con el algoritmo computacional empleado para analizar los datos obtenidos.

Dentro del marco teórico de la Teoría de la Elección Racional haciendo el símil con el comportamiento económico, tomamos las ideas que se emplearon para estudiar la conducta social en la toma de decisiones electorales. De esta forma retomamos el modelo de Downs y el modelo de Wittman. Después hacemos una revisión al Teorema de Arrow, el cual destaca por que con este resultado aparece la primera fundamentación matemática de los métodos de agregación de los intereses individuales.

Sigue la sociofísica, donde presentamos el trabajo de Serge Galam respecto a los modelos electorales, que si bien aporta un enfoque desde la física, observamos que la metodología para modelar considera los mismos supuestos en la conducta de los votantes que en los modelos prevalecientes en el estado del arte, es decir se cambia la herramienta de análisis pero no cambian los supuestos.

Destacan también los trabajos de Simon con su concepto de racionalidad más cercano al que observamos en la realidad: sin óptimos racionales, con estrategias heurísticas y en contraposición con el concepto idealizado por la teoría económica predominante, donde los agentes racionales poseen información completa y capacidad de calcular siempre la elección óptima a sus necesidades.

Su aportación es clave en los trabajos de Brian Arthur quien por primera vez introduce modelos donde las personas interaccionan empleando las ideas de la racionalidad limitada y estudia propiedades emergentes en la interacción de las personas.

La teoría de los sistemas complejos se desarrolla en el segundo capítulo, aquí aclaramos que en el desarrollo de la tesis lo complejo y lo complicado no son sinónimos; adicionalmente encontraremos los conceptos relacionados a la complejidad desde el punto de vista del estructuralismo dinámico.

En cuanto a la metodología de nuestro trabajo, emplearemos el enfoque de abajo hacia arriba (bottom ↑ up) para modelar con agentes las interacciones de un grupo de votantes en un periodo previo a una elección, con ello obtendremos información de el estado de las preferencias electorales en un lapso de un año, reuniremos esa información para aplicar un análisis de información mutua, estudiaremos la correlación de preferencias a las tres opciones de voto y analizaremos los resultados.

A diferencia de los métodos estadísticos donde se emplean encuestas para registrar una "intención" de voto, nuestra metodología permitirá obtener resultados que involucran la preferencia y las interacciones de los votantes, de esta forma aportaremos ideas para un modelo de toma de decisión electoral cuya preferencia mayoritaria evoluciona con el tiempo.

Finalmente, mencionaremos que como resultado del análisis de los datos encontraremos correlaciones no lineales entre las preferencias de los tres partidos, y aunque la *correlación no implica causalidad* (Cellucci, 2005) con esto mostraremos que el fenómeno en cuestión posee las características de un sistema complejo (Solé, 1993) y descartamos la posibilidad de que el resultado de un proceso electoral como el que se analizó sea predecible.

1. Teoría de la elección racional

La teoría de la elección racional, también conocida como teoría de la acción racional, es un marco teórico que se utiliza para entender y modelar formalmente el comportamiento social y económico donde se supone que el individuo o agente tiende a maximizar su utilidad-beneficio y a reducir los costos o riesgos. Para ello, los individuos prefieren más de lo bueno y menos de lo que les cause mal.

Esta teoría se ha usado para interpretar los fenómenos políticos a partir de supuestos básicos que derivan de principios de la economía: el comportamiento de los individuos en el sistema político es similar al de los agentes en el mercado, siempre tienden a maximizar su utilidad o beneficio y a reducir los costos o riesgos. Esta racionalidad tiene que ver con una cierta intuición que lleva a los individuos a optimizar y mejorar sus condiciones.

El actor individual es la unidad de análisis de esta teoría. Se asume que todos los individuos son egoístas, y todo individuo tiene la capacidad racional, el tiempo y la independencia emocional necesarias para elegir la mejor línea de conducta desde su punto de vista. Por tanto, todo individuo se guía racionalmente por su interés personal, independientemente de la complejidad de la elección que deba tomar. Esto no implica necesariamente que estos supuestos sean ciertos, es decir, que los individuos reales se comporten de esa manera en todo y cada momento.

Anthony Downs fue el pionero en la aplicación de los criterios económicos al comportamiento electoral. Otros autores que impulsaron esta corriente dentro de la ciencia política moderna fueron Mancur Olson, Kenneth Arrow, James M. Buchanan, Raymond Aron, Gordon Tullock y William Ricker (Wikipedia, 2017).

Para dar seguimiento al desarrollo de los métodos electorales partiremos de lo racionalidad completa a la cual dedicaremos parte de este trabajo.

1.1. Elecciones y Racionalidad Completa

La teoría clásica de la decisión racional asume que la racionalidad del ser humano es perfecta; ella supone que en una situación de decisión, el medio, la información, las creencias y análisis personales, son óptimos; las estimaciones de probabilidad son fácilmente realizables; el individuo tiene a su alcance información sobre todas las alternativas posibles y dispone de un sistema completo y consistente de preferencias que le permite hacer un perfecto análisis de todas ellas; no presenta dificultades ni límites en los cálculos matemáticos que debe realizar para determinar cual es la mejor, por tanto, garantiza que la alternativa elegida es un óptimo global (Simon, 1957).

Una vez establecida la racionalidad completa podemos dar paso a una breve revisión del desarrollo de los métodos electorales.

1.1.1. Métodos electorales

Como una parte de la teoría de la elección racional estan los métodos electorales y su historia se entrelaza con la historia de la democracia. La palabra *democracia* viene del griego *demos* ("pueblo") y *kratos* ("fuerza, poder"). Fue creada por los atenienses como una forma de mejorar su sistema de gobierno (aproximadamente 508 A.C.). En sus inicios, se pedía a los votantes que emitieran un voto por el político a quien deseaban desterrar por un periodo de diez años (Barcelo, 2007).

La teoría de los métodos electorales fue estudiada por la Academia de Ciencias de París en la época de la revolución francesa, en particular Jean-Charles Borda propuso el "recuento Borda" en 1770 como un método para elegir a los miembros de la Academia mientras que el marqués de Condorcet, por su parte, propuso un método en el que los candidatos se presentan de dos en dos (Barcelo, 2007).

Los sistemas parlamentarios modernos basan su modelo de votación en el modelo de la constitución de los Estados Unidos de 1789. Allí se experimentaron y refinaron los métodos más usados actualmente. En cuanto a modelos de elección, podemos ver la proliferación de diversos trabajos (Wittman, 1973) que reflejan una forma particular de ver el mundo pero que cuentan con un eje en común, buscan no solamente predecir la conducta de los votantes y a los partidos, sino también comprender mejor sus interrelaciones.

Los siguientes apartados describen ejemplos de modelos de elección cuyo eje central es la racionalidad completa.

El modelo de Downs

La formalización del modelo de elecciones de Downs se hace utilizando el modelo espacial de (Hotelling, 1929), quien sugirió la aplicación de su "ciudad lineal"[1]a la competencia política entre dos partidos políticos. La ley de Duverger[2]lleva a esperar que por razones estratégicas haya dos partidos políticos cuando el sistema electoral es de pluralidad simple de votos[3]Este es un modelo de política pre electoral que Downs (1957) incorporó en su marco de dos partidos políticos oportunistas a los cuáles solo les interesa ganar elecciones (aunque a veces es ambiguo y supone que los partidos quieren maximizar votos). Supone que todos los miembros de un partido actúan en pos del objetivo del equipo. Hay además información completa sobre la identidad del votante mediano[4], ambos partidos convergen a la plataforma ideal mediana que separa

[1]El modelo:"Ciudad lineal" es el intervalo [0,1] Los consumidores están distribuidos uniformemente a lo largo de este intervalo. Hay 2 empresas, localizadas a cada extremo que venden el mismo bien. La única diferencia entre las empresas es su localización. c = costo de 1 unidad del bien, t =costo de transporte por unidad de distancia al cuadrado. Este costo es soportado por los consumidores cuando eligen una empresa o la otra. Representa el valor del tiempo, gasolina, etc.

[2]La ley de Duverger es un principio que afirma que el sistema electoral mayoritario conduce a un sistema bipartidista. El descubrimiento de este principio se atribuye a Maurice Duverger, sociólogo francés que observó este efecto y dejó constancia del mismo en diversos textos publicados en los años cincuenta y sesenta del siglo XX. Posteriormente otros politólogos comenzaron a llamar "ley" a este efecto.

[3]Cuando está en juego un solo cargo por distrito, hay razones estratégicas para que compitan solo dos partidos políticos Duverger (1957) es la referencia clásica. Cox (1997), plantea que empíricamente esto se verifica más a nivel local que nacional, aunque hay países bipartidistas como Estados Unidos. Con un sistema electoral mayoritario en cada distrito, como el de pluralidad simple de sufragios, tiene sentido maximizar la probabilidad de ganar las elecciones. Pero si hay un sistema electoral proporcional donde están en juego varios cargos por distrito, puede tener sentido maximizar en cambio el voto esperado.

[4]El supuesto clave del modelo de "voto determinista" es que se sabe a ciencia cierta quién es el votante mediano. Para Downs, la identidad del votante mediano es igual a 50 si todos los votantes participan en la elección. El votante mediano es decisivo cuando el sistema electoral es de simple pluralidad de sufragios, donde gana el candidato con más votos, y compiten solo dos partidos políticos. En ese caso, las probabilidades de que un partido gane las elecciones son discontinuas y cambian abruptamente: cuando está más cerca del mediano que el otro partido, sus posibilidades características de la racionalidad completa de ganar son 1; cuando están a igual distancia, sus probabilidades son 1 ; cuando 2 está más alejado el mediano, sus oportunidades son cero. Dado esto, si hay dos partidos oportunistas a los que sólo les interesa ganar las elecciones, tienen un incentivo a moverse más cerca del mediano que el otro partido, ya que el que está más cerca del mediano tiene una mayoría de votos y gana.

al electorado en dos mitades, una con su punto ideal a la izquierda y otra a la derecha.

Aunque el teorema de Downs a menudo se confunde en la literatura con el teorema del votante mediano de Black, son dos resultados distintos. Black (1948a) consideró un sistema de votación entre un número finito de alternativas, el método de Condorcet de comparar cada par posible y seleccionar como ganadora la que tiene más victorias: con sólo dos alternativas, este método coincide con la simple pluralidad de sufragios que Downs consideró. La innovación fundamental de Black fue introducir las preferencias de un solo tope, que generalizan las preferencias espaciales de Hotelling (1929), mostrando que bajo esas condiciones la propuesta ideal del mediano puede derrotar cualquier otra opción, por lo que hay un único ganador de Condorcet. El problema es que las votaciones son manipulables si hay más de dos opciones.

Si bien Black se concentró en el problema de elegir entre políticas alternativas, el problema característico de las legislaturas o de cualquier cuerpo que tome decisiones directas en forma colectiva, menciona que su teorema se aplica también a la competencia entre candidatos, lo característico de la democracia representativa donde los votantes eligen delegados o representantes para que tomen las decisiones públicas por ellos. Conceptualmente, la diferencia clave entre Downs y Black es que en el primero las posiciones de los candidatos no están dadas: los candidatos eligen endógenamente su posición, por lo que la política pública propuesta se convierte en la variable estratégica clave.

El modelo de Wittman

Wittman (1990) se plantea un modelo más complicado, ahora los partidos deben balancear dos cuestiones, primero las oportunidades electorales y segundo la política que se implementará. Para ver esto supongamos que 'A(pA, pB) la probabilidad de que el partido A gane las elecciones cuando las plataformas son pA y pB, respectivamente. Consideremos además que las preferencias de los partidos vienen dadas por

W(p, j) con j= A, B. Entonces, el problema del partido j es elegir una plataforma que maximice: 'A(pA, pB) · W(pA, j) + [1 'A(pA, pB)] · W(pB, j) mientras que en el modelo de Downs los partidos simplemente maximizan 'A(pA, pB)

A pesar de que el modelo de Wittman es más complicado que el modelo de Downs. La intuición es que aún si los partidos tienen un costado programático[5] y quieren imponer un determinado programa de gobierno, antes tienen que ganar las elecciones para poder implementar su programa preferido[6]. Siempre y cuando estén polarizados a un lado y otro del mediano, los partidos están forzados a moverse hacia el mediano, para evitar que el otro partido se acerque aún más y gane las elecciones con una política opuesta todavía más alejada de su punto ideal que la del mediano. La lógica de la competencia lleva a los partidos a moderarse y converger hacia el mediano. De esta forma tenemos lo siguiente.

- Resultado (convergencia al mediano con partidos polarizados): Supongamos que hay dos partidos con principios, las plataformas electorales determinan las políticas pos electorales, la competencia electoral es sobre una única dimensión y las preferencias de los votantes son de un solo tope. Si hay información completa sobre la identidad del votante mediano, ambos partidos convergen a la plataforma ideal mediana sólo si están polarizados, en el sentido de que la política ideal de un partido está a la izquierda y la del otro a la derecha de la plataforma ideal mediana[7].

Por otro lado un resultado importante de la teoría de la elección racional es el teorema de Arrow, su enunciado y el análisis de sus implicaciones son analizados a continuación.

[5]Los modelos de votación con partidos interesados en las políticas implementadas, usualmente son denominados modelos con partidos “principistas”, también llamados “programáticos” o “ideológicos”, aunque dependiendo del contexto alguna de estas denominaciones podrían ser confusas. De hecho, los partidos podrían tener preferencias sobre las políticas simplemente porque son coaliciones formadas por grupos de ciudadanos con determinados intereses y no debido a que los políticos tengan una ideología propia. Cabe notar también que la distinción entre modelos con partidos oportunistas y modelos con partidos principistas es hasta cierto punto relativa, en un contexto de preferencias consecuencialistas, en tanto un político necesita antes ganar las elecciones para poder definir las políticas a aplicar.

[6]A estos partidos programáticos les interesa las políticas efectivamente implementadas en equilibrio. Si les interesara la fidelidad a los principios del partido como posición que sostienen en la campaña (es decir, les interesara no sólo el resultado final sino el proceso que siguen para competir), las consecuencias podrían ser diferentes ya que estos partidos serían más fundamentalmente principistas.

[7]Ver (Roemer, 2001), para un enunciado y prueba con todos los detalles técnicos.

1.1.2. Teoría de la elección social y el Teorema de Arrow

El teorema de imposibilidad de Kenneth J. Arrow aparece en su libro *Social Choice and Individual Values (Arrow, 1963).* Esta teoría estudia los métodos de agregación de intereses individuales que determinan las preferencias[8] sociales; el enunciado del teorema afirma que:

> "Cualquier regla de votación que respeta la *transitividad*, la *independencia de alternativas irrelevantes* y la *unanimidad* es una dictadura, siempre que la decisión se plantee, al menos, entre tres alternativas". (Arrow, 1963).

