



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

---

FACULTAD DE CIENCIAS

REDES DIFUSAS ALEATORIAS

T E S I S

QUE PARA OBTENER EL GRADO DE:  
LICENCIADO EN MATEMÁTICAS

PRESENTA:  
OCTAVIO BALTASAR ZAPATA FONSECA

DIRECTOR DE TESIS:  
DR. CARLOS GERSHENSON GARCÍA



CIUDAD DE MÉXICO

MAYO, 2014



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# Redes Difusas Aleatorias

por

Octavio Baltasar Zapata Fonseca

Tesis presentada para obtener el grado de

Licenciado en Matemáticas

en la

FACULTAD DE CIENCIAS

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad de México. Mayo, 2014

1. Datos del alumno

Zapata

Fonseca

Octavio Baltasar

55 49 01 13

Universidad Nacional Autónoma de

México

Facultad de Ciencias

Matemáticas

305614430

2. Datos del tutor

Dr

Carlos

Gershenson

García

3. Datos del sinodal 1

Dr

Pedro Eduardo

Miramontes

Vidal

4. Datos del sinodal 2

Dr

Ramsés Humberto

Mena

Chávez

5. Datos del sinodal 3

Dr

Francisco

Hernández

Quiroz

6. Datos del sinodal 4

Dr

David Arturo

Rosenblueth Laguette

7. Datos del trabajo escrito

Redes Difusas Aleatorias

68 p

2014

# Índice general

<b>1. Lógica Difusa</b>	<b>2</b>
1.1. Introducción . . . . .	2
1.2. Objetivos . . . . .	4
1.3. Conjuntos . . . . .	5
1.3.1. Conceptos Básicos sobre Conjuntos Difusos . . . . .	6
1.4. Álgebras de Heyting . . . . .	8
1.4.1. Latices . . . . .	8
1.4.2. Ejemplos . . . . .	12
<b>2. Producto Semi-tensorial de Matrices</b>	<b>15</b>
2.1. Funciones Lógicas . . . . .	15
2.2. Información Multi-dimensional . . . . .	16
2.3. Índices Ordenados . . . . .	18
2.4. Producto Semi-tensorial . . . . .	20
2.5. Representación Matricial de Funciones Lógicas . . . . .	22
2.5.1. Matrices Lógicas . . . . .	23
2.5.2. Ecuaciones Lógicas . . . . .	25
<b>3. Redes Booleanas Aleatorias</b>	<b>29</b>
3.1. El Modelo Clásico . . . . .	30
3.2. Autoorganización en Redes Booleanas . . . . .	32
3.2.1. Comportamiento Dinámico . . . . .	33
3.2.2. Memoria Lejos del Equilibrio . . . . .	34

3.3. Enfoque Sistémico . . . . .	35
3.3.1. Dinámica de redes Booleanas . . . . .	35
3.3.2. Puntos Fijos y Ciclos . . . . .	39
3.4. Transiciones de Fase . . . . .	43
<b>4. Redes Difusas Aleatorias</b>	<b>45</b>
4.1. Modelos Booleanos y Continuos . . . . .	46
4.1.1. Método Sintético . . . . .	46
4.1.2. Idealización Booleana . . . . .	47
4.2. Redes Difusas Aleatorias . . . . .	48
4.2.1. Estructura . . . . .	48
4.2.2. Difusificación . . . . .	49
4.3. Familias de Atractores . . . . .	51
4.3.1. Espacio de Estados . . . . .	51
4.3.2. Atractores . . . . .	51
4.3.3. Familias . . . . .	52
<b>5. Simulaciones y Resultados</b>	<b>55</b>
5.1. Algoritmos . . . . .	55
5.2. Atractores . . . . .	56
5.3. Familias . . . . .	57
5.4. Dinámica . . . . .	58
<b>6. Conclusiones</b>	<b>64</b>

# Prefacio

La ubicua interconectividad de los fenómenos que ocurren en el mundo real aunada a la directa posibilidad de tomar en cuenta una gran variedad de interacciones, ha guiado a científicos e ingenieros a utilizar modelos con organización de *red* para representar procesos complejos e investigar la relación entre sus características funcionales y estructurales. Más aún, es evidente que *la globalización está conduciendo a que todos los sistemas sean cada vez más interdependientes* [Gershenson *et al.*, 2014], así, consideramos conveniente en este trabajo estudiar redes.

Incrementar el entendimiento acerca de cómo ciertos cambios en objetos individuales pueden sobrepasar a escalas más generales e influir en el comportamiento de un sistema es, fundamental para comprender los mecanismos que subyacen el funcionamiento global y, a su vez, una de las principales características del desarrollo intelectual y social del ser humano. Consecuentemente, la ciencia de los sistemas se encuentra claramente frente al reto de representar, con la intención de analizar y controlar, el comportamiento dinámico de estas complejas interdependencias. En este trabajo presentaremos dos modelos de sistemas dinámicos discretos. Ambos son un tipo de red dinámica –redes en las que el *estado* de cada *nodo* cambia a través del tiempo– cuya estructura y funcionalidad son generadas de forma aleatoria.

Las redes booleanas aleatorias (RBAs), creadas por Stuart Kauffman [Kauffman, 1969] para modelar las interacciones en redes de regulación genética, se han convertido en útiles modelos para la investigación de sistemas complejos debido a que tienen la capacidad de exhibir un comportamiento similar al de organismos vivientes. Nosotros extendemos este modelo utilizando conjuntos difusos que nos permiten cuantificar el grado de *complejidad* de una red y medir cuanta información se pierde durante el proceso de *simplificación booleana*. Exponemos algunos conceptos básicos sobre lógica difusa, así como su importancia y utilidad, en el capítulo 1.

En el capítulo 3 damos una introducción a las RBAs y posteriormente presentamos un estudio sistémico de sus principales conceptos, donde una relación lógica es expresada como una ecuación algebraica, y un sistema dinámico lógico, como una red booleana, es convertido en un sistema lineal discreto. Utilizaremos una generalización del producto de matrices llamado *producto semi-tensorial*, para matrices donde la condición de ser conformables no se satisface. Dedicamos el capítulo 2 a la construcción de esta herramienta matemática.

En el capítulo 4 presentamos las redes difusas aleatorias (RDAs), extendiendo el modelo clásico (RBAs) por un modelo difuso donde los nodos pueden pasar de tomar sólo dos estados a tomar una infinidad. De esta forma, estudiamos la *reducción de complejidad* que los modelos Booleanos son capaces de realizar, justificando su utilización en el diseño, análisis y control de sistemas complejos. Los resultados de nuestras simulaciones se encuentran en el capítulo 5, seguidos por comentarios finales en el capítulo 6.

*It would be altogether incredibly inconvenient to have an unambiguous language, and  
therefore mercifully we have not got one.*  
— BERTRAND RUSSELL, *The Philosophy of Logical Atomism* (1918)

168. *The mathematician is an inventor, not a discoverer.*  
— LUDWIG WITTGENSTEIN, *Remarks on the Foundations of Mathematics* (1938)

*... we might change the idea that space is continuous to the idea that space perhaps is  
a simple lattice and everything is discrete and that time jumps discontinuously.*  
— RICHARD FEYNMAN, *Simulating Physics with Computers* (1981)

# Capítulo 1

## Lógica Difusa

El significado de una palabra nunca está completamente determinado. Es el contexto y la forma en la que un término es empleado lo que produce la ilusión de fijar un único significado. En este trabajo usaremos el término *lógica difusa* para referir a todos los aspectos de representar y manipular conocimiento que emplea valores de verdad intermedios. Este general sentido común del significado del término *lógica difusa* abarca, en particular, conjuntos difusos, relaciones difusas y sistemas formales deductivos que admiten valores de verdad intermedios, así como los diversos métodos basados en ellos. Presentaremos a continuación una revisión de las ideas básicas de la lógica difusa dada como en [Belohlavek y Klir \[2011\]](#), así como una introducción a las *álgebras de Heyting* que son estructuras algebraicas capaces de modelar el comportamiento de los operadores difusos.

### 1.1. Introducción

Como es bien sabido, la lógica clásica está basada en las suposiciones que hay exactamente dos valores de verdad, *falso* y *verdadero*, y que el valor de verdad de cualquier fórmula lógica es definido únicamente por el valor de verdad de sus componentes. Estas suposiciones son usualmente llamadas *bivalencia* y *funcionalidad de la verdad*, respectivamente. Las diversas lógicas multivaluadas abandonan la bivalencia mientras se adhieren a la funcionalidad de la verdad. Esto significa que valores de verdad adicionales son reconocidos en cada lógica multivaluada. Aunque no es obvio cómo interpretar estos valores de verdad adicionales, usualmente son vistos

como valores de verdad intermedios entre *falso* y *verdadero* e interpretados como grados de verdad. Las lógicas multivaluadas difieren unas con otras en los conjuntos de valores de verdad que emplean y en las definiciones que usan para las operaciones lógicas básicas, es decir, los operadores lógicos de negación, conjunción y disyunción.

La lógica clásica está cercamente conectada con la teoría de conjuntos clásica. Cada predicado está asociado de forma única a un conjunto clásico. En otras palabras, para cada objeto dado, la proposición formada por el predicado es verdad para este objeto si y sólo si el objeto es un miembro del conjunto asociado. El conjunto asociado juega el papel de la extensión del predicado. Por otro lado, cada operación lógica de predicados tiene una contraparte única –una operación en los conjuntos clásicos asociados. Por ejemplo, las contrapartes de la negación, conjunción y disyunción de predicados son las operaciones de complemento, intersección y unión en los conjuntos asociados, respectivamente.

Cuando la suposición de bivalencia fue abandonada en las diversas lógicas multivaluadas propuestas, la conexión entre predicados y conjuntos quedó perdida. Los conjuntos clásicos simplemente no fueron capaces de jugar el papel de extensiones de predicados multivaluados, es decir, predicados que se aplican a objetos en un grado intermedio. La conexión fue tarde o temprano renovada cuando Lotfi Zadeh introdujo el concepto de conjunto difuso en su artículo seminal [[Zadeh, 1965](#)].

Una lógica difusa puede ser vista, en parte, como una extensión difusa de una lógica multivaluada que constituye un cálculo lógico correcto y completo. Sin embargo, Zadeh también intentó expandir la noción de lógicas difusas en este sentido (usualmente referido como lógicas difusas en el sentido estrecho) con el propósito de desarrollar razonamiento aproximado que haría posible, en última instancia, emular el razonamiento del sentido común humano en el lenguaje natural. Con este fin, introdujo conjuntos difusos apropiados para representar ciertos tipos de términos lingüísticos empleados en el razonamiento humano. Esta noción extendida de cualquiera de las lógicas difusas en el sentido estrecho es usualmente llamada una lógica difusa en el sentido amplio.

Representar y tratar con clases de objetos que no están precisamente definidos fue entonces la principal motivación para introducir conjuntos difusos. Puesto que tales clases son ubicuas en todas las actividades humanas que involucran al lenguaje natural, los conjuntos difusos

abrieron nuevas y potencialmente útiles formas de ver la cognición humana, el razonamiento, la comunicación, la toma de decisión, y similares. Quizá la más importante de estas fue la nueva forma de ver al conocimiento expresado por enunciados en lenguaje natural. Tal conocimiento asumió un nuevo significado debido a la posibilidad de representarlo y tratarlo en una forma rigurosamente matemática. Su utilidad en la ciencia, ingeniería y otras áreas de los asuntos humanos ha sido cada vez más reconocida.

## 1.2. Objetivos

La mayoría de las proposiciones que la gente usa para comunicar información acerca del mundo no son bivalentes. La verdad de tales proposiciones es una cuestión de grado, más que de ser solamente verdaderas o falsas. Asimismo, la mayoría de los objetos a los que la gente se refiere cuando comunican información no tienen fronteras precisas, claramente definidas.

Como la mayoría de las proposiciones acerca del mundo real empleadas en la comunicación y cognición humana no son bivalentes, la lógica clásica es completamente inadecuada para formalizar razonamiento que involucre tales proposiciones. Asimismo, como la mayoría de las colecciones referidas en la comunicación humana no tienen fronteras precisas, los conjuntos clásicos son sumamente inadecuados para representar tales colecciones. El principal objetivo de la lógica difusa es sobreponerse a las previamente descritas incompetencias de la lógica clásica y los conjuntos clásicos.

La principal idea empleada por la lógica difusa es el permitir una escala parcialmente ordenada de valores de verdad, también llamada grados de verdad, que contiene los valores que representan *falso* y *verdadero*, pero también algunos grados adicionales, intermedios, de verdad. Es decir, el conjunto  $\{0, 1\}$  de valores de verdad de la lógica clásica, donde 0 y 1 representan *falso* y *verdadero*, respectivamente, es remplazado en lógica difusa por una escala parcialmente ordenada de grados de verdad siendo el grado más pequeño 0 y el más grande 1. Un ejemplo importante de tal escala es el intervalo  $[0, 1]$  de números reales.

Un grado de una escala dada (e.g., el número 0.9 de  $[0, 1]$ ) que es asignado a una proposición es interpretado como el grado en el que la proposición es considerada verdadera. De forma similar, las escalas de grados de verdad son usadas en los conjuntos difusos para representar los

grados en los que un objeto dado es elemento de una colección con una frontera no precisa.

### 1.3. Conjuntos

El concepto de conjunto es fundamental prácticamente para todas las ramas de las matemática, tanto pura como aplicada. Intuitivamente, un conjunto es cualquier colección de objetos definidos y distintos que son concebidos como un todo. Los objetos que son incluidos en un conjunto son usualmente llamados sus *elementos*.

Los conjuntos clásicos satisfacen dos requerimientos básicos. Primero, los elementos de cada conjunto son distinguibles unos de otros, y segundo, para cualquier objeto dado puede ser especificado si el objeto es, o no es, un elemento de el conjunto. Los conjuntos difusos que fueron introducidos por Zadeh [Zadeh, 1965], difieren de los conjuntos clásicos rechazando el segundo requerimiento. Contrario a los conjuntos clásicos, los conjuntos difusos no requieren fronteras precisas que separen sus elementos de otros objetos. La pertenencia de cualquier objeto en un conjunto difuso no es una cuestión de afirmación o negación, como es requerido para conjuntos clásicos, sino una cuestión de grado.