Para trasladar las preferencias individuales a las decisiones colectivas, Arrow afirma que se deben cumplir las siguientes reglas:

Transitividad.

Si A es preferido a B y B es preferido a C, entonces A tiene que ser preferido a C.

Unanimidad.

Si el colectivo prefiere A respecto a B, entonces, la mayoría de los individuos del colectivo prefieren A respecto a B.

Independencia de las alternativas irrelevantes.

Añadir o retirar alternativas, no cambia el orden de preferencia de las alternativas primeras.

[8] La relevancia de este resultado se deriva de la fundamentación matemática del los métodos de agregación de los intereses individuales que determinan las preferencias, cuya aplicación más importante se da en la Teoría Económica.

En la teoría económica neoclásica se entiende como preferencia del consumidor a la revelación de la elección hecha por un agente individual con base en una combinación de artículos y/o bienes de consumo, empleando racionalidad completa. Como parte de las preferencias del consumidor destacan los siguientes supuestos (Miller, 1985).

- *i*) El consumidor al elegir artículos de consumo puede decir si tiene preferencia por ellos o le son indiferentes
- *ii*) El consumidor es consecuente al elegir entre combinaciones de bienes (propiedad transitiva)
- *iii*) Para el consumidor más es preferible a menos (insaciabilidad)
- *iv*) El consumidor posee información completa y capacidad de cómputo lógico para realizar su elección.

Universalidad.

Ha de ser posible votar a cualquiera de las alternativas.

No imposición.

No ha de haber ningún resultado contradictorio con la naturaleza del sistema de votación.

Ausencia de dictadura.

Ningún votante individual puede tener en sus manos la decisión final.

Dictador.

Aquel individuo que determina las preferencias de una sociedad, es decir, hacia qué estado económico, social o político se dirige dicha sociedad[9].

Esencialmente el teorema establece que ningún sistema de elección social puede determinar las preferencias sociales sin violar al menos una condición de un conjunto específico de criterios *"razonables"* ya mencionados.

Estos criterios tienen sus raíces en la teoría de las elecciones democráticas, en particular en la creencia de que las elecciones sociales deberían, de alguna forma razonable, depender de las preferencias individuales en la sociedad y de ninguna otra cosa.

[9]Se puede consultar una demostración de este teorema en el Anexo A

El proceso de toma de decisión social tiene varias aplicaciones. El sistema de mercado en sí mismo es un ejemplo relevante de elección social. Algunas políticas tomadas por los gobiernos en donde los representantes toman decisiones mediante un sistema de votación sobre cuestiones que impactan a nivel nacional.

Si bien con este resultado Arrow aporta criterios para considerar las condiciones donde es posible dar la toma de decisiones en la elección social, este continúa siendo otro ejemplo de racionalidad completa; llama la atención que Arrow parte de la idea de que las personas efectúan análisis óptimos, que las estimaciones probabilísticas son fácilmente calculables y los individuos tienen a su alcance información completa de todas las alternativas posibles, además de que cuentan con un sistema completo y consistente de preferencias que les permite analizar a todas y cada una de ellas.

A continuación se menciona la aportación de Serge Galam quien fue el pionero del estudio de problemas sociales aplicando herramientas de la física estadística.

1.1.3. Sociofísica de Serge Galam

El Físico francés Serge Galam es el primero en estudiar problemas sociales con herramienta de la física en los años 70's, con ello se convierte en el padre de la Sociofísica.

La Sociofísica esta muy ligada a la física estadística, los problemas que abarca van desde : redes sociales, la evolución del lenguaje, propagación de epidemias, terrorismo, formación de coaliciones y la dinámica de opinión.

Galam (2004a) estudia el comportamiento de una elección con dos partidos y la influencia de un grupo minoritario de votantes inflexibles en el desenlace del proceso. El resultado más notable de Galam es que con dos partidos, y en condiciones de *racionalidad completa*, el número de votantes que favorece a cada candidato tiende al equilibrio (en el sentido de no haber un ganador contundente). Con esto se ha pretendido explicar lo apretado de las votaciones Bush-Gore en 2006 y otras más (Galam, 2007) .

Cabe destacar que en los modelos de Galam donde los votantes interactúan entre ellos, lo hacen socializando en contactos casuales (en el cine, el café, en un restaurante, etc.), si bien estos supuestos han sido criticados por no dejar claro el aspecto de la *interacción de los votantes*, su mérito es haber sido el primero en proponer un método para modelar estos complicados fenómenos sociales.

Hasta ahora la principal diferencia entre los modelos de Wittman, Downs y Galam es la forma de modelar con ecuaciones simples o con modelos adaptados de la física estadística, sin embargo el modelo de Galam continúa usando como idea subyacente que los votantes emplean *racionalidad completa,* esta similitud con los otros modelos nos da la pauta para clasificarlos como *modelos ortodoxos,* pues la herramienta matemática puede diferir pero los supuestos son iguales a los de la teoría de la elección racional.

En contra parte a la racionalidad completa, tenemos otro tipo de racionalidad donde los individuos no poseen información completa para decidir o analizar todas las posibilidades, no tienen capacidad de análisis para estas decisiones y no buscan soluciones óptimas, en la siguiente sección presentamos la racionalidad limitada.

1.2. Elecciones y Racionalidad Limitada

De acuerdo con el concepto de racionalidad limitada de Simon, los individuos tienen la capacidad de análisis limitado y como consecuencia a veces adoptan reglas muy simples de toma de decisión (heurísticas) para alcanzar los propósitos deseados. Estas reglas pueden ser muy exitosas y algunas veces funcionan mejor que los algoritmos de toma de decisión racional (instrumental). Cabe destacar que estas reglas pueden estar dadas por el entorno, por ejemplo costumbres, hábitos, etc.

1.2.1. Racionalidad Limitada según Herbert Simon

Simon (1957) definió la Racionalidad Limitada como el término que describe el proceso de decisión de un individuo considerando limitaciones cognoscitivas tanto de conocimiento como de capacidad computacional. En su obra *"Models of Man"* señala que la mayoría de las personas son sólo parcialmente racionales y que, de hecho, actúan según sus impulsos emocionales y no de forma racional en muchas de sus acciones. Insiste que la racionalidad personal está limitada por tres dimensiones:

1. La información disponible
2. La limitación cognoscitiva de la mente
3. El tiempo disponible para tomar la decisión.

Simon sugiere que los agentes económicos además de emplear la racionalidad limitada, también usan métodos heurísticos para tomar decisiones más que reglas rígidas de optimización. Para Simon esta manera de proceder se debe a lo intrincado de la situación o a la incapacidad de procesar y computar todas las alternativas, cuando los costos de deliberación son altos.

La idea de racionalidad limitada fue propuesta por Herbert Simon como una crítica a la teoría de la racionalidad completa inmersa en la teoría económica neoclásica. La propuesta de Simon implica una triple transformación del modelo de racionalidad completa; en primer lugar, el autor aboga por una concepción procedimental en lugar de la concepción sustantiva; en segundo lugar, reemplaza el concepto de maximización por el de satisfacción, en donde asegura que el decisor no se preocupa tanto por elegir lo óptimo como por elegir una acción cuyo resultado le satisfaga.

Crítica de Simon al concepto de Racionalidad Completa

En el proceso de toma de decisiones, incluso en problemas relativamente simples, no se puede obtener un máximo ya que es imposible verificar todas las posibles alternativas. Las personas difieren tanto en oportunidades disponibles como en deseos (influenciados por factores de su entorno). Cuando un individuo debe decidir, influyen en él, tanto los deseos que posee como las oportunidades que él cree poseer[10]. No es seguro que esas creencias sean correctas: es posible que el individuo no sea consciente de algunas oportunidades que en realidad le son viables o, puede creer que le son propicias ciertas oportunidades que en realidad no lo son, por lo tanto no puede garantizarse que elegirá la mejor alternativa (Elster, 1991).

Según esto, la racionalidad es limitada desde dos direcciones: desde el entorno del decisor, ya que no tiene acceso a la información perfecta, ni a la certidumbre e influyen en él factores exógenos como la cultura, las organizaciones en las que está inmerso etc., y desde el proceso mental del decisor pues este no tiene ni la estructura perfecta de preferencias, ni la capacidad completa de cálculo, y le afectan factores como la experiencia, la memoria, la percepción, las creencias y la sensibilidad personal. La teoría de la racionalidad limitada, no asume al decisor como un ser no racional, sino un ser que trata de ser racional con lo que tiene.

Se reconoce entonces la incapacidad de la teoría racional para captar completamente el proceso de decisión que llevan a cabo los individuos en la realidad. Ante la imposibilidad de optimizar, la teoría de la racionalidad limitada busca caminos satisfactorios para el decisor. Como sabe que la realidad que el decisor percibe es una realidad parcial y simplificada, no pretende tratar el mundo real en toda su extensión y busca soluciones que le sean satisfactorias ante “su realidad”.

[10]En palabras del propio Simon:

> “*...La labor consiste en reemplazar la racionalidad global del hombre económico por un tipo de conducta racional que sea compatible con el acceso a la información y con las capacidades computacionales que realmente poseen los organismos, incluido el hombre, en aquellos contextos en que existen tales organismos*” (Simon, 1957)

Según Simon (1992), *el individuo es fundamentalmente un ser adaptativo a su entorno.* El individuo sólo recoge parte de la información del entorno y tiene que desechar parte de la que le es dada por su complejidad; utiliza representaciones mentales, que tienen que ser soportados en una memoria de trabajo que tiene una capacidad no infinita. Es decir, la resolución está condicionada por la cantidad de elementos que tenga en la memoria de trabajo. Pero, las cadenas de causalidad de los hechos realmente determinantes, son breves y sencillas, por lo tanto, es posible tomar decisiones sin considerar toda la información del entorno y siguiendo una serie de reglas sencillas y manejables.

Factores que limitan la racionalidad

En la racionalidad limitada, el decisor no tiene acceso a la información perfecta, ni a la certidumbre; además en el influyen factores exógenos como la cultura, las organizaciones en las que esta inmerso etc. La cultura se interpreta como un sistema de valores y creencias que establecen una serie de normas sociales que pueden comprometer la decisión de un individuo. Las emociones tales como el amor filial o el disgusto pueden dar también efectivas reglas para modelar la búsqueda. Similarmente, en especies sociales, la imitación y el aprendizaje social pueden ser vistos como mecanismos que permiten un rápido aprendizaje y obvian la necesidad de cálculos individuales de utilidades esperadas.

Ante la imposibilidad de optimizar, la teoría de la racionalidad limitada busca caminos satisfactorios para el decisor. Como sabe que la realidad que el decisor percibe es una realidad parcial y simplificada, no pretende tratar el mundo real en toda su complejidad y busca soluciones que le sean satisfactorias ante “su realidad”.

Aprendizaje y el reconocimiento de patrones en agentes adaptativos

Simon realizó estudios acerca de los procesos de aprendizaje de los individuos desde el punto de vista de la sicología y extendió sus conceptos y postulados a la teoría económica. Investigó principalmente sobre los modelos de comportamiento adaptativo (teorías de aprendizaje). En su opinión, el razonamiento humano puede entenderse como una búsqueda selectiva a través de grandes espacios de posibilidades. Esa selectividad se hace aplicando reglas heurísticas para determinar los patrones que pueden seleccionarse y los que pueden ignorarse. La búsqueda termina cuando se ha encontrado una solución satisfactoria, casi siempre, antes de que todas las alternativas hayan sido examinadas.

Simon utilizó el siguiente ejemplo para desarrollar su teoría de comportamiento adaptativo:

Un organismo (no necesariamente humano), que busca comida para su supervivencia, en un medio sobre el que puede moverse libremente y que presenta puntos aislados donde se encuentra comida. El organismo puede ver en cualquier momento una porción circular de la superficie alrededor del punto donde él está. Puede moverse a una velocidad con un límite máximo. Él necesita comer cierta cantidad de comida en un periodo de tiempo. Si come la cantidad adecuada tendrá cierta energía para buscar la que sigue. El problema de la elección es elegir la ruta de tal forma que él no se muera de hambre. Simon demostró que el organismo requiere solamente de percepciones muy simples y mecanismos de elección para satisfacer sus necesidades y asegurar una alta probabilidad de sobrevivir sobre periodos extensos de tiempo. A medida que busca va recolectando patrones de búsqueda, y entre más patrones recoja, más fácil se le hace su búsqueda y entre más relevantes sean los patrones a su disposición, mejores serán sus decisiones. En particular, el organismo no necesita una función de utilidad, ni siquiera requiere de elaborados procedimientos para calcular las tasas marginales de sustitución entre diferentes metas. Simon demostró, por ejemplo, que la clave en el ajedrez es el reconocimiento de patrones: el buen jugador hace uso de una acumulación de patrones característicos. En otras áreas es similar. Calculó de forma experimental que un “experto” en cualquier área ha almacenado entre cien mil y dos millones de patrones de memoria. La experiencia, como captadora de patrones proporciona soluciones adecuadas en lugar de ideales.

Propuso entonces “modelos adaptativos de búsqueda” y al comparar su propuesta con los modelos de comportamiento racional empleados en economía, estos últimos muestran una mayor diversidad en los mecanismos de elección y una mayor capacidad en los organismos para obtener información y hacer cálculos. Sin embargo, sus modelos adaptativos están más de acuerdo con el comportamiento observado en laboratorio y en el campo que las teorías de comportamiento racional. Esa adaptabilidad está lejos del ideal del postulado de "optimizar" de la teoría económica clásica. Los organismos se adaptan bien para satisfacer más que para optimizar. La teoría del aprendizaje toma en cuenta las limitaciones en la capacidad y variedad de los organismos y el hecho de que, el medio en el cual el organismo debe adaptarse, posee propiedades que permiten simplificaciones de sus mecanismos de elección. Lo que Simon se preguntó es cómo hacer un postulado simplificado sobre un conjunto de mecanismos de elección y obtener aún las características relevantes del comportamiento de elección adaptativo observado. Su respuesta fue que es posible hacer modelos basados en discernir reglas rápidas y efectivas que pueden ser tan exactas como los modelos estadísticos complejos (por ejemplo, las redes bayesianas), los cuales necesitan más información y poder computacional. Las heurísticas simples pueden valerse de estructuras de información del medio ambiente. Su racionalidad es una forma de "racionalidad ecológica", más que de consistencia y coherencia. Un modelo de reglas heurísticas sencillas es más robusto que un modelo con un gran número de parámetros.

Las características de un modelo adaptativo según Simon:

1. Un modelo adaptativo consiste de reglas simples paso por paso que funcionan bien bajo restricciones de búsqueda, conocimiento y tiempo limitados.

2. Esas heurísticas son rápidas y efectivas y computacionalmente baratas más que consistentes, coherentes y generales.

3. Esas heurísticas son adaptables a medios particulares, pasados o presentes, físicos o sociales. Esta "racionalidad ecológica" permite la posibilidad que las heurísticas puedan ser rápidas, efectivas y correctas todo al mismo tiempo explotando la estructura de la información en los medios naturales.

4. El grupo de heurísticas es dirigido por algunos mecanismos que reflejan la importancia de las motivaciones y metas en conflicto.

Es decir, ese modelo del proceso del pensamiento humano puede ser efectivamente establecido en la forma de programas. Se sabe que los mecanismos fisiológicos almacenan esos programas en el cerebro y los ejecutan, pero aún no se conoce cuáles partes de esos programas mentales son inherentes y cuáles partes son adquiridas. Se conoce muy poco acerca del sustrato biológico para los programas y cómo esos programas pueden ser modificados y mejorados a través de educación y entrenamiento. Las computadoras, pueden entonces simular el pensamiento humano y los programas en lenguajes de procesamiento de información ofrecen un poderoso medio para expresar esas teorías (Simon, 1992).

En el siguiente apartado mencionaremos el trabajo de Brian Arthur, quien desde la economía y con la influencia de las ideas de la racionalidad limitada de Simon, se pregunta cómo evolucionan las decisiones de las personas.

1.2.2. Racionalidad desde la óptica de Brian Arthur

Para Arthur (1994) el tipo de racionalidad que se emplea en economía -la que supone el uso de racionalidad completa, lógica y que emplea la deducción, es de suma utilidad para resolver problemas teóricos, pero en los problemas de la vida cotidiana le exige mucho a la conducta de las personas (agentes) cuando tienen que decidir ante situaciones complicadas, de hecho más de lo que ellas pueden hacer.

Existen dos razones por las que la racionalidad completa falla. La primera es que más allá de cierto nivel de dificultad, la capacidad lógica humana deja de funcionar -la racionalidad humana es limitada. La otra es que en en situaciones de complicadas interacciones, los agentes no pueden confiar en que los otros agentes con los que esta lidiando se comporten con racionalidad completa, por lo que se ven obligados a adivinar su comportamiento.

Esto genera un mundo de creencias subjetivas, y de creencias subjetivas acerca de las creencias subjetivas. Los supuestos objetivos, bien definidos y compartidos dejan de aplicarse. A su vez, el razonamiento racional y deductivo (derivando una conclusión por procesos lógicos perfectos a partir de premisas bien definidas) no puede aplicarse.

Los economistas, por supuesto, están al tanto de esta situación. La pregunta no es si la racionalidad completa funciona, más bien es qué poner en su lugar. ¿Cómo funciona un modelo de racionalidad limitada? .

Aunque no hay un consenso general sobre esto, los sicólogos modernos concuerdan que en situaciones que son complicadas o mal definidas los humanos usamos métodos de razonamiento característicos y predecibles. Estos métodos no son deductivos, más bien son *inductivos.*

¿Cómo razonan los humanos en situaciones que son complicadas o mal definidas? La psicología moderna nos dice que, como humanos, solo somos moderadamente buenos para la lógica deductiva y además la usamos moderadamente. Sin embargo somos muy buenos buscando y reconociendo patrones -i.e. conductas que nos confieren beneficios evolutivos.

En la resolución de problemas, buscamos patrones; y simplificamos el problema al usar estos para construir modelos internos temporales o hipótesis o esquemas para trabajar con ellos. Llevamos a cabo deducciones localizadas basadas en nuestras hipótesis actuales y actuamos con base en ellas. A medida que llega la retroalimentación del entorno, podemos fortalecer o debilitar nuestras creencias en nuestras hipótesis actuales, descartando algunas cuando dejan de funcionar y reemplazándolas según sea necesario por otras nuevas.

Arthur (1994) propone una forma de modelar el pensamiento inductivo: En un problema típico que se realiza a corto plazo, nosotros quizás establezcamos una colección de agentes, probablemente heterogéneos, y asumamos que pueden formar modelos mentales, o las hipótesis, o las creencias subjetivas. Estas creencias quizás entren en la forma de expresiones matemáticas sencillas que pueden ser utilizadas para describir o predecir alguna variable o acción; o de modelos complicados de expectativas del tipo utilizados en economía; o de hipótesis estadísticas; o de reglas de condición/predicción ("Si la situación Q se observa/predice el resultado o acción D"). Esto será normalmente subjetivo, esto es, variarán entre los agentes. Un agente puede tener una en mente a la vez, o varias simultáneamente.

Cada agente seguirá el comportamiento de una colección privada de tales modelos de creencia. Cuando viene la hora de tomar decisiones, él actúa sobre lo actualmente más creíble (o posiblemente lo más provechoso). En tanto que los otros, él los mantiene en la parte posterior de su mente. Alternativamente, él puede actuar sobre una combinación de varias creencias (sin embargo, los humanos tienden a mantener en la mente muchas hipótesis y actuar sobre la más plausible). Una vez que las acciones son tomadas el cuadro agregativo es actualizado, y los agentes ponen al día el historial de todas sus hipótesis.

Los agentes "aprenden" cuál de sus hipótesis trabajar, y de vez en cuando pueden desechar hipótesis pobremente comportadas y generar nuevas "ideas" para colocarlas en su lugar. Los agentes se demoran en la actualización de sus hipótesis más creíbles. Esto causa una histéresis incorporada. Se aferra a un modelo de creencia no porque sea "correcto" –no hay manera de saber esto – sino porque se ha trabajado con él en el pasado, y ha acumulado un expediente de fallas antes de que sea desechando. En general, es posible que haya una constante, el movimiento lento de hipótesis afectara. Podríamos hablar de esto como un sistema de expectativas temporalmente satisfechas – las creencias o los modelos o las hipótesis que son cumplidas temporalmente (aunque no perfectamente), que lleva a diversas creencias o hipótesis cuando dejan de ser satisfechas.

“Una pregunta clave que nos queda. ¿De dónde vienen las hipótesis o los modelos mentales? ¿Cómo son generados? Conductivamente, esta es una pregunta profunda en la psicología, tiene que ver con cognición, la representación objetiva y el reconocimiento de formas. . Hay, no obstante, algunas opciones sencillas y prácticas para la modelación. A veces dotaremos a nuestros agentes de modelos focales – pautas o hipótesis que son obvias, simples y fácilmente tratables mentalmente. Nosotros quizás generemos un "banco" de estos modelos y los distribuyamos entre los agentes. Otras veces, dado un modelo-espacio conveniente, quizás admitamos el algoritmo genético o buscar algún dispositivo inteligente similar para generar modelos “más inteligentes". Podríamos también permitir la posibilidad de que los agentes tomen modelos mentales entre ellos (proceso conocido como transferencia en sicología). Cualquiera que sea la opción tomada, es importante aclarar que el marco descrito arriba es independiente de las hipótesis o creencias específicas utilizadas .

Modelo “Asistencia al Bar El Farol”

El problema del bar “El Farol” es importante para ilustrar el razonamiento inductivo, sus implicaciones y cómo se puede modelar.

“Un centenar de personas tiene que decidir de forma independiente cada semana la posibilidad de presentarse en su bar favorito (El Farol, en Santa Fe). La regla es que si una persona predice que más de 60 (por ejemplo) van a asistir, evitará las multitudes y se quedará en casa, y si predice menos de 60 entonces asistirá. Es de interés como los asistentes al bar podrían predecir cada semana el número de los que asistirán y la dinámica resultante del número de participantes”(Arthur, 1994).

Dos características destacan en este problema.

- Primero, nuestros agentes se darán cuenta rápidamente que los pronósticos de cuántos asistirán depende de las predicciones de los demás de cuántos asistirán (porque eso determina su asistencia). Pero las predicciones de los demás, a su vez dependen de sus predicciones y de las predicciones de los demás. Deductivamente hay una regresión infinita. Ningún modelo de expectativas "correcto" puede ser asumido como un conocimiento común, y desde el punto de vista de los agentes, el problema está mal definido. (Esto es cierto para la mayoría de los problemas de expectativas, no sólo para este ejemplo.).

- En segundo lugar, si todos utilizan un modelo de expectativas que predice que pocos asistirán, todos asistirán, invalidando ese modelo. Del mismo modo, si todos creen que la mayoría asistirá, nadie va a ir, lo que invalida la creencia. Las expectativas se verán obligadas a diferir.

Lo que Arthur (1994) se pregunta ¿Cómo sería el comportamiento dinámico de los asistentes a lo largo del tiempo? ¿Convergerá, y si ése es el caso a qué? ¿Llegará a ser caótico?

La explicación que da es : el problema del Bar es una economía de expectativas en miniatura, con una dinámica compleja, en donde los agentes al tomar sus decisiones formulan sus hipótesis ante cambios en el entorno y evidentemente surgen patrones emergentes.

Una vez que hemos revisado los aspectos teóricos, debemos describir el concepto de elección bajo la noción de racionalidad limitada, la siguiente sección se encarga de este concepto.

1.2.3. Elecciones con Racionalidad Limitada

A continuación se describen las características de un procesos de elección empleando los conceptos de racionalidad limitada explicados anteriormente.

Preferencia

- Se entiende como preferencia del elector a la revelación de la elección hecha por un agente individual con base en un conjunto de partidos empleando racionalidad limitada.
- La forma tradicional de recavar información de las preferencias reveladas es mediante encuestas, lo cual predispone a las personas a revelar auténticamente sus preferencias, por lo anterior las información así obtenida no es confiable.

Modelación

- El elector posee información incompleta y su capacidad de análisis es limitada.
- El elector toma decisiones heurísticas influenciado por el ambiente.
- El elector no es consecuente al elegir entre combinaciones de candidatos, es decir, no hay transitividad en sus preferencias.

Un ejemplo práctico Se trata del caso de una persona que vive en una zona de bajos recursos y que al verse en la disyuntiva de sobrevivir sin trabajo, aceptó hacer un trueque con una agrupación política que se dedica a comprar votos. Intercambia su voto por comida, láminas para su casa y de cuando en cuando algo de dinero; sus comentarios al respecto son contundentes, "no es mucho, pero es mejor que lo que el gobierno en turno nos da".

De esta manera su capacidad de tomar decisiones electorales está condicionada irremediablemente por el medio en el que le tocó vivir; lo cual no representa un conflicto de ningún tipo, para ella la decisión es sencilla, a sus 64 años, sin estudios y sin empleo, no tiene otra forma de sobrevivir.

Es así como nos encontramos en un escenario donde la señora no es congruente al elegir de entre los posibles candidatos en una elección, siempre vota por quien le indican; no posee información completa de las opciones a elegir, tiene capacidad de análisis limitado y además toma decisiones pragmáticas. Entonces surgen la preguntas: *¿es irracional su conducta?, ¿bajo que criterios podría calificarse de irracional esta forma de votar?, ¿es un caso aislado o es uno observable en la realidad?.*

2. Sistemas Complejos

En este capítulo desarrollaremos los fundamentos de la teoría de los sistemas complejos que emplearemos como marco teórico principal para nuestra propuesta de modelación.

2.1. Antecedentes

En relación con los sistemas complejos, Miramontes (1999b) afirma que los sistemas son las cosas más los procesos, igualmente explica que un sistema consta de una base material y de un conjunto de relaciones entre los objetos que lo constituyen. A su vez los procesos definen y son definidos por una dinámica, la cual puede o no ser lineal.

Etimológicamente Complejidad viene del Latín *plexus*, que significa *entretejido*, lo que nos remite a algo difícil de separar. La razón de esta dificultad estriba en la codependencia entre los componentes del sistema complejo. Gershenson (2013) menciona que la evolución del sistema depende de las interacciones entre las componentes del mismo. Por esta razón no podemos estudiar de forma aislada las partes para comprender las propiedades del sistema complejo.

Para Gershenson (2013) es importante distinguir lo complejo (que tiene que ver con lo compuesto) de lo complicado (que se relaciona con lo intrincado). Lo complejo no tiene que ser algo difícil de entender.

> "Hay sistemas complejos y complicados (redes genéticas, ecosistemas: hay muchas interacciones relevantes, las cuales son difíciles de desentramar y comprender), complejos y sencillos (modelos computacionales, tales como autómatas celulares (Wolfram, 2002): hay reglas muy sencillas de interacción que pueden producir dinámica muy compleja), simples y complicados (un fotón: estudiado de manera aislada (sin interacciones), su descripción

> es difícil de manejar) y simples y sencillos (una piedra: no tiene interacciones y al no tener dinámica interna, es fácil de describir)" (Gershenson, 2013).

En un proceso lineal los efectos siempre son proporcionales a las causas y en todo proceso lineal es válido el principio de superposición. Por otro lado, los sistemas que no cumplen con las premisas antes mencionadas se conocen como no-lineales.

Los sistemas lineales han sido ampliamente estudiados por su importancia metodológica y formativa en la modelación de problemas, pero sobre todo porque matemáticamente la descripción y manipulación de estos modelos se relaciona con teorías científicas acabadas.

2.2. Características

Un sistema es complejo si cumple lo siguiente (Miramontes, 1999b):

- 1) Esta integrado por componentes simples que interactúan entre sí. Se dice que los componentes de un sistema complejos son simples si su estado se puede describir con pocas variables
- 2)Su estado cambia al transcurrir el tiempo y el cambio es el resultado de una dinámica no-lineal que usualmente tiene dos partes: una local, que modifica el estado de los elementos como resultado de su interacción con elementos vecinos y una dinámica global, que obedece a las interacciones que pesan sobre el sistema y que provienen de las interacciones de este con el resto del universo.

Algunas de las manifestaciones que encontramos en sistemas complejos son:
Ruptura de simetría. Que consiste en la aparición de patrones espacio-temporales en donde antes sólo había homogeneidad. Esta propiedad es exclusiva de sistemas abiertos que intercambian masa, energía e información con su entorno.

Criticalidad auto-organizada (CAO). Estos sistemas evolucionan de una manera natural y espontánea hacia un estado crítico en donde una perturbación pequeña puede causar efectos de cualquier tamaño. De acuerdo con la teoría CAO los mecanismos que producen hechos pequeños son los mismos que producen cambios catastróficos , y los sistemas nunca alcanzan un verdadero estado de equilibrio sino que transitan de un estado meta-estable al siguiente.

Hay dos propiedades fundamentales de los sistemas con CAO: i) la incertidumbre provocada por dos condiciones iniciales distintas crece de manera polinomial; es decir, pierden memoria de su pasado, o capacidad de predicción futura; en contraste con el caos clásico (que lo hace de manera exponencial) este lo hace de forma lenta, llamándose *caos débil*, y se dice que el sistema *evoluciona en el borde del caos.*

Fractalidad. Cuando presentan estructuras discernibles en cualquier escala espacial y sus fluctuaciones; como consecuencia de la CAO, siguen reglas de distribución con auto-escalamiento. (Todo esto se denomina geometría fractal).