En cada aplicación particular de la teoría de conjuntos (clásicos o difusos), el conjunto de todos los objetos que son relevantes para la aplicación es llamado un *conjunto universal* (o un *universo*). Independientemente de que trabajemos con conjuntos clásicos o conjuntos difusos, el conjunto universal relevante es siempre un conjunto clásico.

Los conjuntos difusos están definidos en un conjunto universal dado como funciones que son análogas a las funciones características de los conjuntos clásicos. Cada una de estas funciones asigna a cada objeto en el conjunto universal un grado de verdad. Si los grados de verdad son números reales en el intervalo  $[0, 1]$ , el conjunto difuso definido es llamado un *conjunto difuso común*. Formalmente, un conjunto difuso común  $A$  en un universo  $U$  es una función

$$A : U \rightarrow [0, 1]$$

que asigna a cada elemento  $x$  de  $U$  un número  $A(x)$  en  $[0, 1]$ . Este número es llamado un grado de pertenencia de  $x$  en  $A$  y es visto como el grado en el que  $x$  es considerado un elemento de la colección representada por el conjunto difuso  $A$ . Es decir,  $A(x)$  es interpretado como el grado

de verdad de la proposición “ $x$  es un elemento de  $A$ ”.

Recordamos que una función característica de un conjunto clásico en el conjunto universal  $U$  es cualquier función

$$A : U \rightarrow \{0, 1\}$$

que asigna a cada elemento  $x$  de  $U$  un valor de verdad 0 (si  $x$  no es elemento del conjunto) o 1 (si  $x$  es un elemento del conjunto). Es decir, el concepto de conjunto difuso difiere del concepto de función característica de un conjunto clásico solamente en que el conjunto  $\{0, 1\}$  de valores de verdad es remplazado por una escala general de grados de verdad como  $[0, 1]$ .

Cualquier conjunto difuso especial  $A : U \rightarrow [0, 1]$  para el cual  $A(x) = 0$  o  $A(x) = 1$  para cada  $x$  de  $U$  es llamado un *conjunto difuso crujiante* o simplemente un *conjunto crujiante*. Claramente, todo conjunto clásico puede ser representado por un conjunto difuso crujiante. A saber, un conjunto clásico  $C$  que es un subconjunto de  $U$  es representado de manera única por un conjunto difuso  $A_C$  en  $U$  tal que  $A_C(x) = 0$  si  $x$  no es un elemento de  $C$  y  $A_C(x) = 1$  si  $x$  es un elemento de  $C$ . Inversamente, cada conjunto difuso crujiante en  $U$  representa un conjunto clásico que es un subconjunto de  $U$ . Esta correspondencia nos permite ver a los conjuntos clásicos como conjuntos difusos particulares e identificarlos con conjuntos crujiantes.

Contrario al rol simbólico de los números 0 y 1 en las funciones características de conjuntos clásicos (como falso y verdadero, respectivamente), los grados de pertenencia asignados a objetos por conjuntos difusos comunes tienen una significación numérica. Para cualquier objeto  $x$  en  $U$ , claramente,  $A(x)$  define un grado numérico de pertenencia de  $x$  en el conjunto difuso  $A$ . En conjuntos clásicos, 0 y 1 son considerados como símbolos; en conjuntos difusos, 0 y 1, así como otros grados de verdad, son vistos como números. Una consecuencia importante de este cambio es que los conjuntos difusos (incluyendo a los conjuntos crujiantes) pueden ser manipulados en varias formas numéricas que no tienen contraparte en la teoría clásica de conjuntos.

### 1.3.1. Conceptos Básicos sobre Conjuntos Difusos

Dado un conjunto difuso  $A$  definido en un conjunto universal  $U$ , el conjunto de todos los elementos de  $U$  para el cual  $A(x) > 0$  es llamado un *soporte* de  $A$ , y el conjunto de aquellos para el cual  $A(x) = 1$  es llamado un *núcleo* de  $A$ . Los conjuntos difusos cuyos núcleos no son vacíos son llamados conjuntos difusos *normales*; todos los demás conjuntos difusos son llamados

*subnormales.*

Entre los conceptos más importantes asociados a los conjuntos difusos es el concepto de un *corte de nivel* (también llamado un  $\alpha$ -corte). Dado un conjunto difuso  $A$  definido en un conjunto universal  $U$  y cualquier número  $\alpha$  en el intervalo  $[0, 1]$ , el corte de nivel al nivel  $\alpha$  (o el  $\alpha$ -corte) de  $A$ , denotado por  $A_\alpha$ , es un conjunto clásico de todos los objetos  $x$  de  $U$  para los cuales  $A(x) \geq \alpha$ . Esto puede ser formalmente escrito como

$$A_\alpha = \{x \in U : A(x) \geq \alpha\}, \quad (1-1)$$

donde el símbolo estándar  $x \in U$  significa que un objeto dado  $x$  es un elemento del conjunto universal  $U$  (un conjunto clásico) y los dos puntos denotan la frase *tal que*. Se sigue inmediatamente de esta definición que para cualquier conjunto difuso  $A$  y cualesquiera dos números  $\alpha$  y  $\beta$  en  $[0, 1]$  tales que  $\alpha > \beta$ , la inclusión de conjuntos  $A_\alpha \subseteq A_\beta$  siempre se satisface. De esta manera, todos los cortes de nivel distintos de cualquier conjunto difuso  $A$  forman una familia,

$$\mathbf{A} = \{A_\alpha : \alpha \in [0, 1]\},$$

de conjuntos clásicos anidados. Esto significa que para cualquier secuencia de números  $\alpha_1 > \alpha_2 > \alpha_3 > \dots$  en  $[0, 1]$ ,  $A_{\alpha_1} \subseteq A_{\alpha_2} \subseteq A_{\alpha_3} \subseteq \dots$ . Se ha establecido matemáticamente que esta familia es una representación alternativa del conjunto difuso  $A$ .

Dada la representación por cortes de nivel de un conjunto difuso  $A$ , su representación como función de pertenencia está determinada de forma única para todo  $x \in U$  por la fórmula

$$A(x) = \sup \{\alpha : \alpha \in [0, 1], x \in A_\alpha\}, \quad (1-2)$$

donde “sup” denota el supremo (máximo, si este existe).

La significación de la representación por cortes de nivel de conjuntos difusos es que conecta los conjuntos difusos con los conjuntos clásicos. Considerando que cada conjunto clásico es una colección de objetos que son concebidos como un todo, cada conjunto difuso es una colección (potencialmente infinita) de conjuntos clásicos anidados que son también concebidos como un todo.

Dados dos conjuntos difusos,  $A$  y  $B$ , que están definidos en el mismo conjunto universal  $U$ , decimos que  $A$  es un subconjunto de  $B$  si y sólo si  $A(x) \leq B(x)$  para todo  $x \in U$ . Si  $A$  es un subconjunto de  $B$ , lo escribimos como  $A \subseteq B$ . El conjunto de todos los subconjuntos difusos comunes de  $U$  (a veces llamado un *conjunto potencia difuso*) es denotado por  $[0, 1]^U$ , y el conjunto de todos los subconjuntos crujientes de  $U$  es denotado por  $\{0, 1\}^U$ .

El número de elementos en un conjunto  $A$  definido en un conjunto universal es llamado la *cardinalidad* de  $A$ . Cuando un conjunto difuso definido en un conjunto universal finito  $U$ , su cardinalidad, denotada por  $|A|$ , está definida por la fórmula

$$|A| = \sum_{x \in U} A(x). \quad (1-3)$$

## 1.4. Álgebras de Heyting

La estructura de la lógica difusa es esencialmente la de una álgebra de Heyting [Turunen, 1992]. Esto es conveniente ya que entonces se pueden usar teoremas sobre álgebras de Heyting para deducir consecuencias sobre la lógica proposicional difusa. En esta sección introducimos el concepto de “latiz” y explicamos qué es un álgebra de Heyting. Utilizaremos conjuntos en el sentido de la teoría clásica de conjuntos, o equivalentemente conjuntos crujientes de la teoría de conjuntos difusos, a menos de que se aclare otra cosa.

### 1.4.1. Latices

Comenzamos con la noción de un *conjunto parcialmente ordenado* introducido como en Bell y Slomson [1974]. El lector interesado también puede consultar Sørensen y Uryczyn [2006].

Un conjunto parcialmente ordenado es una pareja ordenada  $\langle X, \leq \rangle$  donde  $X$  es conjunto no vacío, y la relación  $\leq$  es una relación binaria en  $X$  con las siguientes propiedades:

- (a). *Reflexividad*: para todo  $x \in X$ ,  $x \leq x$ ,
- (b). *Antisimetría*: para todos  $x, y \in X$ , si  $x \leq y$  y  $y \leq x$ , entonces  $x = y$ ,
- (c). *Transitividad*: para todos  $x, y, z \in X$ , si  $x \leq y$  y  $y \leq z$ , entonces  $x \leq z$ .

Decimos que un conjunto parcialmente ordenado  $\langle X, \leq \rangle$  es un orden total si, adicionalmente,  $\leq$  tiene la siguiente propiedad:

(d). *Dicotomía*: para todos  $x, y \in X$ , ocurre  $x \leq y$  o  $y \leq x$ .

Si  $x \leq y$  comúnmente usamos lenguaje intuitivo y decimos que “ $y$  es mayor o igual que  $x$ ”. Si  $X$  es un conjunto parcialmente ordenado, una *cadena* en  $X$  es un subconjunto de  $X$  que está totalmente ordenado por la relación de orden en  $X$ .

Si  $A$  es un subconjunto del conjunto parcialmente ordenado  $X$ , una *cota superior* para  $A$  en  $X$  es un elemento de  $X$  mayor que todos los elementos de  $A$ . Así,  $x \in X$  es una cota superior para  $A$  en  $X$  si y sólo si  $a \leq x$  para todo  $a \in A$ . Análogamente  $x \in X$  es una *cota inferior* para  $A$  en  $X$  si para todo  $a \in A$ ,  $x \leq a$ .

Llamamos a  $x \in X$  el *supremo* (o *mínima cota superior*) de  $A$  si  $x$  es una cota superior para  $A$  en  $X$  y es una cota inferior para el conjunto de todas las cotas superiores para  $A$  en  $X$ . Así  $x \in X$  es el supremo de  $A$  si y sólo si para todo  $a \in A$ ,  $a \leq x$  y dada cualquier otra cota superior  $y$ , para  $A$  en  $X$ ,  $x \leq y$ . Análogamente decimos que  $x \in X$  es el *ínfimo* (o *máxima cota inferior*) de  $A$  si es una cota inferior para  $A$  y es mayor que cualquier otra cota inferior para  $A$  en  $X$ .

Denotamos al supremo de  $A$ , si existe, por  $\sup(A)$  y al ínfimo de  $A$ , si existe, por  $\inf(A)$ . Si  $A$  es el conjunto de dos elementos  $\{x, y\}$ , entonces escribimos  $x \vee y$  en lugar de  $\sup(A)$  y  $x \wedge y$  en lugar de  $\inf(A)$ . A  $x \vee y$  se le llama la *conjunción* de  $x$  y  $y$ , y a  $x \wedge y$  se le llama la *disyunción* de  $x$  y  $y$ .

Una *latic* es un conjunto parcialmente ordenado en donde cada subconjunto de dos elementos tiene tanto un supremo como un ínfimo. Así, en una latic  $L$  para todos  $x, y \in L$  tanto  $x \vee y$  como  $x \wedge y$  existen.

**Observación 1** Dados  $x, y, z \in L$  donde  $L$  es una latic, es claro que  $z \geq x \vee y \Leftrightarrow z \geq x$  y  $z \geq y$ , y también  $z \leq x \wedge y \Leftrightarrow z \leq x$  y  $z \leq y$ .

**Proposición 2** En cualquier latic  $L$ , para todos  $x, y, z \in L$  se cumplen las siguientes identidades:

$$L_1. \quad x \vee y = y \vee x \quad x \wedge y = y \wedge x.$$

$$L_2. \quad x \vee (y \vee z) = (x \vee y) \vee z \quad x \wedge (y \wedge z) = (x \wedge y) \wedge z.$$

$$L_3. \quad (x \vee y) \wedge y = y \quad (x \wedge y) \vee y = y.$$

**Demostración.** ( $L_1$ ) Sea  $z \in L$  tal que  $z \geq x \vee y \Leftrightarrow z \geq x, y \Leftrightarrow z \geq y, x \Leftrightarrow z \geq y \vee x$ . Análogamente, sea  $z \in L$  tal que  $z \leq x \wedge y \Leftrightarrow z \leq x, y \Leftrightarrow z \leq y, x \Leftrightarrow z \leq y \wedge x$ .

( $L_2$ ) Sea  $w \in L$  tal que  $w \geq (x \vee y) \vee z \Leftrightarrow w \geq z, x \vee y \Leftrightarrow w \geq x, y, z \Leftrightarrow w \geq x, y \vee z \Leftrightarrow w \geq x \vee (y \vee z)$ . Análogamente, sea  $w \in L$  tal que  $w \leq (x \wedge y) \wedge z \Leftrightarrow w \leq z, x \wedge y \Leftrightarrow w \leq x, y, z \Leftrightarrow w \leq x, y \wedge z \Leftrightarrow w \leq x \wedge (y \wedge z)$ .

( $L_3$ )  $y \geq (x \vee y) \wedge y = x \vee (y \wedge y) = x \vee y \geq y$ . Análogamente,  $y \leq (x \wedge y) \vee y = x \wedge (y \vee y) = x \wedge y \leq y$ . ■

Claramente, si un conjunto parcialmente ordenado tiene una cota superior, entonces esta es única. En particular esto es verdad para una latiz y llamamos a la única cota superior de una latiz, si existe, el *máximo* de la latiz y lo denotamos por 1. Análogamente si una latiz tiene una cota inferior, esta es única. La llamamos el *mínimo* de la latiz y lo denotamos por 0.