Generan propiedades emergentes. Ésta es quizá, una de las características más notables y son el resultado de los procesos en paralelo que se llevan a cabo en un sistema complejo y su naturaleza es intrínsecamente colectiva; surgen en cada nivel sucesivo de complejidad y no se pueden deducir a partir de los componentes del sistema

3. Propuesta de un modelo no ortodoxo

En los siguiente apartados desarrollaremos a detalle la propuesta de modelación de esta tesis siguiendo el enfoque de los sistemas complejos. En primera instancia hablaremos de nuestra herramienta de modelación: los modelos basados en agentes.

3.1. Modelos Basados en Agentes (MBA)

Para Gilbert (2007) los MBA constituyen *un nuevo método analítico para las ciencias sociales*; incluso Axelrod (1997) afirma que constituyen un tercer modo de hacer ciencia, distinto y complementario a los modos científicos tradicionales: la inducción y la deducción.

Un MBA es un tipo particular de modelo científico que se implementa como programa computacional. Gilbert (1996) afirma que los MBA sirven para explicar los fenómenos sociales.

El fundador de los modelos basados en agentes es John Von Neuman, quien además es considerado el pionero de la computación. A finales de la década de los 40´s, Von Neuman estaba interesado en los sistemas que se auto-reproducen. Inspirado por las ideas de su colega Stanislaw Ulam, fue como desarrolla el primer autómata celular.

La idea de Von Neuman fue crear un sistema hecho por un arreglo en forma de celdas constituido por células (Von Neumann, 1966). El tiempo en este sistema era también una variable discreta, y en cada paso cada célula actualiza su estado de acuerdo a un conjunto de reglas basadas en su estado previo y a sus vecinos . Cada celda es una máquina de estado finito simple, pero el comportamiento general del sistema puede volverse bastante complicado.

Von Neuman utilizó este enfoque para diseñar su “constructor universal”: un patrón de celdas que pueden reproducirse así mismas a lo largo del tiempo, proporcionando de este modo un ejemplo notable de cómo una propiedad importante a nivel de sistema (la auto-reproducción) se puede lograr a través de la interacción de las partes individuales que se comportan de forma independiente de la totalidad.

Sin embargo, lo que realmente trajo a los autómatas celulares a un primer plano, fue el Juego de la vida de John Conway (Gardner, 1970). Aunque las células de Von Neuman podían estar en 29 estados, se necesitaban decenas de reglas para describir transiciones entre ellas; en cambio en las células de Conway sólo están “vivas” o “muertas” y sólo se requieren 3 reglas para describir su conducta:

1. Celdas Sobrevivientes. Toda celda con 2 o 3 vecinos sobrevive en la siguiente generación.
2. Celdas Muertas. Toda celda con 4 o más vecinos muere por sobre-población. Cada celda con un vecino o solitaria muere por aislamiento.
3. Nacimientos. Cada celda vacía adyacente a exactamente 3 vecinos da lugar a una célula nueva por nacimiento en la siguiente generación.

Estas simples reglas aplicadas a diferentes patrones de células iniciales dan lugar a una extensa variedad de objetos con conductas complejas; entre ellas están los gliders , los blinkers etc. Los cuales son ejemplos de que un simple arreglo de bloques puede dar lugar a patrones difíciles de predecir.

Un problema fundamental en la teoría de autómatas celulares es la clasificación. Una buena clasificación divide un AC en grupos con propiedades relacionadas.

Los autómatas celulares hasta ahora mencionados sólo modelan fenómenos muy generales: “vida” y “autoreproducción”; sin embargo también pueden aplicarse a fenómenos sociales. Muchos problemas en ciencias sociales pueden modelarse con un conjunto de agentes que interactúan entre ellos en un espacio localizado.

El autómata celular elemental es una colección unidimensional de celdas con dos posibles estados $\{0,1\}$, evoluciona a través de pasos de tiempo discretos de acuerdo a un conjunto de reglas basadas en los estados de sus celdas vecinas.

Una aplicación son los modelos de segregación de Schelling (1969). En estos modelos Schelling explora los mecanismos que originan la formación de arreglos de agentes homogéneos (guetos[1]) en el espacio geográfico. El espacio se modela como un arreglo discreto, igual que en los autómatas de Von Neuman y Conway, sólo que en este caso cada célula individual representa un agente humano. Estos agentes pueden ser estrellas o cualquier otra figura (representando a diferentes etnias), las cuales tienes preferencias respecto a los miembros del grupo que los rodea. Si no están satisfechos, se mueven a un lugar cercano que satisfaga sus necesidades. Schelling exploró la dinámica del modelo para varios patrones iniciales y diversas preferencias en la distribución, sin embargo su conclusión es que aún para agentes que poseen alta tolerancia a convivir con vecinos diferentes a ellos; de manera consistente el sistema genera segregación.

A continuación se indican las propiedades de los modelos basados en agentes.

Propiedades generales de los MBA.

Según Borner *et al.* (2009) los Modelos Basados en Agentes poseen las siguientes propiedades:

1. **Heterogeneidad**. Los agentes pueden diferir de tantas maneras como el rango de parámetros individuales se los permita.
2. **Autonomía**. Los agentes no dependen de un control central.
3. **Espacio Explícito**. Se refiere a que los agentes están situados en un ambiente.
4. **Interacciones Locales**. Los agentes interactúan con otros agentes.
5. **Racionalidad Limitada**. Los agentes no poseen información global y no poseen poder de cómputo infinito; por lo regular hacen uso de reglas simples basadas en la información local.

En el siguiente apartado desarrollamos nuestra propuesta de modelo basado en agentes para estudiar las preferencias electorales.

[1] *gueto*. Adaptación gráfica de la voz italiana ghetto, ‘barrio en que se confinaba a los habitantes judíos de una ciudad’ y ‘barrio o zona en que vive aislada una minoría, normalmente marginada’: «En torno a las grandes ciudades se han creado verdaderos guetos de segregación etnocultural» (Geo [Esp.] 7.95). Deben evitarse las grafías híbridas guetto y gheto, que no son ni italianas ni españolas. (http: //www.rae.es/rae.html)

4. Un modelo de elecciones políticas para tres partidos empleando MBA

En esta sección se presenta una propuesta de estudio del fenómeno de la elección electoral empleando el enfoque de "abajo hacia arriba" (bottom $\uparrow$ up), para ello construiremos un modelo computacional que imita la actividad en colectivo de un número importante de individuos asignando reglas simples de comportamiento individual.

4.1. Supuestos del modelo

4.1.1. Heterogeneidad de los agentes

En el modelo propuesto nuestros votantes (agentes) poseen un índice de susceptibilidad que sirve para definir que tanto un individuo puede de cambiar su preferencia electoral, se trata de un parámetro fijo durante toda la simulación, su valor oscila entre 0 y 1 se asigna aleatoriamente antes del proceso de simulación.

4.1.2. Interacciones locales

La intención de incorporar este parámetro es obtener dos tipos de votantes con los cuales ocuparemos el "espacio geográfico" de la simulación, a saber son:

- votantes necios.- que son todos los agentes cuyo índice de susceptibilidad es cercano a cero. Con esto podemos considerarlos "inmunes" a los cambios de opinión de sus vecinos.
- votantes susceptibles.- que son todos los agentes cuya susceptibilidad se encuentra alrededor de 0.5 y 1. Estos votantes pueden cambiar o no de preferencia electoral.

Estos tienen asociado el grado de susceptibilidad a la influencia de la opinión de sus vecinos.

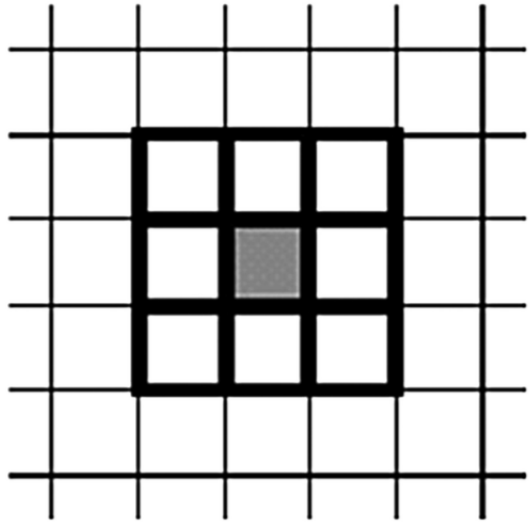

Figura 4.1.: Vecindad de Moore

Fuente: elaboración propia

En esta vecindad, se toman los grupos de nueve celdas en un arreglo cuadrado de 3×3; la celda analizada esta al centro y se compara con sus ocho vecinos adyacentes.

Es importante la influencia del medio ambiente sobre los cambios de preferencia electoral, de forma tal que analizamos la influencia de los vecinos de cada individuo (bajo un arreglo de vecindad de Moore), donde según la opinión mayoritaria y la susceptibilidad los individuos podrían estar cambiando su preferencia a lo largo del tiempo.

4.1.3. Autonomía

Los agentes no poseen un control central, los parámetros de susceptibilidad y preferencia electoral se asignan solamente una vez al comienzo de la simulación y de manera aleatoria.

4.1.4. Espacio explícito

El modelo contará con un espacio en forma de arreglo matricial donde se ubicarán 2500 celdas que almacenarán los parámetros de cada agente.

Para analizar los cambios en las preferencias electorales se tomó como medida de tiempo un año, lo cual se reflejará en el programa en 365 iteraciones.

4.1.5. Racionalidad limitada.

Los agentes no poseen información de todo el sistema, sólo cambian su preferencia electoral si son susceptibles a cambiar de opinión y si se encuentran rodeados por un grupo con opinión mayoritaria.

4.1.6. Programación del modelo multiagentes

4.1.6.1. Características de los agentes

El modelo consiste en un arreglo matricial de celdas de forma cuadrada, su tamaño total será 500×500 para formar un total de 250000 celdas, cada una de las cuales representa la posición de un agente.

Cada uno de estos posee dos cualidades:

- la preferencia electoral. se refiere a la decisión de voto por alguna de las opciones (partidos) que compiten en una campaña electoral, como nos interesa analizar el escenario con tres partidos, los posibles valores de este parámetro son 1, 2 o 3
- la susceptibilidad. Este parámetro toma los valores $0 < s < 1$, los valores cercanos a uno representan agentes muy proclives a cambiar de opinión y los valores cercanos a cero representan agentes con fuerte rigidez en su opinión. En ambos casos, asignamos aleatoriamente los valores a cada agente antes de poner en marcha el modelo multiagentes.

4.1.6.2. Influencia del entorno

Para establecer las reglas del cambio en la preferencia electoral consideraremos una vecindad de Moore, es decir, tomamos los subconjuntos de 3x3 casillas, y estableceremos como criterio de mayoría local lo siguiente:

- La preferencia de la casilla central dependerá de su preferencia inicial, la preferencia mayoritaria de las 8 casillas adyacentes a ella y de que el valor de susceptibilidad permita el cambio de preferencia.
- Si la casilla central es un votante "necio" el programa analiza un nueva casilla sin alterar la preferencia de voto.
- Si la casilla central es "susceptible" el programa analiza las preferencias de sus 8 vecinos, determina la preferencia mayoritaria y mediante un proceso aleatorio (por ejemplo, lanzando una moneda) se determina si a la casilla central se le asigna o no la preferencia de la mayoría.
- En el caso de que la opinión en la vecindad de Moore este empatada, se asigna un partido ganador mediante otro proceso aleatorio entre los partidos involucrados.
- El proceso de revisión de las preferencias se realiza con todos los agentes.

4.1.6.3. Tiempo de interacción

Posteriormente corremos el autómata celular considerando las interacciones de los agentes durante 365 pasos, lo cual representa 1 año de interacciones. Los datos de las preferencias electorales de las 250000 celdas durante este periodo de tiempo se almacenan para ser analizados después.

4.1.6.4. Índice de información mutua

El programa MUTINF se encarga del cálculo de la información mutua para el total de los agentes en función de la preferencia por el partido 1 2 y 3; así se generan 365 series de datos en cada opción de voto, que corresponden a los valores del índice de información mutua durante el periodo de tiempo que interactuaron los agentes. Se genera una gráfica de información mutua para cada partido (componente) con el propósito de analizar estos valores respecto a la opinión por partido en términos de los cambios en el valor de la información mutua.

Se procede entonces a analizar los datos buscando correlaciones no-lineales como lo plantea Cellucci (2005). En nuestro análisis, tomaremos los datos de 365 iteraciones para analizar cualitativamente en un gráfico tridimensional los cambios en la información mutua.

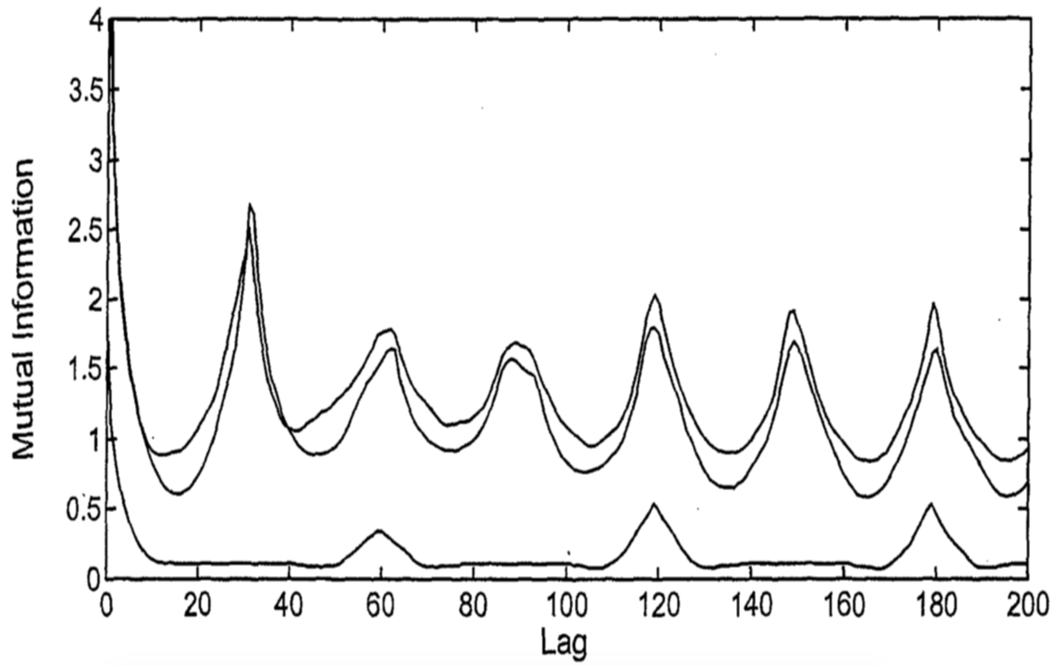

Figura 4.2.: Información mutua vs Lag

Fuente: Cellucci (2005)

El trabajo de Cellucci (2005), muestra cualitativamente las correlaciones no-lineales entre las variables de un sistema; para ello después de usar el algoritmo del cálculo de la información mutua en las variables del sistema, se grafican estos datos y se busca el primer mínimo de información mutua en cada componente.

Siguiendo esta descripción, en la figura se observan las gráficas de la variable X (por encima), de la variable Y (centro) y de la variable Z (abajo). Cellucci establece que el valor de la variable Lag para el cual se encuentra el mínimo global de información mutua equivale a 20, esto es considera el mayor valor del eje Lag (el mayor de los valores mínimos) donde las tres gráficas independientemente toman el primer valor mínimo de información mutua.

4.2. Resultados

En la primera iteración del programa se han asignado de manera aleatoria a cada celda su preferencia y un valor de susceptibilidad.

El arreglo de agentes en condiciones iniciales se asemeja a la distribución de los votantes en el modelo de Galam (2008) con la salvedad que en aquel se están considerando spins magnéticos con dos estados; en nuestro caso tenemos tres preferencias electorales representadas por tres colores.

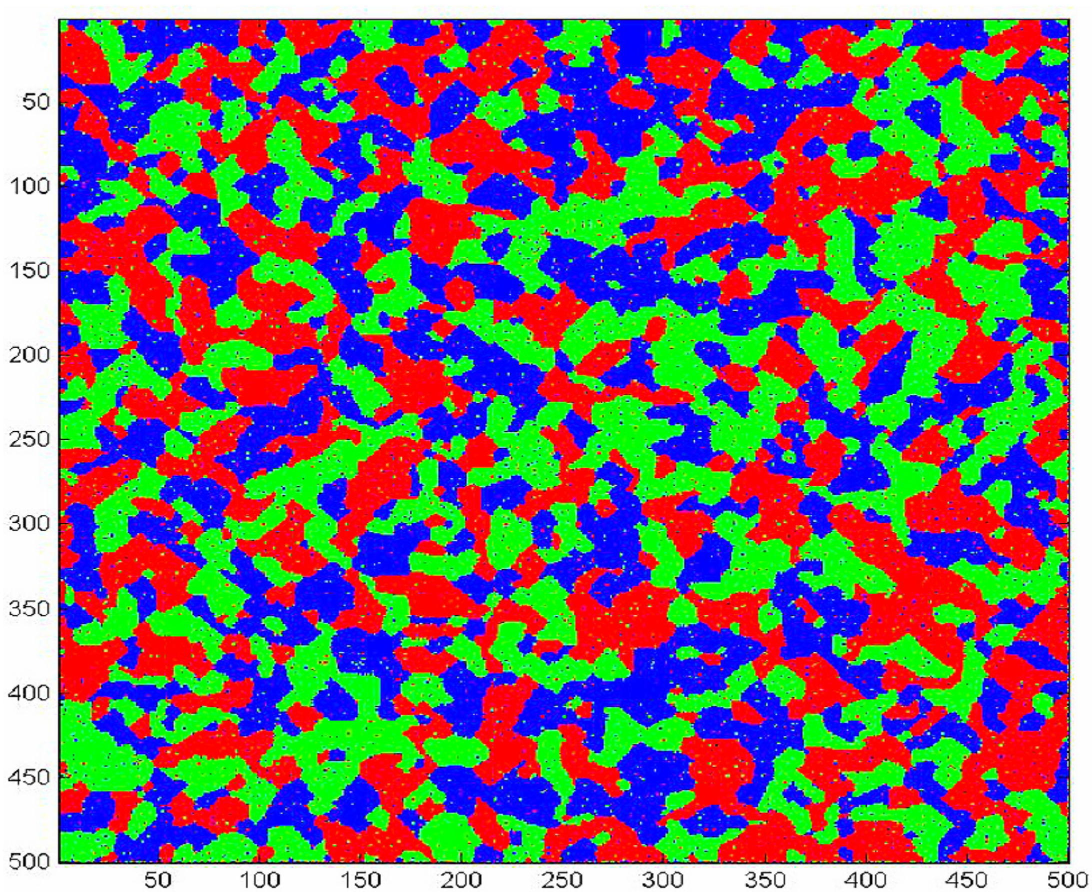

Figura 4.3.: Arreglo de agentes al final de la simulación.

Fuente: elaboración propia.

La figura muestra el arreglo de 250000 agentes a los que se les asignó un color para representar una preferencia, el programa identifica los valores verde= partido 1, azul= partido 2, rojo= partido 3. La asignación de la preferencia se realizó al azar al inicio del programa, se muestra en la figura el estado de las preferencias después de 365 iteraciones del programa.

En esta etapa primera, no hay patrones reconocibles en la distribución de las preferencias electorales. La figura 4.3 muestra el estado de las preferencias electorales al final de la simulación.

Aquí se aprecian regiones donde predomina alguna de las tres preferencias, sin embargo no hay un color que tenga presencia mayoritaria en todo el arreglo.

Se puede notar que dentro de amplias regiones de un solo color persisten agentes que conservan un color diferente. Esto denota la presencia de agentes que no cambiaron su preferencia aunque había una opinión mayoritaria de diferente partido alrededor de ellos.

Una dificultad que encontramos con el modelo computacional es que el consumo de memoria en las iteraciones del autómata nos impidió visualizar en pantalla los cambios de las preferencias del arreglo de agentes en todas las etapas del proceso, pues para poder visualizar una iteración en una computadora de escritorio se requiere del orden de horas de cálculos; por esta razón se almacenaron los datos numéricos de las preferencias para hacer un análisis numérico posterior y sólo se presentó el estado final de las preferencias en el sistema.

Observando el plano ZY en la gráfica, el cambio abrupto del valor de Información Mutua al rededor del valor de susceptibilidad 0.3 (*i.e.* 30×10^{-2}) es relevante pues cualitativamente corresponde con la mayor fluctuación de información mutua presente en la gráfica.

Globalmente destacan los cambios del comportamiento de la información mutua a lo largo de las 365 iteraciones, tenemos valores con mínimas variaciones cuando observamos susceptibilidad menor a 0.3, pero aparecen fluctuaciones importantes después del valor 0.3.

De el total de agentes se ha contabilizado la información mutua y se ha graficado en el eje Z, destaca en la gráfica el cambio abrupto del valor de Información Mutua alrededor del valor 0.3 en susceptibilidad .

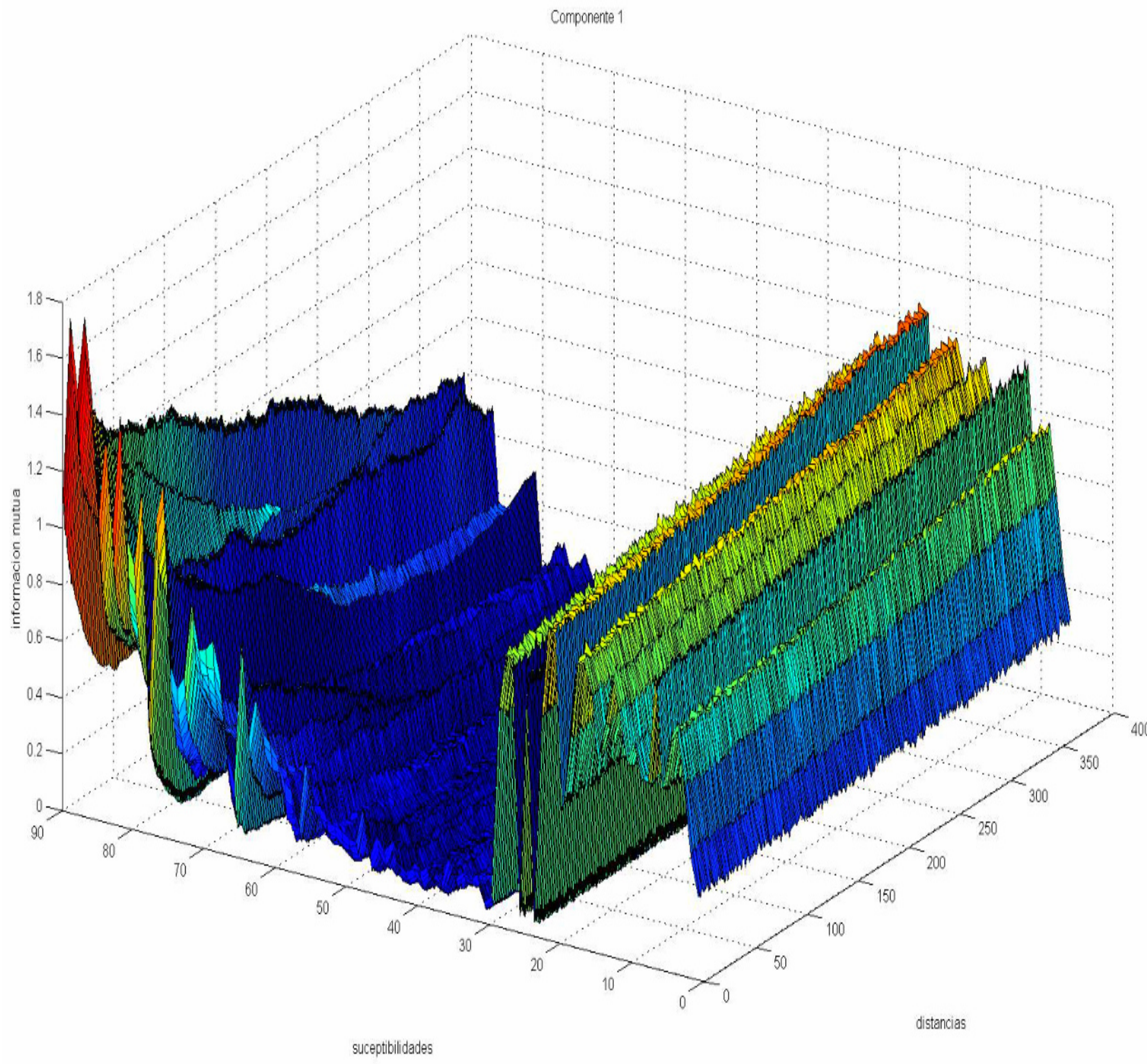

Figura 4.4.: Detalle de las serie de datos generada por el modelo en la primera componente.

Fuente: elaboración propia

Se indica la información mutua (IM) en el eje Z (en bits), en el eje Y tenemos la susceptibilidad (s) en escala de 10^{-2} unidades para apreciar mejor los detalles y en el eje X tenemos las distancias (iteraciones) en bloques de 50 unidades.

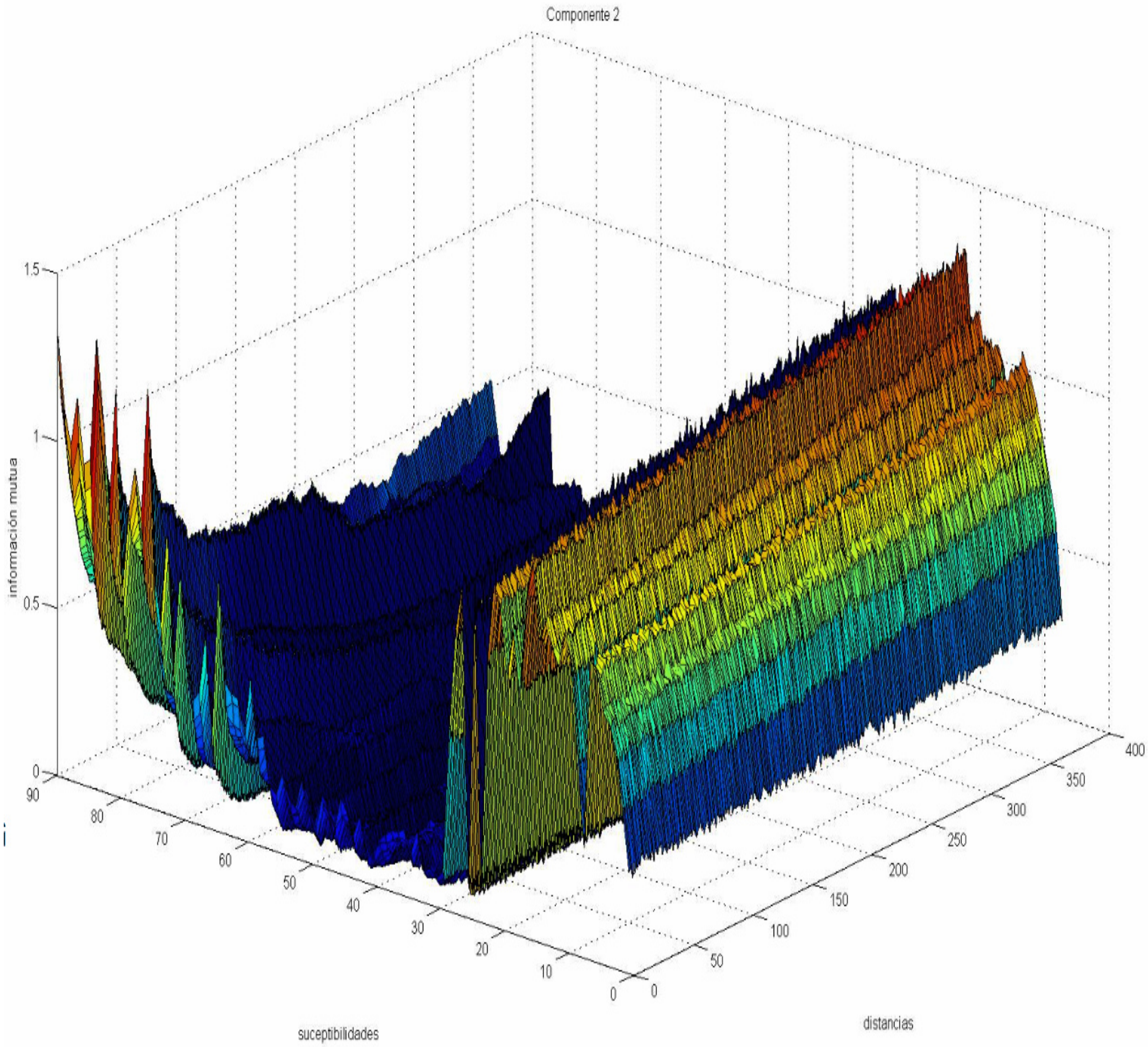

Figura 4.5.: Detalle de las serie de datos generada por el modelo en la segunda componente.

Fuente: elaboración propia.

Se indica la información mutua (IM) en el eje Z (en bits), en el eje Y tenemos la susceptibilidad (s) en escala de 10^{-2} unidades para apreciar mejor los detalles y en el eje X tenemos las distancias (iteraciones) en bloques de 50 unidades.