El elemento máximo de una latiz debe ser cuidadosamente distinguido de un elemento *maximal* de una latiz. Decimos que un elemento  $x$  de una latiz es *maximal* si no hay algún elemento de la latiz estrictamente mayor (mayor y necesariamente diferente) que  $x$ . Podemos recordar aquí que el *lema de Zorn* afirma que si  $X$  es un conjunto parcialmente ordenado en el que cada cadena tiene una cota superior entonces  $X$  contiene un elemento maximal.

Un conjunto parcialmente ordenado no tiene necesariamente un elemento maximal único. Sin embargo, la situación es diferente para las latices.

**Lema 3** *Si la latiz  $L$  contiene un elemento maximal, entonces este elemento maximal es único.*

**Demostración.** Supongamos que  $x$  y  $y$  son elementos maximales de  $L$ . Entonces  $x \leq x \vee y$  y  $y \leq x \vee y$ , y como  $x$  y  $y$  son maximales,  $x = x \vee y$  y  $y = x \vee y$ . Así,  $x = y$  y el lema se sigue. ■

Decimos que una latiz  $L$  es *complementada* si tiene un elemento maximal 1, un elemento minimal 0, y para cada  $x \in L$  existe un elemento  $y \in L$  que satisface:

$$L_4. \quad x \vee y = 1 \quad y \quad x \wedge y = 0.$$

Dicho elemento  $y$  es llamado el *complemento* de  $x$ . En una latiz complementada cada elemento no tiene necesariamente un complemento único.

Dados  $x, y \in L$ , un elemento  $z$  de una latiz  $L$  es un pseudo-complemento relativo de  $x$  con respecto a  $y$  si  $z$  es el mayor elemento que satisface:

$$L'_4. \quad x \wedge z \leq y.$$

Si el pseudo-complemento relativo de  $x$  con respecto a  $y$  existe, lo denotamos por  $x \rightarrow y$ . Decimos que una latiz es *pseudo-complementada* si tiene un elemento maximal 1, un elemento minimal 0, y para cada  $x, y \in L$ , existe  $z = x \rightarrow y$  en  $L$  tal que satisface la propiedad  $L'_4$ .

La condición extra que imponemos en una latiz complementada para asegurar que el complemento sea único es la distributividad. Decimos que una latiz  $L$  es *distributiva* si para todos  $x, y, z \in L$  las siguientes identidades se satisfacen:

$$L_5. \quad \begin{aligned} (x \vee y) \wedge z &= (x \wedge z) \vee (y \wedge z), \\ (x \wedge y) \vee z &= (x \vee z) \wedge (y \vee z). \end{aligned}$$

**Proposición 4** *En una latiz complementada  $L$ , para todo  $x \in L$ ,  $x \wedge 1 = x$ ,  $x \vee 1 = 1$ ,  $x \wedge 0 = 0$ , y  $x \vee 0 = x$ .*

**Demostración.** Sea  $y \in L$  tal que  $x \vee y = 1$  y  $x \wedge y = 0$ . Por  $L_3$  se tiene que  $(y \vee x) \wedge x = x$ , entonces  $x = x \wedge (x \vee y) = x \wedge 1$ . Así,  $x \vee 1 = (x \vee 1) \wedge 1 = 1$ . Análogamente  $(y \wedge x) \vee x = x$ , entonces  $x = x \vee (x \wedge y) = x \vee 0$ . Por lo que  $x \wedge 0 = (x \wedge 0) \vee 0 = 0$  ■

**Lema 5** *En una latiz distributiva complementada cada elemento tiene un único complemento.*

**Demostración.** Supongamos que  $L$  es una latiz distributiva complementada y que tanto  $y$  como  $z$  son complementos del elemento  $x$  de  $L$ . Así,  $x \vee y = 1$  y  $x \wedge z = 0$ . Entonces, usando la proposición 4,

$$\begin{aligned} y &= y \vee 0 = y \vee (x \wedge z) \\ &= (y \vee x) \wedge (y \vee z), \text{ ya que } L \text{ es distributiva,} \\ &= 1 \wedge (y \vee z) \\ &= y \vee z. \end{aligned}$$

Análogamente  $z = y \vee z$ , y por lo tanto  $y = z$ . ■

En una latiz distributiva complementada el único complemento de un elemento  $x$  es denotado por  $\bar{x}$ . Por otro lado, en una latiz distributiva pseudo-complementada se tiene como caso particular que  $\bar{x} = x \rightarrow 0$ .

Estamos ahora listos para la definición principal de esta sección. Una *álgebra de Heyting* es una latiz distributiva<sup>1</sup> pseudo-complementada con al menos dos elementos. Así, insistimos que  $0 \neq 1$  en cualquier álgebra de Heyting.

Introducimos la noción de álgebra de Heyting construyendo gradualmente su estructura iniciando con un conjunto parcialmente ordenado. De manera alternativa también podríamos haber dado la siguiente definición:

Una álgebra de Heyting  $\mathcal{H}$  es una estructura algebraica

$$\mathcal{H} = \langle H, \vee, \wedge, \rightarrow, 0, 1 \rangle,$$

que satisface  $L_1, L_2, L_3, L'_4$  y  $L_5$ .

Si una estructura algebraica  $\mathcal{B} = \langle B, \vee, \wedge, \bar{\cdot}, 0, 1 \rangle$  satisface  $L_1 - L_5$  entonces decimos que es una *álgebra booleana*. Así, toda álgebra booleana es una álgebra de Heyting ya que con  $x \rightarrow y$  definido como  $\bar{x} \vee y$ , para cualesquiera dos de sus elementos, es claro que  $L'_4 \Rightarrow L_4$ .

Dada una álgebra de Heyting  $\mathcal{H}$  observamos que el orden parcial subyacente se encuentra implícito y definido de manera usual por la cláusula:  $x \leq y$  si y sólo si  $x \vee y = y$ , para todos  $x, y \in H$ .

### 1.4.2. Ejemplos

1. Sea  $\mathbf{2}$  el conjunto cuyos únicos dos elementos son los cardinales  $0, 1$  con el orden parcial  $\leq$  definido por  $0 \leq 0, 0 \leq 1, \text{ y } 1 \leq 1$ . Claramente  $\mathbf{2}$  es una álgebra booleana.
2. Sea  $X$  cualquier conjunto. Si  $\subseteq$  es la relación de pertenencia entre conjuntos, entonces  $\langle \{0, 1\}^X, \subseteq \rangle$  es un conjunto parcialmente ordenado y  $\{0, 1\}^X$  una álgebra booleana. En este caso el elemento máximo es el conjunto  $X$ , y el elemento mínimo es el conjunto vacío

---

<sup>1</sup>Se puede demostrar que si  $x \rightarrow y$  existe para cualesquiera  $x, y$  elementos de una latiz  $L$ , entonces  $L$  es distributiva. Esta se omite por cuestiones de practicidad.

$\emptyset$ . El complemento corresponde a la noción de complemento de la teoría de conjuntos clásica.

3. Denotamos por  $\mathcal{O}(\mathcal{T})$  a la familia de todos los conjuntos abiertos de un espacio topológico  $\mathcal{T}$ . Definimos  $A \rightarrow B = \text{Int}(\overline{A} \cup B)$ , para cualesquiera conjuntos abiertos  $A$  y  $B$ , y a  $\cup, \cap$  como en la teoría de conjuntos clásica. Entonces  $\mathcal{H} = \langle \mathcal{O}(\mathcal{T}), \cup, \cap, \rightarrow, \emptyset, \mathcal{T} \rangle$  es una álgebra de Heyting, pero no una álgebra booleana ya que el complemento de un conjunto abierto usualmente no es abierto. Lo mejor que podemos hacer es tomar el mayor (en el sentido de pertenencia) conjunto abierto contenido en el complemento –el interior (Int). Este es exactamente el pseudo-complemento de una álgebra de Heyting.
4. Sean  $T, S : [0, 1]^2 \rightarrow [0, 1]$  funciones que cumplen las propiedades de identidad [ $T(x, 1) = x; S(x, 0) = x$ ], conmutatividad [ $T(x, y) = T(y, x); S(x, y) = S(y, x)$ ], asociatividad [ $T(x, T(y, z)) = T(T(x, y), z); S(x, S(y, z)) = S(S(x, y), z)$ ] y monotonía [ $x \leq u, y \leq v \Rightarrow T(x, y) \leq T(u, v), S(x, y) \leq S(u, v)$ ]. Sea  $N : [0, 1] \rightarrow [0, 1]$  una función continua tal que es una involución [ $N(N(x)) = x$ ], es estrictamente decreciente [ $x < y \Rightarrow N(x) > N(y)$ ], y satisface las condiciones de frontera  $N(0) = 1$  y  $N(1) = 0$ . Entonces, decimos que  $T$  es una *norma triangular* ( $t$ -norma),  $S$  es una *conorma triangular* ( $t$ -conorma) y  $N$  una *negación fuerte*. Denotando por  $I$  al intervalo  $[0, 1]$  es claro que  $\mathcal{H} = \langle I, T, S, N, 0, 1 \rangle$  es una álgebra de Heyting.

**Nota 6** En su artículo seminal [Zadeh, 1965], Zadeh menciona: “... esencialmente, los conjuntos difusos en (un universo de discurso)  $X$  constituyen una latiz distributiva con 0 y 1.”. Se define como una tripleta de De Morgan a una tripleta  $(S, T, N)$  donde  $S$  es una  $t$ -conorma,  $T$  es una  $t$ -norma,  $N$  es una negación fuerte, y además se satisface que  $S(x, y) = N(T(N(x), N(y)))$  para todos  $x, y \in I$ . La tripleta de De Morgan (mín, máx,  $1 - x$ ), que es usualmente llamada  $t$ -norma y  $t$ -conorma de Gödel, constituye a lo que nos referiremos más adelante como los operadores de Zadeh.

Todas las operaciones de conjuntos difusos están inducidas por operaciones asociadas en el intervalo unitario  $[0, 1]$  de la siguiente manera. Una operación  $n$ -aria en  $[0, 1]$ , digamos la operación  $s$ , se extiende a una operación  $n$ -aria en conjuntos difusos (que por simplicidad,

también denotamos por  $s$ ) asignando a los conjuntos difusos  $A_1, A_2, \dots, A_n$  un conjunto difuso  $s(A_1, A_2, \dots, A_n)$  cuya función de pertenencia está definida para todo  $x \in U$  por

$$[s(A_1, A_2, \dots, A_n)](x) = s(A_1(x), A_2(x), \dots, A_n(x)). \quad (1-4)$$

Decimos que una operación en conjuntos difusos definida de esta forma está basada en la operación asociada en  $[0, 1]$ .

De esta manera, por el ejemplo 4 y la nota 6, es claro que dado cualquier universo  $U$ , la colección de todos los conjuntos difusos comunes en  $U$  con los operadores de Zadeh es una álgebra de Heyting, i.e.

$$\mathcal{F} = \langle [0, 1]^U, \text{mín}, \text{máx}, 1 - x, 0, 1 \rangle$$

es una estructura algebraica que satisface los requisitos (axiomas  $L_1, L_2, L_3, L'_4$  y  $L_5$ ) para ser una álgebra de Heyting.

## Capítulo 2

# Producto Semi-tensorial de Matrices

En este capítulo construiremos una herramienta matemática que definiremos como el *producto semi-tensorial* de matrices, con el cual una función lógica puede ser convertida en un mapeo multilinear llamado la *expresión matricial de relaciones lógicas*. Bajo esta construcción, la *dinámica* de una red booleana puede ser expresada como un sistema dinámico lineal discreto. Mediante esta representación lineal, ciertas propiedades importantes de la topología de la red, como *puntos fijos*, *ciclos*, *tiempos de transcurso*, y *cuencas de atracción*, pueden ser fácilmente descubiertas a través de un conjunto de fórmulas. Seguiremos el enfoque desarrollado por [Cheng et al. \[2011\]](#) y [Zhao et al. \[2014\]](#).

### 2.1. Funciones Lógicas

En la lógica clásica se asume que una proposición es o “verdadera”, o “falsa”. Como la estructura de dicha lógica es esencialmente la de una álgebra booleana [[Bell y Slomson, 1974](#)], es usual denotar por 1 al valor *verdadero* y por 0 al valor *falso*. Los conectivos  $\vee$ ,  $\wedge$  y  $\neg$  son también llamados *operadores lógicos*, y es usual nombrarlos **o**, **y** y **no** respectivamente.

**Definición 7** Consideremos el conjunto  $\mathcal{D} = \{1, 0\}$ .

1. Una variable lógica es una variable que puede tomar valores en  $\mathcal{D}$ .
2. Las variables lógicas  $x_1, \dots, x_n$  son independientes si para cualquier valor fijo  $x_j$ ,  $j \neq i$ , la variable lógica  $x_i$  puede aún valer tanto 1 como 0.

3. Una función lógica de variables lógicas  $x_1, \dots, x_n$  es una expresión lógica que involucra a  $x_1, \dots, x_n$ , y a algunos posibles enunciados (llamados constantes), mediante conectivos.

Por lo tanto, una función lógica es un mapeo  $f : \mathcal{D}^n \rightarrow \mathcal{D}$ . Por la forma en que están definidas las funciones lógicas usualmente también son llamadas *operadores  $n$ -arios*. Así, observamos que la *negación*  $\bar{\phantom{x}}$  es un operador 1-ario (unario), y la *disyunción*  $\vee$  y *conjunción*  $\wedge$  son operadores 2-arios (binarios).

Es claro que para un operador  $n$ -ario hay  $n$  variables lógicas y cada una de estas puede tomar dos valores posibles, por lo tanto un operador  $n$ -ario es un mapeo cuyo *dominio* es un conjunto de  $2^n$  elementos diferentes y cuyo *codominio* es un conjunto de dos elementos. Por tanto existen  $2^{2^n}$  operadores  $n$ -arios diferentes.