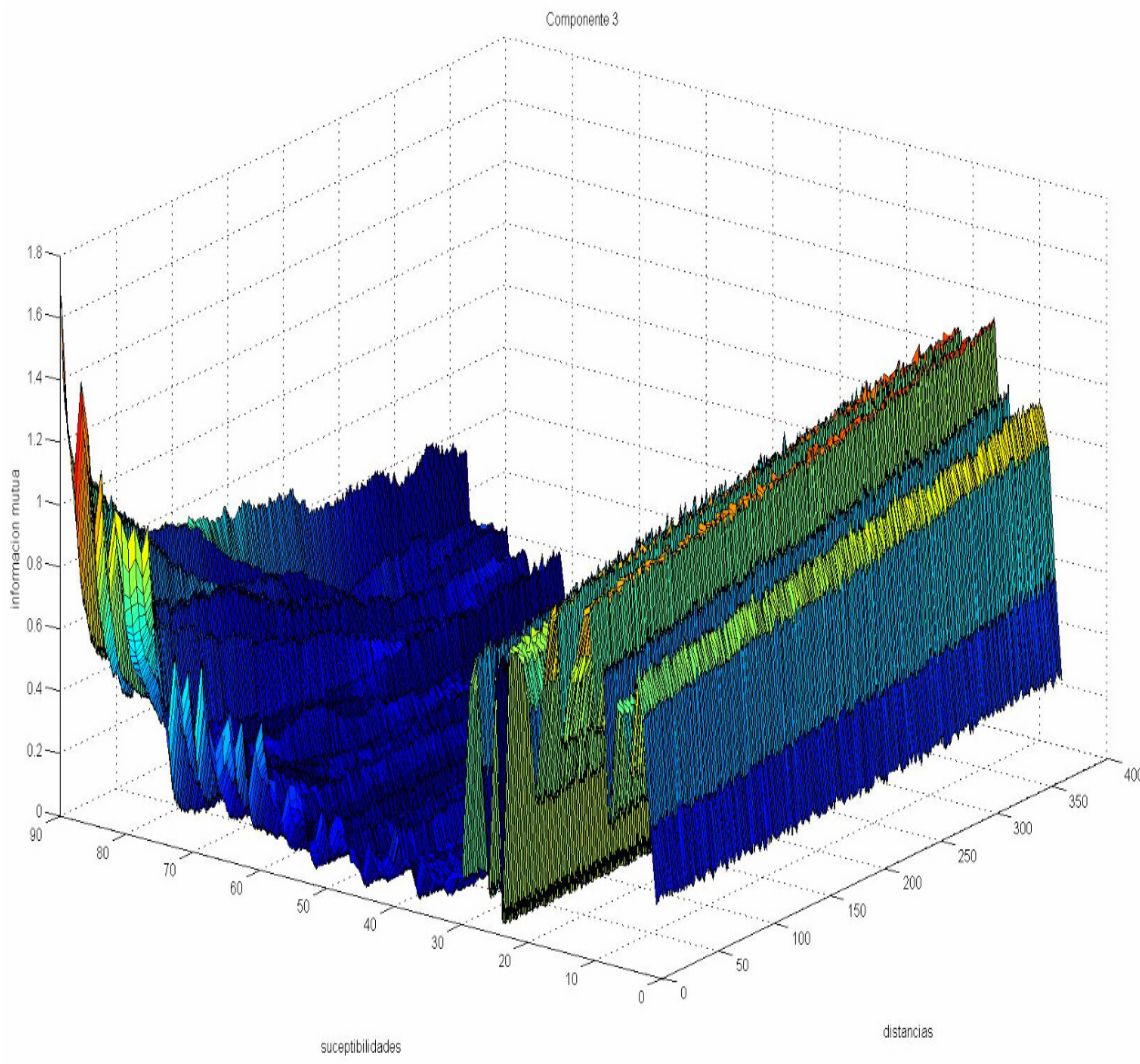

Figura 4.6.: Detalle de la serie de datos generada por el modelo en la tercera componente.

Fuente: elaboración propia

Se indica la información mutua (IM) en el eje Z (en bits), en el eje Y tenemos la susceptibilidad (s) en escala de 10^{-2} unidades para apreciar mejor los detalles y en el eje X tenemos las distancias (iteraciones) en bloques de 50 unidades. Destaca en la gráfica el cambio abrupto del valor de información mutua al rededor del valor $s = 0.3$.

Siguiendo el criterio planteado por Cellucci (2005) y de manera análoga a la figura 4.2 podemos establecer lo siguiente:

1. el valor mínimo de información mutua en las tres componentes (variables) coincide con el momento en que se alcanza un valor de susceptibilidad cercano a 0.3 para nuestra simulación.
2. Las tres componentes que representan las preferencias por cada partido presentan correlaciones no-lineales bajo el análisis efectuado en las simulaciones.

Conclusiones

Como primer aporte de este trabajo tenemos la construcción de un modelo no-ortodoxo que estudia la forma en que cambian las preferencias electorales basado en la racionalidad limitada de los agentes en el sentido de Simon (1957), basándose en las interacciones entre miembros de una comunidad (vecindad) y la susceptibilidad de los votantes empleando el enfoque de los sistemas complejos.

Entre las diferencias de nuestro modelo respecto a otros tenemos primero, el enfoque de racionalidad limitada empleada por los agentes; en segundo lugar la incorporación de un parámetro de susceptibilidad mediante el cual podemos representar una nueva gama de conductas de votantes que no se ha representado con los modelos tradicionales, con ello obtuvimos agentes "necios" que no cambian su preferencia y votantes "susceptibles", es decir aquellos que eran bastante proclives a cambiar de preferencia electoral.

Nuestro modelo puede servir para estudiar los procesos electorales de los últimos años en nuestro país, donde hemos observado una tendencia a desarrollar elecciones (presidenciales) entre tres fuerzas políticas; mientras estos procesos en países anglosajones se desarrollan entre dos fuerzas políticas. Aunque esto no impide usar nuestro modelo para un análisis con dos fuerzas políticas haciendo ajustes a las hipótesis.

Cabe señalar que con este modelo también podríamos estudiar las elecciones presidenciales de México para el presente proceso del 2018 considerando los escenarios pertinentes en las hipótesis, como el voto útil, el abstencionismo, las coaliciones de partidos etc.

Considerando que la información mutua alcanzó su valor mínimo (global) en las tres componentes cuando la susceptibilidad registró un valor de 0.3, podemos afirmar que hemos verificado la existencia de correlaciones no-lineales entre las 3 preferencias electorales del modelo. Sobre este hallazgo concluimos entonces que la toma de decisiones electorales posee las características de un sistema complejo como lo explica Solé (1993).

Con esta herramienta hemos podido determinar que el comportamiento de las decisiones electorales en el caso estudiado de tres partidos empleando racionalidad limitada, no se puede entender como un fenómeno que tiende al equilibrio como lo plantea desde la sociofísica Galam (2008), así queda corroborada nuestra hipótesis inicial, en donde las interacciones entre los votantes influyen en el proceso de toma de decisión; pero además, otro resultado producto de este trabajo es que bajo estas condiciones la elección se torna impredecible.

En nuestro cálculo del índice de información mutua, hemos observado correlaciones no-lineales en la evolución durante un año de las preferencias electorales para tres partidos. Queda pendiente para un futuro trabajo estudiar la posibilidad de que este fenómeno presente un comportamiento al "borde del caos" (Miramontes, 1999b).

La propuesta por lo pronto deja abierto el campo para que científicos sociales, ajusten las hipótesis de partida para otros procesos electorales de interés, pero sobre todo, para que permitan re-interpretar los resultados obtenidos en escenarios observables en la realidad y que no han sido considerados en las simulaciones computacionales.

Bibliografía

Arrow, Kenneth J. 1963. *Social Choice and Individual Values.* John Wiley and Sons.

Arthur, Brian. 1994. *Inductive Reasoning and Bounded Rationality.* Papers and Proceedings of the Hundred and Sixth Annual Meeting of the American Economic Association, vol. 84, no. 2. University of Michigan Press.

Arthur, Brian. 1997. *The Economy as an Evolving Complex System II with Steven Durlauf and David Lane.* Addison-Wesley, Reading MA.

Axelrod, Robert. 1997. Advancing the Art of Simulation in Social Sciences. *In: Simulating Social Phenomena.*

Barcelo, Bartolome. 2007. Sistemas Electorales. *MATerials MATematics.*

Black, Duncan. 1948a. On the rationale of group decision making. *Journal of Political Economy.*

Black, Duncan. 1948b. On the rationale of group decision-making. *Journal of Political Economy*, 23,34.

Borghesi A, Galam Serge. 2006. Chaotic staggered and polarized dynamics in opinion forming: the contrarian effect. *Physical Review E*, **73**(066118), 1 ,9.

Borner, Katy, *et al.* . 2009. Models of Science Dynamics Encounters Between Complexity Theory and Information Sciences. **3**, 161–172.

Cellucci, et al. 2005. Statistical validation of mutual information calculations. *Physical Review E*, **71**(006208).

Cocho, Germinal, *et al.* . 2017. *Ciencia Humanismo Sociedad. De los sistemas complejos a la imaginación heterodoxa.* México CDMX: CopIt-arXives y EditoraC3.

Cox, Gary. 1997. *Making votes count Strategic coordination in the worlds electoral system.* Cambridge University.

Downs, Anthony. 1957. *An economic theory of democracy*. New York.

Duverger, Maurice. 1957. *Political parties: Their organization and activity in the modern state*. New York.

Elster, Jon. 1991. *Juicios Salomónicos: las limitaciones de la racionalidad como principio de desición*. Barcelona.

Epstein. 2006. *Generative social science: Studies in agent-based computational modeling*. Princeton Studies in Complexity.

Galam. 2004a. The Dynamics of minority opinion in democratic debate. *Physica A 336*, 56–62.

Galam. 2007. From 2000 Bush-Gore to 2006 Italian elections: voting at fifty-fifty and the contrarian effect. *Quality and Quantity Journal*, **41**, 579–589.

Galam, S. 2004b. Contrarian deterministic effect: the hung elections scenario. *Physica A 333*, 453–360.

Galam, S. 2008. Sociophysics: A review of Galam models. *International Journal of modern Physics C19*, 409–440.

Gardner, M. 1970. Mathematical games: The fantastic combinations of John Conway's new solitaire game 'Life'. *Scientific American*, **223**(4), 120–123.

Gary, Cox. 1997. *Making votes count. strategic coordination in the world's electoral systems*. Cambridge University Press.

Gershenson, Carlos. 2013. Como hablar de complejidad. *Lengua, Societat i Comunicació*.

Gilbert, N. & Terna P. 2000. How to build and use agent-based models in social science. *Mind & Society Springer*, **1**(57). 1: 57. doi:10.1007/BF02512229.

Gilbert, Nigel. 1996. Simulation as a Research Strategy. *Pages 448–454 of: Social Science Microsimulation*.

Gilbert, Nigel. 2007. *Agent-Based Models. Quantitative Applications in the Social Sciences*. SAGE Publications.

Hotelling, Harold. 1929. Stability in competition. *Economic Journal*.

JM. Epstein, R. Axtell. 1996. *Growing artificial societies Social science from the bottom up*. Complex Adaptive Systems Brookings Institution Press, Washington, DC.

Miller, Leroy Roger. 1985. *Microeconomía*. Edit Trillas.

Miramontes, Octavio. 1999a. Los sistemas complejos como instrumentos de conocimiento y transformacion del mundo. *In: Perspectivas sobre la teoria de sistemas*. S. Ramirez.

Miramontes, Pedro. 1999b. El estructuralismo dinamico. *In: Perspectivas en la teoria de sistemas*. Santiago Ramirez Siglo XXI Editores y CEIICH UNAM.

Ponce, R. 2010. *Competencia política y las finanzas públicas de los gobiernos estatales y locales*. Universidad Autónoma de Ciudad Juárez.

Rissanen, Y. 1992. *Stochastic Complexity in Statistical Inquiry.* World Scientific.

Roemer, J. 2001. *Political competition Theory and applications.* Cambridge, Harvard University Press.

Schelling, T. 1969. Models of segregation. *Am Econ Rev*, **59**(2), 488–493.

Simon, Herbert. 1957. *Models of Man, Social and Rational: Mathematical Essays of Rational Human Behavior in a Social Setting*. Wiley NY. Chap. A Behavioral Model of Rational Choice.

Simon, Herbert. 1992. *Economics, bounded rationality and the cognitive revolution. With: M. Egidi, R. Marris y R. Viale*. M. Egidi y R. Marris.

Solé, Ricard. 1993. *Modelado de sistemas complejos mediante simulacion basada en agentes y mediante dinamica de sistemas*. Edit UPC.

Von Neumann, John. 1966. *Theory of self-reproducing automata*. University of Illinois Urbana.

Wikipedia. 2017 (Dec.). *Wikipedia*. `https://goo.gl/dV4Sgw`. Accesado 12-abril-2018.

Wittman, D. 1983. Candidate motivation: A synthesis of alternatives theories. *American Political Science Review*, 142–157.

Wittman, D. 1990. *Spatial strategies when candidates have policy preferences*. Cambridge University Press.

Wittman, Donald. 1973. Parties as utility maximizers. *American Political Science Review*, 490–498.

Wolfram, Stephen. 2002. *A New kind of Science*. Wolfram.

Anexos

A. Teorema de Arrow

Antes de definir un sistema de votación de manera precisa daremos las definiciones y suposiciones necesarias.

Sea $S = \{v_1, v_2, \ldots, v_k\}$, *con* $|S| = k \in \mathbb{N}$ el conjunto finito de ***votantes*** en una sociedad. Sea $C = \{c_1, c_2, \ldots c_m\}$, *con* $|C| = m \in \mathbb{N}$ el conjunto finito de ***candidatos*** (conocido como alternativas u opciones de voto)

Cada votante v_i es capaz de elegir racionalmente sus preferencias de voto en el sentido de la Racionalidad Completa explicada en el capítulo 1

Definimos el ***orden de preferencia*** $\underset{v_i}{>}$ para cada votante $v_i \in S$ de la siguiente forma: para dos candidatos $a, b \in C$, escribimos $\mathbf{a} \underset{\mathbf{v_i}}{>} \mathbf{b}$ siempre que el votante v_i prefiere la opción a sobre la opción b , para denotar lo contrario simplemente escribimos $\mathbf{b} \underset{\mathbf{v_i}}{>} \mathbf{a}$.

Con esta definición los *ordenes de preferencia definidos individualmente por los votantes son transitivos*, esto es: $\forall\, v_i \in S$ y para $a, b, c \in C$ si $a \underset{v_i}{>} b \wedge b \underset{v_i}{>} c \Rightarrow a \underset{v_i}{>} c$.

Lema 1 $E = \left\{ \underset{v_j}{>} :\ v_j \in S,\ con\ \underset{v_j}{>}\ transitiva \right\}$ El conjunto de las relaciones de preferencia de todos los votantes en el sistema es un conjunto totalmente ordenado

Demostración del Lema1 Cada elemento de E es un orden $c_1, c_2 \ldots c_m$ que refleja la preferencia de algún votante v_j; el conjunto E contiene todas las preferencias de los votantes v_j de S respecto a las opciones de voto en C.
Supongamos que E no esta totalmente ordenado.