Para el caso en el que existan más de dos valores de verdad definimos los conjuntos:

$$\mathcal{D}_m = \left\{ 1, \frac{m-2}{m-1}, \frac{m-3}{m-1}, \dots, 0 \right\}$$

para lógicas multivaluadas y

$$\mathcal{D}_f = \{r : 0 \leq r \leq 1\}$$

para la lógica difusa. De esta forma,  $\mathcal{D}_2 = \mathcal{D}$  es el conjunto definido previamente para la lógica booleana.

Un sistema lógico es llamado una *lógica  $m$ -valuada (multivaluada)* si sus variables lógicas toman cualquier valor de  $\mathcal{D}_m$ . Por otro lado, un sistema lógico es llamado una *lógica difusa* si sus variables lógicas toman cualquier valor de  $\mathcal{D}_f$ . Así, un operador lógico  $\sigma : \mathcal{D}_m^s \rightarrow \mathcal{D}_m$  es un *operador  $s$ -ario lógico  $m$ -valuado*, y un operador lógico  $\sigma : \mathcal{D}_f^s \rightarrow \mathcal{D}_f$  es un *operador  $s$ -ario lógico difuso*.

## 2.2. Información Multi-dimensional

A grandes rasgos, el álgebra lineal principalmente se interesa por dos tipos de objetos: vectores y matrices. Un vector  $n$ -dimensional es expresado como  $X = (x_1, x_2, \dots, x_n)$ . Sus elementos son clasificados por un índice  $i$ , donde  $x_i$  es el  $i$ -ésimo elemento de  $X$ . Para una

matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

los elementos son clasificados por dos índices,  $i$  y  $j$ , donde  $a_{i,j}$  es el elemento de  $A$  localizado en el  $i$ -ésimo renglón y la  $j$ -ésima columna. De esta manera, es fácil relacionar la dimensión de un conjunto de datos con el número de índices.

Un conjunto de datos, clasificados por  $k$  índices, es llamado un conjunto de *información  $k$ -dimensional*. Precisamente,

$$X = \{x_{i_1, i_2, \dots, i_k} : 1 \leq i_j \leq n_j, j = 1, 2, \dots, k\} \quad (2-1)$$

es un conjunto de infotmación  $k$ -dimensional. El *número cardinal* de  $X$  es su cardinalidad o número de elementos  $|X| = n_1 n_2 \cdots n_k$ .

Ahora, consideremos en un espacio vectorial  $n$ -dimensional  $V$  un mapeo lineal  $\mu : V \rightarrow \mathbb{R}$ . Si escogemos una base  $\{e_1, e_2, \dots, e_n\}$  para  $V$ , las *constantes estructurales*  $\{c_1, c_2, \dots, c_n\}$  de  $\mu$  están dadas de la siguiente forma:

$$\mu(e_i) = c_i, \quad i = 1, \dots, n. \quad (2-2)$$

Enotnces,  $\mu$  puede ser expresado como  $\mu = c_1 e_1^* + c_2 e_2^* + \cdots + c_n e_n^*$ , donde para cada  $i$  se tiene que  $e_i^* : V \rightarrow \mathbb{R}$  satisface:

$$e_i^*(e_j) = \delta_{i,j} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \quad (2-3)$$

Se puede ver fácilmente que el conjunto de mapeos lineales en  $V$  forman un espacio vectorial, llamado el *espacio dual* de  $V$  y denotado por  $V^*$ .

Cualquier vector  $X \in V$  puede ser expresado como  $X = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$ . De igual forma, cualquier mapeo lineal  $\mu \in V^*$  puede ser expresado como  $\mu = \mu_1 e_1^* + \mu_2 e_2^* + \cdots + \mu_n e_n^*$ . Cuando la base y la base dual están fijas,  $X$  puede ser expresado como un *vector columna* y  $\mu$

puede ser expresado como un *vector fila*, i.e.,

$$X = (a_1, a_2, \dots, a_n)^T, \quad \mu = (c_1, c_2, \dots, c_n).$$

Usando esta forma vectorial, la *acción* de  $\mu$  sobre  $X$  puede ser expresada como su matriz producto:

$$\mu(X) = \mu X = \sum_{i=1}^n a_i c_i, \quad \mu \in V^*, X \in V. \quad (2-4)$$

**Nota 8** Aunque las constantes estructurales  $\{c_i : i = 1, 2, \dots, n\}$  dependen de la elección de la base, estas determinan de forma única a  $\mu$ . Así, en general, para definir un mapeo multilinear es suficiente dar sus constantes estructurales.

Sea  $\tau : \underbrace{V^* \times \dots \times V^*}_t \times \underbrace{V \times \dots \times V}_s \rightarrow \mathbb{R}$  un mapeo multilinear en una  $(s + t)$ -variedad. Entonces, decimos que  $\tau$  es un *tensor* sobre  $V$  con *grado covariante*  $s$  y *grado contravariante*  $t$ . Denotamos por  $\mathcal{T}_t^s$  al conjunto de tensores sobre  $V$  con grado covariante  $s$  y grado contravariante  $t$

Definiendo  $c_{j_1, j_2, \dots, j_t}^{i_1, i_2, \dots, i_s} = \tau(e_{i_1}, e_{i_2}, \dots, e_{i_s}, e_{j_1}^*, e_{j_2}^*, \dots, e_{j_t}^*)$ , tenemos que

$$\{c_{j_1, j_2, \dots, j_t}^{i_1, i_2, \dots, i_s} : 1 \leq i_1, \dots, i_s, j_1, \dots, j_t \leq n\}$$

es el conjunto de constantes estructurales de  $\tau$ . Las constantes estructurales de  $\tau \in \mathcal{T}_t^s$  forman un conjunto de información  $(s + t)$ -dimensional.

## 2.3. Índices Ordenados

Ahora consideraremos cómo acomodar información multidimensional de forma uno- o dos-dimensional, es decir, como un vector o como una matriz.

**Definición 9** Supongamos que tenemos un conjunto de datos  $S$  con  $\prod_{i=1}^k n_i$  elementos clasificados como en (2-1) por  $k$  índices. Además, supongamos que los elementos de  $S$  están acomodados en un vector fila (o columna). Decimos que la información está clasificada por los índices  $i_1, \dots, i_k$  de acuerdo a un multi-índice ordenado denotado por

$$Id(i_1, \dots, i_k; n_1, \dots, n_k),$$

si los elementos están clasificados por  $i_1, \dots, i_k$  y ordenados de la siguiente manera: Dejamos correr  $i_t$ , con  $t = 1, \dots, k$ , desde 1 hasta  $n_t$  en orden tal que primero  $t = k$ , después  $t = k - 1$ , y siguiendo así hasta que  $t = 1$ .

De acuerdo a la definición anterior,  $x_{\alpha_1, \dots, \alpha_k}$  antecede a  $x_{\beta_1, \dots, \beta_k}$  si y sólo si existe  $1 \leq j \leq k$  tal que:

$$\alpha_i = \beta_i, \quad i = 1, \dots, j - 1, \quad \alpha_j < \beta_j. \quad (2-5)$$

Si los números  $n_1, \dots, n_k$  de  $i_1, \dots, i_k$  son iguales, podemos usar

$$Id(i_1, \dots, i_k; n) = Id(i_1, \dots, i_k; n, \dots, n),$$

y si todo  $n_i$  es obviamente conocido, la expresión  $Id$  puede ser simplificada como

$$Id(i_1, \dots, i_k) = Id(i_1, \dots, i_k; n_1, \dots, n_k).$$

Dados  $a, b$  y  $c$  números enteros, con  $b > 0$ , denotamos por  $a \% b$  al residuo de dividir  $a$  entre  $b$ , que claramente siempre es no negativo. También, denotamos por  $\lfloor c \rfloor$  al mayor entero que es menor o igual que  $c$ . Es decir,  $a/b = p + a \% b$  para algún  $p \in \mathbb{Z}$ , y  $\lfloor c \rfloor = \max\{q \in \mathbb{Z} : q \leq c\}$ .

**Proposición 10** *Sea  $S$  un conjunto de datos con  $\prod_{i=1}^k n_i$  elementos. Los datos están clasificados por un índice individual como  $\{x_i\}$ , y por un índice  $k$ -variado dado por el multi-índice ordenado  $Id(\lambda_1, \dots, \lambda_k; n_1, \dots, n_k)$ , de forma que podemos convertir al índice de*

$$S = \{s_p : p = 1, \dots, n\} = \{s_{\lambda_1, \dots, \lambda_k} : 1 \leq \lambda_i \leq n_i; i = 1, \dots, k\},$$

1. *De índice individual a multi-índice. Definiendo  $p_k = p - 1$ , el índice individual  $p$  puede ser convertido en el orden del multi-índice ordenado  $Id(i_1, \dots, i_k) = Id(i_1, \dots, i_k; n_1, \dots, n_k)$  como  $(\lambda_1, \dots, \lambda_k)$ , donde  $\lambda_i$  puede ser calculado recursivamente como*

$$\begin{cases} \lambda_k = p_k \% n_k + 1, \\ p_j = \left\lfloor \frac{p_{j+1}}{n_{j+1}} \right\rfloor, \quad \lambda_j = p_j \% n_j + 1, \quad j = k - 1, \dots, 1. \end{cases} \quad (2-6)$$

2. *De multi-índice a índice individual. Del multi-índice  $(\lambda_1, \dots, \lambda_k)$  en el orden de  $Id(i_1, \dots, i_k) =$*

$Id(i_1, \dots, i_k; n_1, \dots, n_k)$  de regreso al índice individual, tenemos que

$$p = \sum_{j=1}^{k-1} (\lambda_j - 1) n_{j+1} n_{j+2} \cdots n_k + \lambda_k. \quad (2-7)$$

Así, un conjunto de información multidimensional, clasificado por un multi-índice, puede ser convertido en un conjunto de información 1-dimensional, clasificado por un índice individual. Una matriz, como conjunto de información 2-dimensional, puede ciertamente ser convertida en un conjunto de información 1-dimensional. Consideremos la matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

La forma *apilada por fila* de A, denotada por  $V_f(A)$ , es un  $nm$ -vector acomodado fila por fila, i.e.,  $V_f(A) = (a_{11}, a_{12}, \dots, a_{1n}, \dots, a_{m1}, a_{m2}, \dots, a_{mn})^T$ . La forma *apilada por columna* de A, denotada por  $V_c(A)$ , es el siguiente  $nm$ -vector:  $V_c(A) = (a_{11}, a_{21}, \dots, a_{m1}, \dots, a_{1n}, a_{2n}, \dots, a_{mn})^T$ . De esta forma, es claro que:  $V_c(A) = V_f(A^T)$  y  $V_f(A) = V_c(A^T)$ .

## 2.4. Producto Semi-tensorial

El producto convencional de matrices puede por lo general tratar solo con información uno- o dos-dimensional. Además, en el caso dos-dimensional, es necesario que la información esté acomodada en matrices conformables<sup>1</sup>. Para extender el producto a conjuntos de información  $n$ -dimensional, con  $n$  cualquier número finito, definimos el concepto de *producto semi-tensorial*.

### Definición 11

1. Sean  $X = (x_1, \dots, x_s)$  un vector fila y  $Y = (y_1, \dots, y_t)^T$  un vector columna.

Caso 1: Si  $t$  es un factor de  $s$ , es decir,  $s = t \times n$ , entonces el vector columna

---

<sup>1</sup>Decimos que las matrices  $A_{p \times q}$ ,  $B_{r \times s}$  son conformables si se satisface que  $q = r$ .

$n$ -dimensional definido como

$$X \times Y = \sum_{k=1}^t X^k y_k \in \mathbb{R}^n \quad (2-8)$$

es llamado el producto interior semi-tensorial izquierdo de  $X$  y  $Y$ , donde

$$X = (X^1, \dots, X^t), \quad X^i \in \mathbb{R}^n, \quad i = 1, \dots, t.$$

Caso 2: Si  $s$  es un factor de  $t$ , es decir,  $t = s \times n$ , entonces el vector columna  $n$ -dimensional definido como

$$X \times Y = \sum_{k=1}^t x^k Y_k \in \mathbb{R}^n \quad (2-9)$$

es llamado el producto interior semi-tensorial izquierdo de  $X$  y  $Y$ , donde

$$Y = ((Y^1)^T, \dots, (Y^t)^T)^T, \quad Y^i \in \mathbb{R}^n, \quad i = 1, \dots, t.$$

2. Sean  $M \in \mathcal{M}_{m \times n}$  y  $N \in \mathcal{M}_{p \times q}$ . Si  $n$  es un factor de  $p$  o  $p$  es un factor de  $n$ , entonces  $C = M \times N$  es llamado el producto semi-tensorial izquierdo de  $M$  y  $N$ , donde  $C$  consiste en bloques de  $m \times q$  como  $C = (C^{ij})$ , y

$$C^{ij} = M^i \times N_j, \quad i = 1, \dots, m, \quad j = 1, \dots, q,$$

donde  $M^i$  es la  $i$ -ésima fila de  $M$  y  $N_j$  es la  $j$ -ésima columna de  $N$ , i.e.  $M^i = \text{Fil}_i(M)$  y  $N_j = \text{Col}_j(N)$ .

En el primer punto de la definición 11, si  $t = s$ , el producto interior semi-tensorial izquierdo se convierte en el producto interior convencional. Así, en el segundo punto de la definición 11, si  $n = p$ , el producto interior semi-tensorial izquierdo se convierte en el producto convencional de matrices. De esta manera, el producto interior semi-tensorial izquierdo es una generalización del producto convencional de matrices.

En esta tesis, el producto estándar para matrices será el producto interior semi-tensorial izquierdo, por tanto lo llamaremos simplemente “producto semi-tensorial” (o todavía más simple “producto”). Cuando nos refiramos al producto tensorial (o *producto de Kronecker*), lo haremos explícito denotándolo, como es usual, por el símbolo  $\otimes$ .

Así, dadas  $A \in \mathcal{M}_{m \times n}$  y  $B \in \mathcal{M}_{p \times q}$ ,

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}.$$

Decimos que  $A$  y  $B$  son *conformables* si  $n = p$ , y diremos que son *multi-conformables* si  $n = tp$ , o  $p = tn$  para algún  $t \in \mathbb{N}$ . Cuando se tiene que  $n = tp$ , lo denotamos por  $A \succ_t B$ , y cuando  $p = tn$ , lo denotamos por  $A \prec_t B$ . De esta manera, observamos que si definimos a  $s = mcm\{n, p\}$ , el *mínimo común múltiplo* de  $n$  y  $p$ , se tiene que [Meng y Feng, 2014]:

$$A \times B = (A \otimes I_{s/n})(B \otimes I_{s/p}). \quad (2-10)$$

**Definición 12** Una matriz de intercambio  $W_{[m,n]}$  es una matriz de  $mn \times mn$ , definida de la siguiente manera. Sus filas y columnas están clasificadas por un índice doble  $(i, j)$ , las columnas están acomodadas por el orden del multi-índice  $Id(i, j; m, n)$ , y las filas están acomodadas por el orden del multi-índice  $Id(j, i; n, m)$ . El elemento en la posición  $[(I, J), (i, j)]$  es entonces

$$w_{(I,J),(i,j)} = \delta_{i,j}^{I,J} = \begin{cases} 1, & I = i \text{ y } J = j, \\ 0, & \text{en otro caso.} \end{cases} \quad (2-11)$$

## 2.5. Representación Matricial de Funciones Lógicas

Para el caso Booleano donde una variable lógica puede tomar valores de  $\mathcal{D} = \{0, 1\}$ , identificamos a los elementos de  $\mathcal{D}$  con los vectores

$$1 \sim \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad 0 \sim \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (2-12)$$

Denotamos por  $\delta_k^i$  a la  $i$ -ésima columna de la matriz identidad  $I_k$ , y así definimos el conjunto  $\Delta_k = \{\delta_k^i : i = 1, 2, \dots, k\}$ . Para el caso en que  $k = 2$  denotamos a  $\Delta_2$  por  $\Delta$ , por lo tanto

$$\Delta = \{\delta_2^1, \delta_2^2\} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\},$$

y así, un operador lógico  $r$ -ario es un mapeo  $\sigma : \Delta^r \rightarrow \Delta$ .

### 2.5.1. Matrices Lógicas

Una matriz  $L \in \mathcal{M}_{n \times m}$  es llamada una matriz lógica si  $\text{Col}(L) \subset \Delta_n$ . El conjunto de matrices lógicas de  $n \times m$  es denotado por  $\mathcal{L}_{n \times m}$ . Es claro que si  $L \in \mathcal{L}_{n \times m}$ , entonces es de la forma  $L = [\delta_n^{i_1} \delta_n^{i_2} \dots \delta_n^{i_m}]$ . Por practicidad esto lo escribiremos de manera más compacta como  $L = \delta_n[i_1 \ i_2 \ \dots \ i_m]$ .

**Definición 13** Una matriz  $M_\sigma$  de  $2 \times 2^r$  es llamada la matriz estructural del operador lógico  $r$ -ario  $\sigma$  si se tiene que

$$\sigma(p_1, \dots, p_r) = M_\sigma \times p_1 \times \dots \times p_r = M_\sigma \times_{i=1}^r p_i. \quad (2-13)$$

Así, las matrices estructurales de los operadores lógicos conjunción, disyunción y negación ( $\vee, \wedge, \bar{\phantom{x}}$ ), denotadas por  $M_c, M_d, M_n$  están dadas de la siguiente forma:

$$M_c = \delta_2[1 \ 2 \ 2 \ 2],$$

$$M_d = \delta_2[1 \ 1 \ 1 \ 2],$$

$$M_n = \delta_2[2 \ 1].$$

Entonces, dadas dos variables lógicas en su forma vectorial  $p, q \in \Delta$ , tenemos que

$$p \vee q = M_c \ pq,$$

$$p \wedge q = M_d \ pq,$$

$$\bar{p} = M_n \ p.$$

**Nota 14** Usando las matrices estructurales correspondientes, es evidente que una función lógica  $f(p_1, p_2, \dots, p_r)$  con variables lógicas  $p_1, p_2, \dots, p_r$  puede ser expresada como

$$f(p_1, p_2, \dots, p_r) = \times_i \xi_i, \quad (2-14)$$

donde  $\xi_i \in \{M_c, M_d, M_n, p_1, p_2, \dots, p_r\}$ .

**Proposición 15** Una matriz  $M_r \in \mathcal{L}_{4 \times 2}$  es llamada matriz de reducción si  $M_r = \delta_4[1 \ 4]$ .

Si  $p \in \Delta$ , entonces  $p^2 = M_r p$ .

**Demostración.** Sea  $p = [t, 1 - t]^T$ . Entonces,  $p^2 = [t^2, t(1 - t), (1 - t)t, (1 - t)^2]^T$ . Como  $t \in \{0, 1\}$  es claro que  $t^2 = t$ ,  $(1 - t)^2 = 1 - t$ , y  $t(1 - t) = 0$ . Entonces,  $p^2 = [t, 0, 0, 1 - t]^T = M_r p$ .

■

**Teorema 16** Dada una función lógica  $f(p_1, p_2, \dots, p_r)$  con variables lógicas  $p_1, p_2, \dots, p_r$ , existe una única matriz  $M_f$  de  $2 \times 2^r$ , llamada matriz estructural de  $f$ , tal que

$$f(p_1, p_2, \dots, p_r) = M_f p_1 p_2 \cdots p_r. \quad (2-15)$$

**Demostración.** Primero probaremos la existencia de  $M_f$ . Siguiendo la nota 14, sólo tenemos que probar que  $\times_i \xi_i$  puede ser expresado como el lado derecho de (2-15). Usando el hecho de que

$$pM = (I_2 \otimes M)p \quad (2-16)$$

podemos mover todos los factores de las matrices estructurales ( $M_j$  o bien  $I_2 \otimes M_j$ ) hacia el frente, y mover todas las variables ( $p_i$ ) hacia atrás del producto  $\times_i \xi_i = \times_j N_j \times_k p_{i_k}$ , donde  $N_j \in \{I_{2^s} \otimes M_c, I_{2^s} \otimes M_d, I_{2^s} \otimes M_n : s = 0, 1, 2, \dots\}$ ,  $i_k \in \{1, 2, \dots, r\}$ .

Usando una matriz de intercambio podemos cambiar el orden de dos variables lógicas:  $W_{[2]} p_i p_j = p_j p_i$ . Usando (2-16) es fácil obtener la siguiente forma:  $\times_k p_{i_k} = M p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$ .

Usando una matriz de reducción, el exponente de las  $p_i$ 's puede ser reducido hasta 1. Nuevamente usando (2-16), las matrices de coeficientes generadas al reducir los ordenes, pueden ser movidas hacia el frente. De esta manera habremos construido la matriz estructural  $M_f$ .

Para probar la unicidad, supongamos que hay dos matrices estructurales  $M_f \neq M'_f$ . Estas tendrían que diferir en al menos una columna, digamos la  $i$ -ésima columna,  $c_i \neq c'_i$ . Escogemos  $p_1, \dots, p_r$  tales que  $\times_{i=1}^r p_i = \delta_{2^r}^i$ . Entonces se tiene que  $f(p_1, \dots, p_r) = M_f \times_{i=1}^r p_i = c_i$ , al mismo tiempo que  $f(p_1, \dots, p_r) = M'_f \times_{i=1}^r p_i = c'_i$  lo cual es una contradicción, y la unicidad se sigue. ■

### 2.5.2. Ecuaciones Lógicas

Una variable lógica  $p$  es llamada el *argumento lógico* o *incógnita* si puede tomar un valor  $p \in \mathcal{D} = \{1, 0\}$  para satisfacer ciertos requerimientos lógicos. Una variable lógica  $c$  es llamada *constante lógica* si es un valor fijo  $c \in \mathcal{D}$ .

**Definición 17** *Un sistema estándar de ecuaciones lógicas es un sistema de ecuaciones expresado como*

$$\begin{cases} f_1(p_1, p_2, \dots, p_n) = c_1, \\ f_2(p_1, p_2, \dots, p_n) = c_2, \\ \vdots \\ f_m(p_1, p_2, \dots, p_n) = c_m, \end{cases} \quad (2-17)$$

donde  $f_i$  con  $i = 1, \dots, m$  son funciones lógicas,  $p_i$  con  $i = 1, \dots, n$  son argumentos lógicos, y  $c_i$  con  $i = 1, \dots, m$  son constantes lógicas.

Un conjunto de constantes lógicas  $d_1, d_2, \dots, d_n$  tales que  $p_i = d_i$  con  $i = 1, \dots, n$ , son llamadas una *solución* de (2-17) si satisfacen a (2-17). La idea para resolver este sistema es, básicamente, convertirlo en una ecuación lineal equivalente y después resolver esta ecuación algebraica, obteniendo así la solución o soluciones del sistema de ecuaciones lógicas.

**Lema 18** *Sean  $p_i$  con  $i = 1, \dots, n$  variables lógicas en forma vectorial, i.e.,  $p_i \in \Delta$ . Definimos  $x = \times_{i=1}^n p_i$ . Entonces,  $p_i$  con  $i = 1, \dots, n$  están determinadas de forma única por  $x$ .*

**Demostración.** Para probar esto, construiremos una fórmula para calcular  $p_i$ . Como  $p_i \in \Delta$ , se sigue que  $x \in \Delta_{2^n}$ . Podemos suponer sin pérdida de generalidad que  $x = \delta_{2^n}^i$ , y dividirlo en dos segmentos del mismo tamaño, de forma que  $x = [x_1^T, x_2^T]^T$ , donde  $0 \neq x_1 \in \Delta_{2^{n-1}}$  y

$x_2 = 0$ , o bien  $x_1 = 0$  y  $0 \neq x_2 \in \Delta_{2^{n-1}}$ . De acuerdo con la definición de producto semi-tensorial (2-8 , 2-9), si  $x_2 = 0$ , entonces  $p_1 = 1$ , y si  $x_2 = 1$ , entonces  $p_1 = 0$ . Ahora, podemos dividir a un segmento diferente de cero,  $x_1 \neq 0$ , en dos partes iguales como  $x_1 = [x_{11}^T, x_{12}^T]^T$ . Entonces, aplicamos el mismo juicio a  $p_2$ , y así sucesivamente, calculamos  $p_i$  para  $i = 1, \dots, n$ . ■

Basándonos en el argumento de la prueba para el lema anterior, dado  $\times_{j=1}^n p_j = \delta_{2^n}^i$ , con  $p_j \in \Delta$ , podemos dar un algoritmo para calcular inductivamente la forma escalar de  $\{p_j\}$  a partir de  $i$ :

Paso 1. Definir  $q_0 = 2^n - i$ .

Paso 2. Calcular  $p_j$  y  $q_j$ ,  $j = 1, 2, \dots, n$ , recursivamente por

$$\begin{cases} p_j = \lfloor \frac{q_{j-1}}{2^{n-j}} \rfloor, \\ q_j = q_{j-1} - p_j 2^{n-j}, \quad j = 1, 2, \dots, n. \end{cases} \quad (2-18)$$

Por otro lado,  $i$  puede ser calculado a partir de la forma escalar de  $\{p_j\}$  por la fórmula:

$$i = \sum_{j=1}^n (1 - p_j) 2^{n-j} + 1. \quad (2-19)$$

Ahora construiremos una matriz, que llamaremos *matriz de reducción de grupo*, de la siguiente manera. Para  $j \geq 1$ , definimos

$$\Theta_j = \prod_{i=1}^j I_{2^{i-1}} \otimes [(I_2 \otimes W_{[2,2^{j-i}]}) M_r]. \quad (2-20)$$

**Lema 19** Si  $z_j = p_1 p_2 \cdots p_j$ , con  $p_i \in \Delta$ ,  $i = 1, 2, \dots, j$ , entonces

$$z_j^2 = \Theta_j z_j. \quad (2-21)$$

**Demostración.** Por inducción matemática, cuando  $j = 1$ , usando la proposición 15 se tiene que  $z_1^2 = p_1^2 = M_r p_1$ . En la fórmula (2-20) se tiene que  $\Theta_1 = (I_2 \otimes W_{[2,1]}) M_r$ . Como  $W_{[2,1]} = I_2$ , se sigue que  $\Theta_1 = M_r$ . Por lo tanto, (2-21) se cumple para  $j = 1$ . Ahora, suponiendo que (2-21)

es verdad para  $j = s$ , entonces para  $j = s + 1$  se tiene que

$$\begin{aligned}
z_{s+1}^2 &= p_1 p_2 \cdots p_{s+1} p_1 p_2 \cdots p_{s+1} \\
&= p_1 W_{[2,2^s]} p_1 [p_2 \cdots p_{s+1}]^2 \\
&= (I_2 \otimes W_{[2,2^s]}) p_1^2 [p_2 \cdots p_{s+1}]^2 \\
&= [(I_2 \otimes W_{[2,2^s]}) M_r] p_1 [p_2 \cdots p_{s+1}]^2.
\end{aligned}$$

Aplicando la hipótesis de inducción a la última igualdad, se tiene que

$$\begin{aligned}
z_{s+1}^2 &= (I_2 \otimes W_{[2,2^s]}) M_r p_1 \times \left( \prod_{i=1}^s I_{2^{i-1}} \otimes [(I_2 \otimes W_{[2,2^{s-i]}) M_r] \right) p_2 p_3 \cdots p_{s+1} \\
&= [(I_2 \otimes W_{[2,2^s]}) M_r] \times \left( \prod_{i=1}^s I_{2^i} \otimes [(I_2 \otimes W_{[2,2^{s-i]}) M_r] \right) p_2 p_3 \cdots p_{s+1}.
\end{aligned}$$

Así la conclusión entonces se sigue. ■

Otro concepto útil es el llamado *operador ficticio*, denotado por  $\sigma_d$  y definido como  $\sigma_d(p, q) = q$ , para todos  $p, q \in S$ . Es fácil notar que la matriz estructural del operador ficticio  $\sigma_d$  es

$$E_d = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}. \quad (2-22)$$

Se sigue de la forma en que está definida, que para cualesquiera dos variables lógicas  $X, Y$ ,

$$E_d X Y = Y \quad \text{o} \quad E_d W_{[2]} X Y = X. \quad (2-23)$$

Una variable lógica que aparece formalmente en una función lógica, pero que no afecta el valor de esa función, es llamada una *variable fabricada*.