Sean $a, b \in C$ cualesquiera dos candidatos y una relación de preferencias $\succ \in E$.

$\succ$ es transitivo, por lo cual $\exists\, p \in C \perp a \succ p \ \wedge\ p \succ b \Rightarrow a \succ b \quad (l_1)$
con el mismo razonamiento podemos afirmar $\exists\, q \in C \perp b \succ q \ \wedge\ q \succ a \Rightarrow b \succ a \quad (l_2)$

Siendo a *y* b arbitrarios, de (l_1) y (l_2) tenemos que E esta completamente ordenado, contradiciendo la hipótesis inicial
$\therefore E$ es un conjunto totalmente ordenado. $\square$

Dado el conjunto E del Lema 1 y $E^k = (E \times E \times \ldots \times E) \quad k - veces\, E$ definimos como sistema de votaciones el mapeo $\boldsymbol{V} : E^k \longrightarrow E$.

Es decir, un *sistema de votación* V es una función cuyas entradas son las preferencias $\underset{v_j}{>}$ de cada votante v_j y cuya salida devuelve algún orden de los candidatos.

De acuerdo con *"Social Choice and Individual Values "(1951) de Kenneth Arrow, un sistema de votaciones debe contener las siguientes propiedades.*

- **Unanimidad (U)** Si para algunos $a, b \in C$ tenemos que $\forall v \in S$ sucede $a \underset{v}{>} b$ entonces $\exists\, \hat{e} \perp \hat{e} = \boldsymbol{V}(E_1, E_2, \ldots E_k)$ y $\hat{e}$ satisface la preferencia $a \underset{v}{>} b$. En otras palabras, dado el sistema de votacion $\boldsymbol{V}$, cuando se aplica a la sociedad S que por unanimidad prefiere al candidato a sobre b , se tiene como salida $\hat{e}$. Podemos decir que, S siempre prefiere $a \underset{v}{>} b$, si cada individuo prefiere $a \underset{v}{>} b$.
- **Racionalidad (R) o Transitividad**. Todas las preferencias de los electores estan totalmente ordenadas (por el Lema1), entonces la salida del sistema $V(E_1 \times E_2 \times \ldots \times E_k) = \hat{e}$ cumple con la propiedad transitiva.
- **No existe un dictador (ND)**. Para algún $1 \leq i, j \leq k$, $i \neq j$ y para cualesquiera $a, b \in C$, siempre que e_i satisfaga $a \underset{v_i}{>} b$ y que exista algun otro e_j que satisfaga $b \underset{v_j}{>} a$, si el sistema de votación $V(E_1 \times E_2 \times \ldots \times E_k) = \hat{e}$ da como resultado ganador $a > b$ entonces el votante v_i se conoce como *dictador*. Un sistema de votación «razonable» no debe tener dictadores. Por lo tanto el resultado del sistema de votación no debe satisfacer los deseos de un votante anulando los deseos de los demás. Para que la propiedad ND tenga sentido, se requiere de *monotonicidad,* esto es: si $a, b \in C$ son tales que $V(E_1 \times E_2 \times \ldots \times E_k) = \hat{e}$ da como resultado ganador $b > a$ entonces disminuir la posición relativa de a (con

respecto a b) en cualquiera de los preferencias e_i , no debería alterar la posición relativa de a en el resultado de la votación.

- **Independencia de las alternativas irrelevantes (IAI)** Para $a, b, c \in C$ la preferencia entre a y b en el resultado de la votación $V(E_1 \times E_2 \times \ldots \times E_k) = \hat{e}$ debe ser idéntico a la preferencia correspondiente cuando la posición de c en $\hat{e}$ se ha cambiado. Esto es, si socialmente se ha elegido $a > b$ entonces la posición que c tiene en $\hat{e}$ no debería ser relevante considerando que se realizó la elección $a > b$ en $\hat{e}$. Si el sistema de votaciones tiene sólo 2 candidatos la propiedad IAI se obtiene de inmediato, pues no hay terceras alternativas que elegir

A.1. Demostración del Teorema de Arrow

Teorema (Arrow). Sea un sistema de votaciones como en el apartado anteiror, Si $|C| \geq 3$ **, entonces las propiedades** U, R, ND, IAI **son inconsistentes.**[1]

Definición. Para algún $a, b \in C$ llamamos *Quorum al conjunto* $Q \subseteq S$, de forma que si la elección de todo votante $v_q \in Q$ satisface $a \underset{q}{>} b$ y la elección de todo $v \in \{S \setminus Q\}$ satisface $b \underset{s}{>} a$, entonces la salida del sistema de votaciones $V(E_1 \times E_2 \times \ldots \times E_k) = \hat{e}$ satisface $a > b$.

Esto significa, que los miembros del quorum que prefieren $a \underset{q}{>} b$ pueden forzar $a > b$ en la elección social por medio de la elección unánime $a > b$ en sus elecciones individuales. Sea $R(a, b) = \left\{ Q \subseteq S : \quad \forall q \in Q \; con \; a \underset{q}{>} b \;\wedge\; \forall v \in \{S \setminus Q\} \;\; b \underset{v}{>} a \right\}$

el conjunto de todos los quorums de a, b. Notar que por (**U**), S es un Quorum para cualesquiera dos candidatos.

Por (**IAI**), la posibilidad de que $Q \in R(a, b)$ pueda o no forzar $a > b$ depende solode las preferencias individuales de a, b para cada $q \in Q$. Por lo anterior la noción de Quorum esta bien definida; en otras palabras, un Quorum no puede forzar $a > b$ a veces sí y a veces no, si se observa que todos sus miembros han elegido $a > b$.

Cabe notar que cada par de candidatos tiene una noción distinta de Quorum, Q no tiene que ser quorum para otro par c, d si $Q \in R(a, b) \nRightarrow Q \in R(c, d)$, esto es, *existe imparcialidad de candidatos*

Lema 2: La intersección de quorums es transitiva Si $Q_1 \in R(a, b)$ y $Q_2 \in R(b, c) \Rightarrow Q_1 \cap Q_2 \in R(a, c)$ para cualesquiera candidatos $a, b, c \in C$

Demostración del Lema 2:

Supongamos $\forall\, v \in Q_1 \in R(a, b)$ *tenemos* $a > b$ ***(i)***

$\forall\, s \in Q_2 \in R(b, c)$ *tenemos* $b > c$ ***(ii)***

$\forall\, w \in Q_3 \in S \setminus (Q_1 \cap Q_2)$ *tenemos* $c > a$

De (i) y (ii) tenemos $a > b \quad \wedge \quad b > c \;\Rightarrow a > c$ por transitividad

[1] consultar T. Tao « Arrow´s Theorem» UCLA University

$$\Rightarrow \left\{ a,\, b,\, c \in C \; con \; a \underset{v_j}{>} b \; \wedge b \underset{v_j}{>} c \Rightarrow a \underset{v_j}{>} c \; : \quad a \tilde{>} b \right\}$$
$$= \left\{ a,\, b,\, c \in C \; : \; v_j \in [Q_1 \in R(a,a)] \cap [Q_2 \in R(b,c)] \right\}$$
$$\Rightarrow \left\{ Q_1 \cap Q_2 \right\} \in R(a,c) \qquad \Rightarrow Q_1 \cap Q_2 \text{ es quorum } \square.$$

Lema 3: Para cualquier $\hat{s} \in S$ **el conjunto** $\left\{ S \setminus \{\hat{s}\} \right\}$ **es quorum de cualesquiera candidatos** $a, b, c \in C$**.**

Demostración del lema 3: Supongamos lo contrario: $\left\{ S \setminus \{\hat{s}\} \right\}$ no es quorum.

Sean $a, b \in C$; como $\left\{ S \setminus \{\hat{s}\} \right\}$ no es quorum *tenemos* $\neg \left\{ S \setminus \{\hat{s}\} \right\}$ es quorum , esto siginifica:

$$\neg \left[\exists v_j \in \left\{ S \setminus \{\hat{s}\} \right\} \quad \wedge \quad \exists v_k \in \{\hat{s}\} \; : \; a \underset{v_j}{>} b \; \wedge b \underset{v_k}{>} a \Rightarrow a \tilde{>} b \right] \qquad \text{esto equivale a}$$

$$\forall v_j \in \left\{ S \setminus \{\hat{s}\} \right\} \quad a \underset{v_j}{>} b \; \vee \forall v_k \in \{\hat{s}\} \quad : \; b \underset{v_k}{>} a \; \vee a \underset{v_k}{>} b \Rightarrow \neg \left(a \tilde{>} b \right) \quad \Rightarrow b \tilde{>} a$$

$$\Rightarrow \{\hat{s}\} \; es \; Quorum \Rightarrow \hat{s} \in R(a,b) \; y \left\{ S \setminus \{\hat{s}\} \right\} \in R(a,c).$$

Por el Lema2 $\{\hat{s}\} \cap \left\{ S \setminus \{\hat{s}\} \right\} = \emptyset \in R(a,c)$ lo cual no puede suceder pues se ***contradice la condición (U)*** .

$\therefore \quad \left\{ S \setminus \{\hat{s}\} \right\}$ es quorum para cualquier $a, b, c \in C$ $\square$

Afirmación: $S \setminus \{\hat{s}\}$ debe ser quorum para $a, b \in C$ con $c \neq b$ del caso anterior
Supongamos lo contrario, entonces $\hat{s}$ es dictador con $a \underset{v_j}{>} c$ además *por unanimidad* $(a, c_1), (a, c_2) \ldots, (a, c_k)$ $\forall k \Rightarrow \hat{s}$ es dictador $\forall c \in C$ lo cual **contradice la condición (ND).**

Como $\hat{s} \in R(a,b) \quad \wedge \quad S \setminus \{\hat{s}\} \in R(a,c)$ el conjunto $\{\hat{s}\} \bigcap (S \setminus \{\hat{s}\}) = \emptyset \in R(b,c)$lo cual **contradice la condición (U)**

$\therefore$ para cualquier $\hat{s} \in S$ el conjunto $S \setminus \{\hat{s}\}$ es quorum para cualquier $a, b, c \in C$.

Cabe destacar, si $S = \{v_1, v_{2,} \cdots, v_k\}$ y $a, b, c \in C$ con $|C| \geq 3$ entonces tenemos $S \setminus \{v_1\} \in R(a, b)$ y además $S \setminus \{v_i\} \in R(b, c) \quad i \neq 1$ implican que $\bigcap_{i=1}^{k}(S \setminus \{v_i\}) = \emptyset$ es quorum ya sea de $Ra, b)$ *o* $R(a, c)$ lo cual **contradice la condición (U).**

Por todo lo anterior concluimos por lo tanto, **las condiciones (U), (T), (ND), (IAI)** son inconsistentes $\square$.

B. Información Mutua

Definamos un sistema X donde el *alfabeto* esta formado por el conjunto de símbolos $\{x_1, x_{2,\ldots}, x_{N_X}\}$, del cual se puede obtener una secuencia de salida, o mensaje generado por el sistema X, denotamos N_x el número de elementos del conjunto de símbolos y empleamos N_D para denotar la longitud del mensaje.

La probabilidad asociada a cada uno de los elementos del conjunto de símbolos es $\{P_X(1), P_X(2), \ldots, P_X(N_x)\}$, el cual tiene la probabilidad $\sum_{i=1}^{N_X} P_X(i) = 1$

El símbolo x_i denota el evento en la i-ésima componente del histograma el cual contiene en total N_X partes. De esta forma $P_X(i)$ es la probabilidad de que el i-ésimo espacio se encuentre ocupado.

$I(i)$ **denota la *información mutua*** Este describe el contenido de información del i-ésimo símbolo y se calcula con la fórmula: $I(i) = -\log_2 P_X(i)$ notemos que el cálculo de la información se mide en bits. Supongamos que la probabilidad asociada a un símbolo es de 1, en este caso, no se aprende nada de la observación de este símbolo. La correspondiente información es $-\log_2(1) = 0$. Conforme un símbolo se vuelve crecientemente improbable, la información que se le asocia se incrementa.

$H(X)$ **denota la entropía de un sistema** La información es una propiedad asociada a un símbolo. En contraste, la entropía es la propiedad asociada a un sistema, $H(X)$ es la cantidad promedio de información ganada por la observación de x , en otras palabras $H(X)$ es la incertidumbre promedio de x antes de su observación.

Para calcular la entropía tenemos

$$H(X) = \sum_{i=1}^{N_X} P_X(i) I(i) = -\sum_{i=1}^{N_X} P(i) \log_2(P_X(i))$$

Supongamos que en un cálculo de entropía podemos realizar una observación dentro de un total de sesenta y cuatro posibles casos, tendremos de aquí los símbolos $\{x_1, x_{2,\ldots}, x_{N_{64}}\}$, partiendo de que $P(i) = \frac{1}{64}$ para todo i.

$$H(X) = -\sum_{i=1}^{N_{64}} P(i)\log_2(P_X(i)) = -64\left(\frac{1}{64}\right)\log_2\left(\frac{1}{64}\right) = 6 \text{ bits}$$

El valor máximo de la entropía de un sistema se obtiene cuando todos los símbolos poseen la misma probabilidad. Por otro lado $\lim_{x\to 0} x\log x = 0$ con lo cual si $P_X(j) = 1$ y $P_X(k) = 0$ para toda $k \neq j$, entonces $H(X) = 0$.

Probabilidad conjunta $P_{XY}(i,j)$ Consideremos dos sistemas X, Y con alfabetos

$\{x_1, x_2, \ldots, x_{N_D}\}$ y $\{y_1, y_2, \ldots, y_{N_D}\}$ respectivamente. Cuando definimos una distribución de probabilidad conjunta, debemos hacer la distinción entre N_X el número de símbolos diferentes que pueden presentarse en el sistema X , $N_Y$el número de símbolos diferentes que pueden presentarse en el sistema Y, y N_D el número de pares (x, y) observados. La probabilidad de cada símbolo independiente del sistema X esta denotada por $\{P_X(1), P_X(2), \ldots, P_X(N_x)\}$. Similarmente la distribución de probabilidad independiente de Yes $\{P_Y(1), P_Y(2), \ldots, P_Y(N_Y)\}$.

La distribución de probabilidad conjunta $P_{XY}(i,j)$ es la probabilidad asociada a un par (x, y) del i-ésimo símbolo del sistema X y del j-ésimo símbolo del sistema Y. En donde:

$$P_{XY}(i,j) = P_X(i)P_Y(j) \Longleftrightarrow X, Y \text{ son independientes.}$$

Lema. Relación entre distribución de probabilidad conjunta y distribución de una variable

$$P_X(i) = \sum_{j=1}^{N_Y} P_{XY}(i,j)$$

Demostración del Lema Mediante la suma de todos los posibles valores de probabilidad, sin importar el caso, la probabilidad de y es 1, por lo tanto el valor remanente de la suma es la probabilidad del sistema X $\square$.