**Lema 20** Sea  $x = \times_{i=1}^n p_i$ . Usando la forma vectorial, cada ecuación lógica  $f_i(p_1, p_2, \dots, p_n) = c_i$ , con  $i = 1, 2, \dots, m$ , en el sistema (2-17) puede ser expresada como

$$M_i x = c_i, \quad i = 1, 2, \dots, m, \quad (2-24)$$

donde  $M_i \in \mathcal{L}_{2 \times 2^n}$ .

**Demostración.** Supongamos que  $f_i$  es una función lógica de  $p_1, \dots, p_n$ . Sea  $M_i$  la matriz estructural de  $f_i$ . Entonces, (2-24) se sigue inmediatamente. Suponiendo que algunas  $p_j$ 's no aparecen en  $f_i$ , usando la técnica del operador ficticio, podemos obtener (2-24) introduciendo variables fabricadas. ■

**Teorema 21** Sean  $x = \times_{i=1}^n p_i$ , y  $b = \times_{i=1}^n c_i$ . El sistema de ecuaciones lógicas (2-17) puede ser convertido en una ecuación algebraica lineal como

$$Lx = b, \quad (2-25)$$

donde

$$L = M_1 \times_{j=2}^n [(I_{2^n} \otimes M_j)\Theta_n], \quad (2-26)$$

con  $M_i$  definida como en (2-24).

**Demostración.** Observamos que del lema 20 se tiene que  $x^2 = \Theta_n x$ . Multiplicando (2-24) obtenemos

$$\begin{aligned} b &= M_1 x M_2 x \cdots M_n x \\ &= M_1 (I_{2^n} \otimes M_2) x^2 M_3 x \cdots M_n x \\ &= M_1 (I_{2^n} \otimes M_2) \Theta_n x M_3 x \cdots M_n x \\ &= \cdots \\ &= M_1 (I_{2^n} \otimes M_2) \Theta_n (I_{2^n} \otimes M_3) \Theta_n \cdots (I_{2^n} \otimes M_n) \Theta_n x. \end{aligned}$$

Así, (2-26) se sigue. ■

Como  $L \in \mathcal{L}_{2^m \times 2^n}$  y  $b \in \Delta_{2^m}$ , es claro que la ecuación algebraica (2-25) tiene como solución a  $x \in \Delta_{2^n}$  si y sólo si  $b \in \text{Col}(L)$ . Expresando a  $L$  de forma condensada como  $L = \delta_{2^m} [i_1, i_2, \dots, i_{2^n}]$ , definimos el conjunto  $\Lambda = \{\lambda : \delta_{2^m}^{i_\lambda} = b, 1 \leq \lambda \leq 2^n\}$ . Usando esta notación, la solución de (2-25) es

$$x = \delta_{2^n}^\lambda, \quad \lambda \in \Lambda. \quad (2-27)$$

## Capítulo 3

# Redes Booleanas Aleatorias

La biología de sistemas no estudia genes individuales, proteínas o células aisladas. Más bien, investiga el comportamiento y las relaciones de todas las células, proteínas, ADN y ARN en un sistema biológico llamado *red celular*. Probablemente las redes más activas sean las *redes de regulación genética* [Kitano, 2002], las cuales, en respuesta a los cambios en su entorno, regulan crecimiento, replicación y muerte celular.

Un gen está constituido por una secuencia de bases nitrogenadas (*adenina, guaina, citosina, tiamina*) que codifican cómo un grupo de aminoácidos deben de ser acomodados para producir una proteína específica. Decimos que un gen es expresado cuando su proteína asociada es sintetizada. Este proceso es regulado por un conjunto de genes “reguladores” que actúan como *interruptores* y que pueden *encenderse* o *apagarse* unos con otros. El papel que juegan las proteínas en la diferenciación celular es determinante [Kappler *et al.*, 2003], por lo tanto, es importante estudiar cómo es regulada la expresión genética.

Las *redes booleanas aleatorias* (RBAs), introducidas por Stuart Kauffman [Kauffman, 1969], se han convertido en una poderosa herramienta para describir, analizar y simular redes celulares. En este modelo, un gen es cuantificado por sólo dos niveles: verdadero o falso. El estado de cada gen es determinado a partir del estado de sus genes *vecinos* mediante el uso de reglas lógicas. Las RBAs son genéricas [Gershenson, 2004a], porque al construirlas no suponemos ninguna funcionalidad ni conectividad de los nodos en particular. Estas, como su nombre lo indica, son generadas aleatoriamente. Las propiedades genéricas encontradas en el modelo pueden ser aplicadas a un sistema en particular, con la intención de entender su mecanismo subyacente.

### 3.1. El Modelo Clásico

Kauffman originalmente propuso el modelo de RBA, para apoyar la hipótesis de que los organismos vivos pueden ser construidos a partir de una forma aleatoria, sin la necesidad de elementos previamente programados [Kauffman, 1969]. Una RBA consiste en  $n$  nodos  $x_1, x_2, \dots, x_n$  que pueden tomar uno de los valores Booleanos 0 o 1. El *estado* de cada nodo es el valor que toma en un momento dado. Dicho estado es determinado por  $k$  conexiones provenientes de otros nodos o de él mismo. La forma en la que los nodos se afectan entre sí es a través de funciones lógicas, una para cada nodo, y cuyos argumentos son los estados de sus  $k$  nodos *vecinos*. Las conexiones y funciones son escogidas aleatoriamente cuando la RBA es generada y permanecen fijas durante la dinámica de la red. En el modelo clásico, el *esquema de actualización* es *sincrónico*, i.e., el estado de un nodo en el tiempo  $t + 1$  depende de los estados de sus  $k$  vecinos en el tiempo  $t$ .

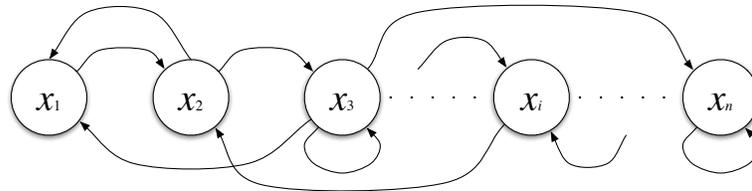


Figura 3-1: Topología de una red booleana aleatoria con  $k = 2$ . Cada nodo tiene exactamente 2 *entradas*, pero no necesariamente 2 *salidas*. El nodo  $x_3$  tiene tres salidas mientras que otros nodos como  $x_n$  tienen sólo una.

Generalmente, un estado inicial para una RBA es escogido aleatoriamente y la dinámica fluye de acuerdo con las funciones y el esquema de actualización. Como las RBAs son finitas ( $n$  nodos), la cantidad de estados posibles también es finita ( $2^n$  estados). Como la dinámica es determinista, tarde o temprano un estado será revisitado. Entonces, la red habrá llegado a un *atractor*. El número de estados en un atractor determina el *período* del atractor. Los *puntos fijos* o atractores *puntuales* tienen período uno (un solo estado), mientras que los *ciclos de estados* o atractores *cíclicos* tienen período mayor que uno (múltiples estados). Una RBA

puede tener uno o varios atractores. En Zhao [2005], se muestra que encontrar los puntos fijos y ciclos de una red booleana es un problema NP-completo<sup>1</sup>. Es importante notar la diferencia que hay entre la *topología* de la red (fig. 3-1), donde la estructura formada por los  $n$  nodos es representada de acuerdo a su *gráfica de conectividad*, y por otro lado, la *dinámica* de la red (fig. 3-2), donde lo que se observa es la trayectoria de los estados que toma la red a través del tiempo.



Figura 3-2: Evolución temporal (50 pasos) de una RBA con  $n = 20$  y  $k = 3$ , a partir de un estado inicial aleatorio. Los cuadros (*bits*) blancos representan 1 y los cuadros (*bits*) negros representan 0. La condición inicial es representada por la columna de hasta la izquierda y la dinámica fluye hacia la derecha. De esta manera, las *columnas* representan los estados de la red en un instante particular, mientras que cada *fila* representa la evolución temporal del estado de un nodo. Tomada de RBNLab [Gershenson, 2005].

En este modelo, se tiene que para cada nodo existen  $2^{2^k}$  funciones de actualización posibles, y cada uno tiene  $\frac{n!}{(n-k)!}$  combinaciones ordenadas posibles para  $k$  conexiones diferentes. Entonces, dadas  $n$  y  $k$ , la cantidad de redes posibles será [Harvey y Bossomaier, 1997]:

$$\left( \frac{2^{2^k} n!}{(n-k)!} \right)^n,$$

y aunque muchas de ellas sean equivalentes, es claro que el espacio de redes es inmenso. A este conjunto usualmente se le llama el *espacio de parámetros*.

El conjunto de estados visitados antes de llegar a un atractor es llamado *transcurso*, y el tiempo que pasa la red en estos estados se denomina *tiempo de transcurso*. El conjunto de estados que convergen hacia un atractor constituyen su *cuenca de atracción*. Las cuencas de diferentes atractores dividen el espacio de estados como un *campo de cuencas de atracción*. Las RBAs son *disipativas* en el sentido que varios estados pueden converger en un único estado (un

<sup>1</sup> Es decir, que no es posible encontrar una solución en un período razonable de tiempo.

estado puede tener varios predecesores), pero a partir de un estado la transición es determinista hacia un solo estado (un estado puede tener sólo un sucesor). El número de predecesores es también llamado *ingrado*. Los estados que no tienen predecesores son usualmente llamados estados “Jardín del Edén” (*ingrado* = 0), ya que estos únicamente pueden ser alcanzados si se encuentran en una condición inicial. La figura 3-3 ilustra los conceptos sobre las *transiciones de estado* presentados en este párrafo.

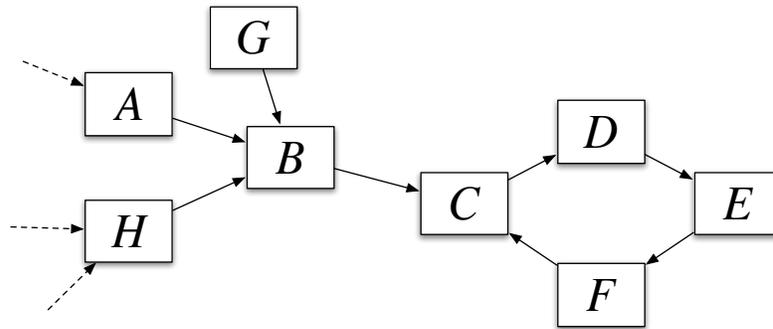


Figura 3-3: Red de estados. *B* es un sucesor del estado *A* y un predecesor del estado *C*. Los estados pueden tener varios predecesores (e.g., *B*), pero sólo un sucesor. *G* es un estado Jardín del Edén ya que no tiene predecesores. El atractor  $C \rightarrow D \rightarrow E \rightarrow F \rightarrow C$  tiene período 4.

## 3.2. Autoorganización en Redes Booleanas

Podemos describir a una RBA como un sistema autoorganizante simplemente porque tiene atractores [Gershenson, 2012]. Si interpretamos la presencia de atractores como una forma de “organización”, entonces podemos decir que la dinámica se autoorganiza hacia ellos.

En una RBA, la dinámica es dirigida por la iteración de un procedimiento global de actualización, que resulta en una sucesión de estados globales llamada la *trayectoria* del sistema. Aunque cada estado global determina una trayectoria, puede ocurrir que esta sea altamente impredecible, ya que como otros sistemas dinámicos, una RBA puede tener diferentes comportamientos dinámicos. Existen dos fases dinámicas: la fase *ordenada* y la fase *caótica*. La transición de fase está caracterizada por su *criticalidad*.

### 3.2.1. Comportamiento Dinámico

En la fase ordenada, la mayoría de los nodos no cambian su estado. A nivel global, estados similares tienden a converger al mismo atractor, y en promedio cada estado tiene muchos predecesores, por lo que muchos estados convergen hacia unos cuantos. Las RBAs en esta fase son robustas, ya que las perturbaciones no se esparcen debido a que la mayoría de los nodos se encuentran estáticos. Sus tiempos de transcurso son cortos, y la densidad de estados Jardín del Edén es alta.

En la fase caótica, la mayoría de los nodos están cambiando su estado constantemente. A nivel global, estados similares tienden a divergir hacia atractores diferentes, y en promedio cada estado tiene muy pocos predecesores, por lo que la convergencia es baja. Las RBAs en esta fase son frágiles, ya que las perturbaciones se esparcen a través de la red. Los tiempos de transcurso, en promedio, son muy largos, y la densidad de estados Jardín del Edén es relativamente baja.

En el régimen crítico, cerca de la transición entre las fases caótica y ordenada, los extremos de ambas son balanceados. Algunos nodos cambian constantemente, y algunos permanecen estáticos. Por lo tanto, las perturbaciones pueden esparcirse, pero no necesariamente a través de toda la red. A nivel global, estados similares tienen a tener trayectorias paralelas, y muchos estados tienen pocos predecesores mientras que pocos estados tienen muchos predecesores. De hecho, la distribución de los ingresos se aproxima a una ley de potencias [Wuensche, 1998; Aldana *et al.*, 2003], por lo que la red de estados presenta una topología libre de escala.

Una RBA, especialmente cuando se encuentra cerca de la transición de fase, no puede ser descrita de manera completa a través de la descripción de todos sus nodos. Se necesita más información debido a que las conexiones, que representan las interacciones del sistema, aportan datos relevantes en diferentes escalas [Gershenson, 2010] tanto dinámicas como funcionales. Si estructuralmente cada nodo representa un gen, entonces los tipos celulares, que son patrones generados por la regulación de dichos genes, pueden ser representados por los diferentes atractores a los que una red puede llegar. Cualquier configuración topológica arbitraria del campo de cuencas de atracción es posible, dado el conjunto correcto de *parámetros*.

### 3.2.2. Memoria Lejos del Equilibrio

En biología, la dinámica de los sistemas vivos no es una cuestión de equilibrio estable o inestable, sino de procesos *lejos del equilibrio* [Buiatti y Longo, 2013], i.e. que experimentan un flujo de energía o materia, pero que aún así son sistemas “estructuralmente estables”. Este comportamiento no conservativo (en el sentido termodinámico) que simultáneamente exhibe estabilidad estructural es una *rara* combinación de estabilidad e inestabilidad debida a la coexistencia de propiedades opuestas como *orden/caos* o *integración/diferenciación*.

El tamaño de una proteína sintetizada puede ser medido a través del número de aminoácidos que contiene. En general, el código genético especifica 20 aminoácidos esenciales. De esta manera, para formar una proteína de longitud 200, existen  $20^{200} \approx 10^{260}$  posibles configuraciones (arreglos de aminoácidos) diferentes. El *intervalo de tiempo* más pequeño que puede ser medido es el *tiempo* de Planck ( $10^{-43}$ s). Si todas las partículas del universo ( $10^{80}$ ), ignorando cuestiones como la separación<sup>1</sup>, trabajaran constantemente reacomodándose en cadenas de 200 elementos, tomaría  $10^{39}$  veces la edad del universo ( $10^{17}$ s) explorar todas las configuraciones posibles para formar una proteína de longitud 200 una sola vez.

Esto significa que físicamente no es posible que durante su evolución, la naturaleza explore todas las posibilidades de convertirse en lo que es. La evolución no es causal [Časlav, 2014] ni determinada en el sentido cuántico, sin embargo, tampoco puede ser aleatoria [Bailly y Longo, 2007]. Una característica fundamental de los sistemas vivientes es que funcionan en un estado que puede describirse como un proceso lejos del equilibrio, al *borde del caos*, donde cierta cantidad de información es preservada, cristalizada en algún patrón que se mantiene estable a lo largo del tiempo, pero esto, sin detener por completo la búsqueda de un óptimo global del sistema.

La noción de *memoria lejos del equilibrio* puede responder a la difícil tarea de explicar como las redes celulares, pese a las astronómicas longitudes de sus transcurros, llegan en un tiempo extremadamente rápido a un atractor. Por tanto, es plausible decir que el campo de cuencas de atracción de una RBA representa el substrato cognitivo de la red [Wuensche, 1998]. La falta de restricciones en la topología del campo de cuencas de atracción, y su enorme espacio de parámetros, sugieren que estas redes aleatorias son vehículos ideales para la categorización emergente y

---

<sup>1</sup>El argumento es originalmente de Kauffman (cf. <http://edge.org/response-detail/10193>).

puede ser dicho que tienen el potencial para representar el comportamiento cerebral. La relación que se observa entre *campo de cuencas/red* proporciona una fructuosa metáfora para analizar, y quizá entender mejor, la intrincada relación que hay entre los conceptos *mente/cerebro*.

### 3.3. Enfoque Sistémico

La teoría de las RBAs se ha investigado exhaustivamente tanto por deducciones teóricas como por simulaciones computacionales. Sin embargo, la descripción detallada de la evolución de una RBA no ha sido desarrollada de una manera propiamente teórica [Yingjun *et al.*, 2007], sino hasta años recientes mediante el desarrollo de nuevas herramientas matemáticas como el producto semi-tensorial de matrices [Cheng *et al.*, 2011; Meng y Feng, 2014; Zhao *et al.*, 2014].

A continuación, presentaremos algoritmos para obtener todos los puntos fijos, ciclos, transcurso, y cuencas de atracción. Teóricamente, estos algoritmos pueden darnos soluciones completas, sin embargo es difícil usarlos para redes grandes, ya que están limitados por la complejidad computacional que estos representan. Para hacerlo lo más general posible, supondremos que  $k = n$ , sin embargo en la práctica es usual que  $k < n$ .

#### 3.3.1. Dinámica de redes Booleanas

**Definición 22** *Una red booleana aleatoria es un conjunto de nodos,  $x_1, x_2, \dots, x_n$ , que simultáneamente interactúan entre sí. En cada tiempo determinado  $t = 0, 1, 2, \dots$ , un nodo toma sólo uno de dos diferentes valores posibles: 1 o 0. Así, la red puede ser descrita por un sistema de ecuaciones:*

$$\begin{cases} x_1(t+1) = f_1(x_1(t), x_2(t), \dots, x_n(t)), \\ x_2(t+1) = f_2(x_1(t), x_2(t), \dots, x_n(t)), \\ \vdots \\ x_n(t+1) = f_n(x_1(t), x_2(t), \dots, x_n(t)), \end{cases} \quad (3-1)$$

donde  $i = 1, 2, \dots, n$ , y cada  $f_i$  es una función lógica  $n$ -aria.

Una *gráfica de red*,  $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$ , está compuesta por un conjunto de nodos,  $\mathcal{N} = \{x_i : i = 1, \dots, n\}$ , y por un conjunto de aristas  $\mathcal{E} \subset \{x_1, \dots, x_n\} \times \{x_1, \dots, x_n\}$ . Si  $(x_i, x_j) \in \mathcal{E}$ , existe

una arista  $x_i \rightarrow x_j$ , lo cual significa que el nodo  $x_j$  es afectado por el nodo  $x_i$ . La gráfica de red es también conocida como la *gráfica de conectividad*.

La gráfica de red puede ser también expresada como una matriz de  $n \times n$ , llamada la *matriz de incidencia*, definida como

$$\mathcal{I} = (b_{ij}), \quad \text{donde} \quad b_{ij} = \begin{cases} 1, & (x_i, x_j) \in \mathcal{N}, \\ 0, & \text{en otro caso.} \end{cases} \quad (3-2)$$

**Ejemplo 23** Consideremos una red booleana,  $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$ , de tres nodos, dada por

$$\begin{cases} \sigma_1(t+1) = \sigma_2(t) \wedge \sigma_3(t), \\ \sigma_2(t+1) = \overline{\sigma_1(t)}, \\ \sigma_3(t+1) = \sigma_2(t) \vee \sigma_3(t). \end{cases} \quad (3-3)$$

Su conjunto de nodos es  $\mathcal{N} = \{x_1 := \sigma_1, x_2 := \sigma_2, x_3 := \sigma_3\}$ , y su conjunto de aristas es  $\mathcal{E} = \{(x_1, x_2), (x_2, x_1), (x_2, x_3), (x_3, x_1), (x_3, x_3)\}$ . Su matriz de incidencia es

$$\mathcal{I}(\mathcal{G}) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}. \quad (3-4)$$

En este caso, se tiene que  $k_1 = k_3 = 2$ , mientras que  $k_2 = 1$ . En el modelo clásico de Kauffman,  $k_i = k_j \quad \forall i, j \in \{1, 2, \dots, n\}$ <sup>2</sup>. Su gráfica de conectividad está ilustrada en la figura 3-4 mientras que la figura 3-5 representa un atractor de período 4.

Nuestra primer misión será convertir la dinámica de una red booleana en una forma algebraica, es decir, expresarla como un sistema lineal discreto convencional. Usando la forma vectorial de  $x_i(t) \in \Delta = \{\delta_2^1, \delta_2^2\}$ , donde  $\delta_2^i$  es la  $i$ -ésima columna de la matriz identidad  $I_2$ , definimos  $x(t) = x_1(t)x_2(t) \cdots x_n(t) = \times_{i=1}^n x_i(t)$ .

---

<sup>2</sup>Si esta condición se satisface para una gráfica, entonces decimos que la gráfica es regular.

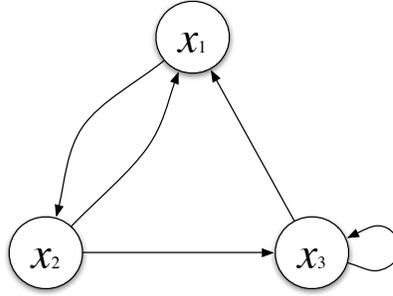


Figura 3-4: Gráfica de la red (3-3).

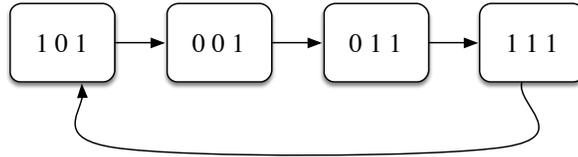


Figura 3-5: Trayectoria de la condición inicial 101 en la red (3-3).

Usando el teorema 16, existen matrices estructurales,  $M_i = M_{f_i}$  con  $i = 1, \dots, n$ , tales que

$$x_i(t + 1) = M_i x(t), \quad i = 1, 2, \dots, n. \quad (3-5)$$

Multiplicando todas las ecuaciones en (3-5) se tiene que

$$x(t + 1) = M_1 x(t) M_2 x(t) \cdots M_n x(t). \quad (3-6)$$

Y usando el teorema 21, la ecuación (3-6) puede ser expresada como

$$x(t + 1) = Lx(t), \quad (3-7)$$

donde

$$L = M_1 \prod_{j=2}^n [(I_{2^n} \otimes M_j) \Theta_n]$$

es llamada la *matriz de transición*.

**Teorema 24** *La dinámica de la red booleana (3-1) está determinada de forma única por el*

sistema dinámico lineal (3-7).

**Demostración.** En (3-7) es claro ver que  $x(t) = L^t x(0)$ , para  $t = 0, 1, 2, \dots$ . Entonces, se tiene que  $x_i(t) = M_i L^{t-1} x(0)$ , con  $i = 1, 2, \dots$ . Por lo tanto, (3-7) determina de forma única la dinámica de (3-1). ■

La ecuación (3-7) es llamada la *forma algebraica* de la red (3-1). La ecuación (3-5) es llamada la *forma algebraica por componentes* de la red (3-1). La ecuación (3-7) es un sistema lineal estándar con  $L$  una matriz booleana cuadrada. Por tanto, todos los métodos y resultados clásicos para sistemas lineales pueden ser usados para analizar la dinámica de la red booleana.

**Ejemplo 25** La dinámica de la red booleana  $\mathcal{G}$  del ejemplo 23 está dada por (3-3). Expresándola en forma algebraica, se tiene

$$\begin{cases} \sigma_1(t+1) = M_c \sigma_2(t) \sigma_3(t), \\ \sigma_2(t+1) = M_n \sigma_1(t), \\ \sigma_3(t+1) = M_d \sigma_2(t) \sigma_3(t). \end{cases} \quad (3-8)$$

Definiendo  $x(t) = \sigma_1(t) \sigma_2(t) \sigma_3(t)$ , podemos calcular  $L$  como

$$\begin{aligned} x(t+1) &= M_c \sigma_2 \sigma_3 M_n \sigma_1 M_d \sigma_2 \sigma_3 \\ &= M_c (I_4 \otimes M_n) \sigma_2 \sigma_3 \sigma_1 M_d \sigma_2 \sigma_3 \\ &= M_c (I_4 \otimes M_n) (I_8 \otimes M_d) \sigma_2 \sigma_3 \sigma_1 \sigma_2 \sigma_3 \\ &= M_c (I_4 \otimes M_n) (I_8 \otimes M_d) W_{[2,4]} \sigma_1 \sigma_2 \sigma_3 \sigma_2 \sigma_3 \\ &= M_c (I_4 \otimes M_n) (I_8 \otimes M_d) W_{[2,4]} \sigma_1 \sigma_2 W_{[2]} \sigma_2 \sigma_3 \sigma_3 \\ &= M_c (I_4 \otimes M_n) (I_8 \otimes M_d) W_{[2,4]} (I_4 \otimes W_{[2]}) \sigma_1 M_r \sigma_2 M_r \sigma_3 \\ &= M_c (I_4 \otimes M_n) (I_8 \otimes M_d) W_{[2,4]} (I_4 \otimes W_{[2]}) (I_2 \otimes M_r) (I_4 \otimes M_r) \sigma_1 \sigma_2 \sigma_3. \end{aligned} \quad (3-9)$$

Así, el sistema (3-3) puede ser expresado en forma matricial como

$$x(t+1) = Lx(t),$$

donde la matriz de transición de la red es

$$\begin{aligned} L &= M_c(I_4 \otimes M_n)(I_8 \otimes M_d)W_{[2,4]}(I_4 \otimes W_{[2]})(I_2 \otimes M_r)(I_4 \otimes M_r) \\ &= \delta_8[3 \ 7 \ 7 \ 8 \ 1 \ 5 \ 5 \ 6]. \end{aligned}$$

### 3.3.2. Puntos Fijos y Ciclos

Es común usar  $L$  para denotar tanto a una matriz lógica, como a su mapeo lineal correspondiente. Considerando la red booleana descrita por la ecuación (3-7) podemos enunciar el siguiente resultado.

**Lema 26**  $\text{Col}_i(L) \in \Delta_{2^n}$ , donde  $L \in \mathcal{L}_{2^n \times 2^n}$ .

**Demostración.** Lo único que hay que mostrar es que  $\text{Col}(L) \subset \Delta_{2^n}$ . Supongamos que existe  $j$ , con  $1 \leq j \leq 2^n$ , tal que  $\text{Col}_j(L) \notin \Delta_{2^n}$ . Entonces, cuando  $x(t) = \delta_{2^n}^j$ , se tiene que  $x(t+1) = Lx(t) = \text{Col}_j(L) \notin \Delta_{2^n}$ , lo cual es una contradicción. ■

Ahora, podemos definir la noción de *atractor* de la siguiente manera:

1. Un estado  $x_0 \in \Delta_{2^n}$  es llamado un *punto fijo* o atractor puntual del sistema (3-7) si  $Lx_0 = x_0$ .
2. El conjunto  $\{x_0, Lx_0, \dots, L^p x_0\}$  es llamado un *ciclo* o atractor cíclico del sistema (3-7) con *longitud* o período  $p$ , si  $L^p x_0 = x_0$ , y los elementos del conjunto  $\{x_0, Lx_0, \dots, L^{p-1} x_0\}$  son distintos dos a dos.