Distribución de Probabilidad Condicional Dados los sistemas X, Y antes definidos la distribución de probabilidad condicional se denota por $P_{X|Y}$. Escribimos $P_{X|Y}(i,j)$ para calcular la probabilidad de que $X = x_i$, y análogamente sucede lo mismo para el caso $Y = y_j$. La relación entre la distribución de probabilidad conjunta y la distribución de probabilidad condicional se establece en el siguiente lema.

Lema La relación entre la distribución condicional de probabilidad y la distribución de probabilidad conjunta esta dada por

$$P_{XY}(i,j) = P_{X|Y}(i,j)P_Y(j)$$

Demostración del lema $P_Y(j)$ es la probabilidad de que $Y = y_j$. $P_{X|Y}(i,j)$ es la probabilidad de que $X = x_i$ si ya se sabe $Y = y_j$. Por lo tanto el producto de estas probabilidades es la probabilidad de que $X = x_i$ junto con $Y = y_j$, la cual por definición es $P_{XY}(i,j)$.

Si X y Yson independientes entonces $P_{XY}(i,j) = P_X(i)P_Y(j)$, y $P_X(i) = P_{X|Y}(i,j)$. Esto es, si X y Yson independientes, entonces la probabilidad de que $X = x_i$ esta determinada solamente por el sistema X $\square$.

Entropía conjunta Dados los sistemas X, Y antes definidos, la entropía conjunta se define como

$$H(X,Y) = -\sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} P_{XY}(i,j) \log_2(P_{XY}(i,j))$$

$H(X,Y)$ es la cantidad promedio de información ganada por observar un par (x, y).

Lema La entropía conjunta de un sistema respecto a si mismo esta dada por

$$H(X,X) = H(X)$$

Demostración: Por definición $H(X,Y) = -\sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} P_{XY}(i,j) \log_2(P_{XY}(i,j))$. Si $Y = X$, entonces $P_{XY}(i,j) = \delta_{ij} P_X(i)$ donde δ_{ij} es la delta de Kronecker

$H(X,Y) = -\sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} \delta_{ij} P_X(i) \log_2(\delta_{ij} P_X(i))$ tomando $\lim_{z \to 0} z \log z = 0$ con lo cual obtenemos

$H(X,X) = -\sum_{i=1}^{N_X} P_X(i) \log_2(P_X(i)) = H(X)$ $\square$

Lema La función de entropía conjunta es simétrica, esto es:

$$H(X,Y) = H(Y,X)$$

Demostración: Por definición $H(X,Y) = -\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y} P_{XY}(i,j)\log_2(P_{XY}(i,j))$, además la función de distribución conjunta no es una distribución de probabilidad condicional, por ello $P_{XY}(i,j) = P_{YX}(j,i)$. Por lo tanto $H(X,Y) = H(Y,X)$ $\square$.

Entropía condicional Dado un valor de Y digamos $Y = y_j$, la entropía condicional se define como

$$H(X \mid j) = -\sum_{i=1}^{N_X} P_{X|Y}(i,j)\log_2(P_{X|Y}(i,j))$$

$H(X \mid j)$ es la cantidad promedio de información obtenida por la observación de X cuando $Y = y_j$. $H(X \mid Y)$ la entropía condicional promedio, es el cálculo de $H(X \mid j)$ promediada sobre todos los posibles valores de y .

$$H(X \mid Y) = -\sum_{j=1}^{N_Y} P_Y(j)H(X \mid j)$$

$H(X \mid Y)$ es la información promedio obtenida por observar X después de que es conocida Y . En otros términos, $H(X \mid Y)$ es el número promedio de bits adicionales requeridos para especificar X si Y es conocida.

Lema La relación entre la entropía condicional y la entropía conjunta esta dada por

$$H(X \mid Y) = H(X,Y) - H(Y)$$

Demostración: Por definición $H(X \mid Y) = -\sum_{j=1}^{N_Y} P_Y(j)H(X \mid j)$ donde $H(X \mid j) = -\sum_{i=1}^{N_X} P_{X|Y}(i,j)\log_2(P_{X|Y}(i,j))$, además como ya hemos demostrado $P_{XY}(i,j) = P_{X|Y}(i,j)P_Y(j)$ entonces

$$H(X \mid Y) = \sum_{j=1}^{N_Y} P_Y(j)(-1)\sum_{i=1}^{N_X} \frac{P_{XY}(i,j)}{P_Y(j)}\log_2\frac{P_{XY}(i,j)}{P_Y(j)}$$

$$= -\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y} P_{XY}(i,j)\log_2 \frac{P_{XY}(i,j)}{P_Y(j)}$$

$$= -\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y} P_{XY}(i,j)\log_2(P_{XY}(i,j)) + \sum_{i=1}^{N_X}\sum_{j=1}^{N_Y} P_{XY}(i,j)\log_2(P_Y(j))$$

El primer termino de la derecha es la entropía conjunta por lo que podemos reescribir

$H(X \mid Y) = H(X,Y) + \sum_{j=1}^{N_Y}\left\{\sum_{j=1}^{N_Y} P_{XY}(i,j)\right\}\log_2 P_Y(j)$

Anteriormente demostramos $P_Y(j) = \sum_{i=1}^{N_X} P_{XY}(i,j)$ por lo tanto

$H(X \mid Y) = H(X,X) + \sum_{j=1}^{N_Y} P_Y(j)\log_2 P(j)$

$H(X \mid Y) = H(X,X) - H(Y) \qquad \square.$

Definición. Información Mutua $I(X,Y)$:

Retomando conceptos, tenemos

$H(X)$ = cantidad promedio de información obtenida por una observación de X

$H(X \mid Y)$ = cantidad promedio de información sobre X obtenida al observar X después de que se conoce Y.

Finalmente tenemos que la información mutua $I(X,Y)$ = cantidad promedio de información sobre X obtenida por la observación de Y.

$H(X) = I(X,Y) + H(X \mid Y)$

$I(X,Y) = H(X) - H(X \mid Y)$

B.1. Propiedades de la Información Mutua

Lema 1. $H(X) = I(X, X)$ **se conoce como autoinformación.**

Lema 2. $I(X, Y) = H(X) + H(Y) - H(X, Y)$ **,**

aquí la información mutua depende de la entropía conjunta

Lema 3. $I(X, Y) = I(Y, X)$,**la información mutua es simétrica**

Lema 4. $I(X, Y) = H(Y) - H(Y \mid X)$

Lema 5. $I(X, Y) = \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} P_{XY}(i, j) \log_2 \left\{ \frac{P_{XY}(i,j)}{P_X(i) P_Y(j)} \right\}$

Lema 6.$I(X, Y) = 0$ **Si** X **y** Y **son independientes.**

De interés para los cálculos es el lema 5 cuya demostración se describe a continuación.

Demostración Lema5. En términos prácticos este lema dice que podemos expresar la información mútua en términos de probabilidad y de la distribución de probabilidad conjunta.

Desarrollando la expresión derecha de la ecuación tenemos:

$* = \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} P_{XY}(i, j) \log_2 P_{XY}(i, j)$

$- \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} P_{XY}(i, j) \log_2 P_X(i) - \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} P_{XY}(i, j) \log_2 P_Y(j)$

Empleando la definición de entropía conjunta y reescribiendo lo términos obtenemos:

$* = H(X, Y) - \sum_{i=1}^{N_X} \log_2 P_X(i) \sum_{j=1}^{N_Y} P_{XY}(i, j) - \sum_{j=1}^{N_Y} \log_2 P_Y(j) \sum_{i=1}^{N_X} P_{XY}(i, j)$

Anteriormente demostramos $P_X(i) = \sum_{j=1}^{N_Y} P_{XY}(i, j)$ asi como $P_Y(j) = \sum_{i=1}^{N_X} P_{XY}(i, j)$ con lo cual podemos simplificar la última expresión

$* = -H(X, Y) - \sum_{i=1}^{N_X} P_X(i) \log_2 P_X(i) - \sum_{j=1}^{N_Y} P_Y(j) \log_2 P_Y(j)$

$* = H(X) + H(Y) - H(X, Y) = I(X, Y) \qquad \square.$

C. Listado del programa principal

Listados del programa totalmutinf.m en Matlab:

```
function [ salida ] = totalmutinf( st,cp,nb1 )
%Calcula la función de informacion mutua
%para cada valor de la suceptibilidad 'st'
%y cada partido 'cp'.
%Utiliza el metodo propuesto por Cellucci
%et al. para elegir el tamaño optimo
%de los histogramas. Puede utilizarse
%tambien el metodo sugerido
%por Rissanen.
% %Referencias
%T. M. Cover & J. A. Thomas, "Elements of
%Information Theory", Wiley, 1991.
%Cellucci, et al., "Statistical validation
%of mutual information calculations"
%Physical Review E, vol. 71, 066208, 2005.
%Y. Rissanen, "Stochastic Complexity in
%Statistical Inquiry", World Scientific 1992, pag. 76.
%Definiendo entrada y salida de datos
path='C:\tesis\joaquin\datos\eleccN1000\'; s=st; comp=cp;
%Numero de divisiones del histograma p(i,j)
Nb=nb1; %<—— Actualizar aqui el tamaño del histograma
%Aqui se puede utilizar el metodo de Rissanen o el propuesto por
%Cellucci et al. usando una particion adaptativa.
coor=str2double(comp); file=strcat('datos',s);
% file=strcat('largo'); fid1=fopen(strcat(path,file,'.txt'),'rt');
```

```
if fid1~=-1
[x,count]=fscanf(fid1,'%g %g %g',[3,1095]);z=x';end;dat=z(:,coor);
[n1 n2]=size(dat);      nn=max(n1,n2);
filename2=strcat('C:\mutinfN1000\','mif_',s,'_',comp,'.txt');
fid2=fopen(filename2,'w');    %<---- Actualizar aqui salida
DistMax=min(floor(nn/3),1000);    %DistMax=500;
fprintf('%g\r',DistMax);  %<---- Optativo.
%Solo para saber la cantidad de  distancias a calcular.
%************************************************
%INICIO DEL PROGRAMA PRINCIPAL
%************************************************
%Inicializando las estructuras de datos
%Creando la particion para los histogramas
Exti=min(dat);   Exts=max(dat);
Hp=(Exts-Exti)/Nb; edge=zeros(1,Nb+1);
for k=1:Nb+1,
edge(k)=Exti+(k-1)*Hp;
end;
[e1 e2]=size(edge);      ee=max(e1,e2);
m=zeros(1,DistMax);
for d=1:DistMax,           x=dat(1:nn-d);           y=dat(1+d:nn);
%Calculando los histogramas
ox=histc(x,edge);          ox=ox/sum(ox);
oy=histc(y,edge);          oy=oy/sum(oy);
oxy=zeros(ee,ee);
for i=1:nn-d,
c1=floor( ((x(i)-Exti)/Hp) )+1;
c2=floor( ((y(i)-Exti)/Hp) )+1;
oxy(c1,c2)=oxy(c1,c2)+1;
end;
oxy=oxy/sum(sum(oxy));
for i=1:ee
 for j=1:ee
  if (ox(i)*oy(j)~=0) && (oxy(i,j)~=0)
```

```
        m(d)=m(d)+oxy(i,j)*log(oxy(i,j)/(ox(i)*oy(j)));
end;
end;
end;
fprintf(fid2,'%f\n',m(d));          fprintf('%g\r',d);
end;
%*************************************************
%                          FIN DEL PROGRAMA PRINCIPAL
%*************************************************
%Cierre del file de datos.
fclose(fid1);     fclose(fid2);
salida=count;
end
```

C.1. Listado del programa Mutinf

```
%Program Mutinf
%Calcula la funcion de informacion mutua
%para la serie de los datos proveniente de
%las componentes del sistema. Utiliza el
%metodo propuesto por Cellucci et al. para
%elegir el tamaño optimo de los histogramas.
%Puede utilizarse tambien el metodo sugerido
%por Rissanen.
%   %Referencias
%T. M. Cover & J. A. Thomas, "Elements of
%Information Theory", Wiley, 1991.
%Cellucci, et al., "Statistical validation
%of mutual information calculations"
%Physical Review E, vol. 71, 066208, 2005.
%Y. Rissanen, "Stochastic Complexity in
%Statistical Inquiry", World Scientific 1992, pag. 76.
%Definiendo entrada y salida de datos
path='C:\datos\';    s='01';    comp='3';
%Numero de divisiones del histograma p(i,j)
Nb=34;%<----- Actualizar aqui el tamaño del histograma
%Aqui se puede utilizar el metodo de Rissanen o el propuesto por
Cellucci et al. usando una particion adaptativa.
coor=str2double(comp);    file=strcat('datos',s);
%file=strcat('largo');   fid1=fopen(strcat(path,file,'.txt'),'rt');
if fid1~=-1
[x,count]=fscanf(fid1,'%g %g %g',[3,1095]);
z=x';
end;
dat=z(:,coor); [n1 n2]=size(dat);  nn=max(n1,n2);
filename2=strcat('C:\','mif_',s,comp,'.txt');
fid2=fopen(filename2,'w');  %<---- Actualizar aqui salida
DistMax=min(floor(nn/3),1000);     %DistMax=500;
```

```
fprintf('%g\r',DistMax);%<----Optativo.
%Solo para saber la cantidad de distancias a calcular.
%**************************************************************
%                       INICIO DEL PROGRAMA PRINCIPAL
%**************************************************************
% Inicializando las estructuras de datos
% Creando la particion para los histogramas
Exti=min(dat);     Exts=max(dat);     Hp=(Exts-Exti)/Nb;
edge=zeros(1,Nb+1);
for k=1:Nb+1,
edge(k)=Exti+(k-1)*Hp;
end;
[e1 e2]=size(edge);     ee=max(e1,e2);     m=zeros(1,DistMax);
for d=1:DistMax,
x=dat(1:nn-d);
y=dat(1+d:nn);
% Calculando los histogramas
ox=histc(x,edge);          ox=ox/sum(ox);
oy=histc(y,edge);          oy=oy/sum(oy);          oxy=zeros(ee,ee);
for i=1:nn-d,
c1=floor( ((x(i)-Exti)/Hp) )+1;
c2=floor( ((y(i)-Exti)/Hp) )+1;
oxy(c1,c2)=oxy(c1,c2)+1;
end;
oxy=oxy/sum(sum(oxy));
for i=1:ee
for j=1:ee
if (ox(i)*oy(j)~=0) && (oxy(i,j)~=0)
m(d)=m(d)+oxy(i,j)*log(oxy(i,j)/(ox(i)*oy(j)));
end;
end;
end;
fprintf(fid2,'%f\n',m(d));
fprintf('%g\r',d);
```

```
end;
primin=find(m==min(m));
plot(m);
%*******************************************************************************
%                              FIN DEL PROGRAMA PRINCIPAL
%*******************************************************************************
%Cierre del file de datos.
c1=fclose(fid1);      c2=fclose(fid2);
end
```