**Teorema 27** Consideremos la red booleana (3-1). Se tiene que  $\delta_{2^n}^i$  es un punto fijo si y sólo si, en su forma algebraica (3-7), el elemento diagonal  $\ell_{ii}$  de la matriz de transición de la red  $L$  es igual a 1.

**Demostración.** Supongamos que  $\delta_{2^n}^i$  es un punto fijo. Se tiene que  $L\delta_{2^n}^i = \text{Col}_i(L)$ . Es evidente que  $\delta_{2^n}^i$  es punto fijo de (3-1) si y sólo si  $\text{Col}_i(L) = \delta_{2^n}^i$ , así el teorema se sigue. ■

Con el resultado anterior, es claro que el número de puntos fijos de la red (3-1), denotado por  $N_e$ , es igual al número  $i$  para el cual  $\ell_{ii} = 1$ . Equivalentemente,  $N_e = \text{tr}(L)$ , donde  $\text{tr}$  denota

la traza de una matriz. Si  $\ell_{ii} = 1$ , entonces  $\text{Col}_i(L)$  es llamada una columna *diagonal diferente de cero* de  $L$ .

Ahora consideremos los ciclos de la red Boolean (3-1). Dados dos enteros positivos  $p, s \in \mathbb{Z}_+$ , decimos que  $s$  es un *factor propio* de  $p$  si  $s < p$ , y  $p/s \in \mathbb{Z}_+$ . El conjunto de factores propios de  $p$  es denotado por  $\mathcal{P}(p)$ .

**Teorema 28** *El número de ciclos de longitud  $s$ , denotado por  $N_s$ , esta determinado inductivamente por*

$$\begin{cases} N_1 = N_e, \\ N_s = \frac{\text{tr}(L^s) - \sum_{p \in \mathcal{P}(s)} p N_p}{s}, \quad 2 \leq s \leq 2^n. \end{cases} \quad (3-10)$$

**Demostración.** Si  $\delta_{2^n}^i$  es un elemento de un ciclo de longitud  $s$ , entonces  $L^s \delta_{2^n}^i = \delta_{2^n}^i$ . Por el teorema 27,  $\text{Col}_i(L^s)$  es una columna diagonal diferente de cero de  $L^s$ , la cual suma 1 a  $\text{tr}(L^s)$ . Observemos que si  $\delta_{2^n}^p$  es un ciclo de longitud  $p \in \mathcal{P}(s)$ , entonces también se tiene que  $L^s \delta_{2^n}^p = \delta_{2^n}^p$ , y  $\text{Col}_p(L^s)$  también sumará 1 a  $\text{tr}(L^s)$ . Estos elementos diagonales tendrán que ser restados de  $\text{tr}(L^s)$ . De esta forma, la segunda parte de la fórmula (3-10) es obvia.

Para la cota superior de  $s$ , como  $x(t)$  puede tomar a lo más  $2^n$  valores posibles, es evidente que la longitud de cualquier ciclo es menor o igual que  $2^n$ . ■

Ahora, veremos como encontrar ciclos. Si  $\text{tr}(L^s) - \sum_{p \in \mathcal{P}(s)} p N_p > 0$ , entonces llamamos a  $s$  una *potencia no trivial*. Suponiendo que  $s$  es una potencia no trivial, denotamos por  $\ell_{ii}^s$  a la  $(i, i)$ -ésima entrada de la matriz  $L^s$ . Así, definimos  $C_s = \{i : \ell_{ii}^s = 1\}$ , con  $s = 1, 2, \dots, 2^n$ , y  $D_s = C_s \cap_{i \in \mathcal{P}(s)} \overline{C_i}$ , donde  $\overline{C_i}$ , como es usual, denota el complemento de  $C_i$ . Definimos ahora al conjunto

$$\Omega = \bigcup_{i=1}^p C_i,$$

donde  $\{C_i : i = 1, \dots, p\}$  es el conjunto de todos los atractores de la red.

**Ejemplo 29** *En referencia al ejemplo 23, es fácil ver que  $\text{tr}(L^t) = 0$  cuando  $t \leq 3$ , y  $\text{tr}(L^t) = 4$  cuando  $t \geq 4$ . Usando el teorema 28, concluimos que existe sólo un ciclo de longitud 4 (fig. 3-5). Más aún, observamos que*

$$L^4 = \delta_8[1 \ 3 \ 3 \ 1 \ 5 \ 7 \ 7 \ 3].$$

Cada columna diagonal diferente de cero puede entonces generar al ciclo. De esta forma, escogiendo  $Z = \delta_8^1$ , tenemos que

$$LZ = \delta_8^3, \quad L^2Z = \delta_8^7, \quad L^3Z = \delta_8^5, \quad L^4Z = Z.$$

Mediante el algoritmo que usa la ecuación (2-18), convirtiendo las formas vectoriales de regreso a las formas escalares de  $\sigma_1(t), \sigma_2(t)$ , y  $\sigma_3(t)$ , se tiene el ciclo como  $(1, 1, 1) \rightarrow (1, 0, 1) \rightarrow (0, 0, 1) \rightarrow (0, 1, 1) \rightarrow (1, 1, 1)$ .

A continuación, consideraremos el período de transcurso de una red, i.e., el mínimo número de pasos en los que cualquier punto llega al conjunto límite  $\Omega$ , el cual consiste en todos los puntos fijos y ciclos. Primero, cabe mencionar que  $L$  tiene a lo más  $r = 2^n \times 2^n$  valores posibles independientes. Entonces, construimos un serie de matrices de  $r+1$  como  $L_0 = I_{2^n}, L, L^2, \dots, L^r$ , donde debe de haber dos matrices iguales. Sea  $r_0 < r$  el menor  $i$  tal que  $L^i$  aparece otra vez en la serie. Es decir, existe  $p > i$  tal que  $L^i = L^p$ , o más precisamente,

$$r_0 = \arg \min_{0 \leq i \leq p} \{L^i \in \{L^{i+1}, L^{i+2}, \dots, L^r\}\}. \quad (3-11)$$

Entonces, tal  $r_0$  existe y es obvio que empezando una trayectoria desde cualquier estado, la trayectoria entrará en un ciclo después de  $r_0$  iteraciones.

Para cualquier estado  $x_0$ , el período de transcurso de  $x_0$ , denotado por  $T_t(x_0)$ , es el menor  $p$  que satisface  $x(0) = x_0$  y  $x(p) \in \Omega$ . El período de transcurso de una red booleana, denotado por  $T_t$ , está definido como  $T_t = \max_{\forall x \in \Delta_{2^n}} (T_t(x))$ .

**Teorema 30** La  $r_0$  definida en la ecuación (3-11) es el período de transcurso del sistema (3-1), es decir,  $T_t = r_0$ .

**Demostración.** Supongamos que

$$L^{r_0} = L^{r_0+T} \quad (3-12)$$

y que  $T > 0$  es el menor número positivo que verifica (3-12). Por definición,  $r_0 + T \leq r$ . Afirmamos que si existe un ciclo de longitud  $t$ , entonces  $t$  es un factor de  $T$ . Por contradicción,

supongamos que  $T \pmod t = s$  y  $1 \leq s \leq t$ . Sea  $x_0$  un estado en el ciclo.  $L^{r_0}x_0$  es entonces también un estado en el mismo ciclo. Entonces,  $L^{r_0}x_0 = L^{r_0+T}x_0 = L^T L^{r_0}x_0 = L^s(L^{r_0}x_0) \neq L^{r_0}x_0$ , lo cual es una contradicción.

De la ecuación (3-12) y de la definición de  $T_t$ , es obvio que  $T_t \leq r_0$ . Para mostrar que  $T_t = r_0$ , supongamos que  $T_t < r_0$ . Por definición, para todo  $x$ ,  $L^{T_t}x$  está en un ciclo cuya longitud es un factor de  $T$ . Por lo tanto,

$$L^{T_t}x = L^T L^{T_t}x = L^{T_t+T}x, \quad \forall x. \quad (3-13)$$

Es fácil ver que si para todo  $x \in \Delta_{2^n}$ , la ecuación (3-13) se cumple, entonces  $L^{T_t} = L^{T_t+T}$ , lo cual contradice la definición de  $r_0$ . ■

Una red booleana es *globalmente convergente* si su conjunto límite  $\Omega$ , consiste en un único punto fijo. La convergencia global también es llamada *estabilidad global*. Esta definición de convergencia global, significa que desde cualquier estado inicial, la trayectoria de la red converge a un único punto fijo.

### Definición 31

1.  $S_i$  es llamado la cuenca del atractor  $C_i$ , si  $S_i$  es el conjunto de estados que convergen a  $C_i$ . Es decir,  $p \in S_i$  si y sólo si la trayectoria  $x(t, p)$  con  $x(0, p) = p$  satisface que  $x(t, p) \in C_i$  para  $t \geq T_t$ .
2. El estado  $q$  es llamado predecesor del estado  $p$ , si  $p = x(1, q)$ .

### Nota 32

- Dados  $C \subset \Delta_{2^n}$ , y  $L^{-1}(C) = \{q : Lq \in C\}$ , el conjunto de predecesores del estado  $p$  es el conjunto  $L^{-1}(p)$ .
- $\Delta_{2^n} = \bigcup_{i=1}^p S_i$ . Más aun, como  $\{S_i : i = 1, \dots, k\}$  son disjuntos, es una partición del espacio de estados  $\mathcal{X} = \Delta_{2^n}$ .
- Una forma alternativa de describir las cuencas de atracción es de la siguiente manera. Consideremos el sistema (3-7). Sea  $p \in \mathcal{X}$ . Definimos al conjunto de sucesores de  $p$  como

$Sc(p) = \{q : \exists m \geq 0, q = L^m p\}$ . Si  $Sc(p) \cap Sc(q) \neq \emptyset$ , entonces decimos que  $p$  y  $q$  son equivalentes, y lo denotamos por  $p \sim q$ .

- Es fácil ver que  $\sim$  es una relación de equivalencia. Así, la clase de equivalencia de  $p$  está dada por  $[p] = \{q : q \sim p\}$ , y entonces  $\{[p] : p \in \mathcal{X}\}$  es una partición de  $\mathcal{X}$ .
- En esta partición, cada componente contiene exactamente un atractor. Entonces, cada componente es la cuenca del único atractor contenido en ella.
- Para un atractor  $C$ , sea  $p \in C$ . Su cuenca de atracción es entonces  $[p]$ .

Para encontrar  $S_i$ , comenzamos de cada punto  $p \in C_i$ . Si podemos encontrar sus predecesores  $L^{-1}(p)$ , entonces para cada estado  $p_1 \in L^{-1}(p)$ , también podemos encontrar  $L^{-1}(p_1)$ . Continuando este proceso, después de  $T_t$  iteraciones, obtenemos un árbol de estados que convergen en  $p$ . Es decir,

$$S_i = C_i \cup L^{-1}(C_i) \cup L^{-2}(C_i) \cup \dots \cup L^{-T_t}(C_i). \quad (3-14)$$

Y para encontrar  $L^{-1}(p)$ , es fácil ver que la siguiente fórmula se cumple:

$$\begin{cases} L^{-1}(p) = \{\delta_{2^n}^j : \text{Col}_j(L) = p\}, \\ L^{-m}(p) = \{\delta_{2^n}^j : \text{Col}_j(L^m) = p\}, \quad m = 2, \dots, T_t. \end{cases} \quad (3-15)$$

### 3.4. Transiciones de Fase

Las transiciones de fase pueden ser estudiadas de varias formas. Estas pueden ser obtenidas estadísticamente o también de manera analítica. Derrida y Pomeau fueron los primeros en determinar analíticamente que la fase crítica se encuentra cuando  $k = 2$  [Derrida y Pomeau, 1986]. El método que usaron es llamado *aproximación recocida de Derrida*, el cual toma dos configuraciones iniciales aleatorias, y mide su traslape, usualmente con la *distancia de Hamming normalizada*, definida como

$$H(A, B) = \frac{1}{n} \sum_i^n |a_i - b_i|. \quad (3-16)$$

Entonces, un paso de la dinámica es computado, y el traslape es vuelto a medir. Después, se escoge nuevamente un conjunto de funciones y conexiones aleatoriamente. Se puede mostrar que esto evoluciona en un mapeo uno-dimensional. Cuando  $k < 2$ , el mapeo converge a un punto fijo estable  $H = 0$ , lo que significa que estados diferentes tienden a converger (dinámica ordenada). En  $k = 2$ , este punto fijo se vuelve inestable, por lo que para  $k > 2$ , estados diferentes tienden a divergir (dinámica caótica) [Gershenson, 2004a].

Desde antes ya se había encontrado en simulaciones que RBAs con  $k \leq 2$  se encuentran en el régimen ordenado, mientras que redes con  $k \geq 3$  se encuentran en el régimen caótico. Más recientemente, se observó que en la práctica, el tamaño de las redes puede jugar un papel importante en las transiciones de fase. Para RBAs grandes, midiendo el promedio de las diferencias de la distancia de Hamming normalizada (3-16) para  $t = 0$  y  $t \rightarrow \infty$ , la transición de fase está dada para  $k$  recorrida hacia 1, mientras que para redes muy pequeñas, para  $k$  recorrida hacia 3 [Gershenson, 2004b,c].

En la figura 3-6 podemos apreciar los distintos regímenes dinámicos de RBAs en diferentes fases dependiendo de su conectividad  $k$ .

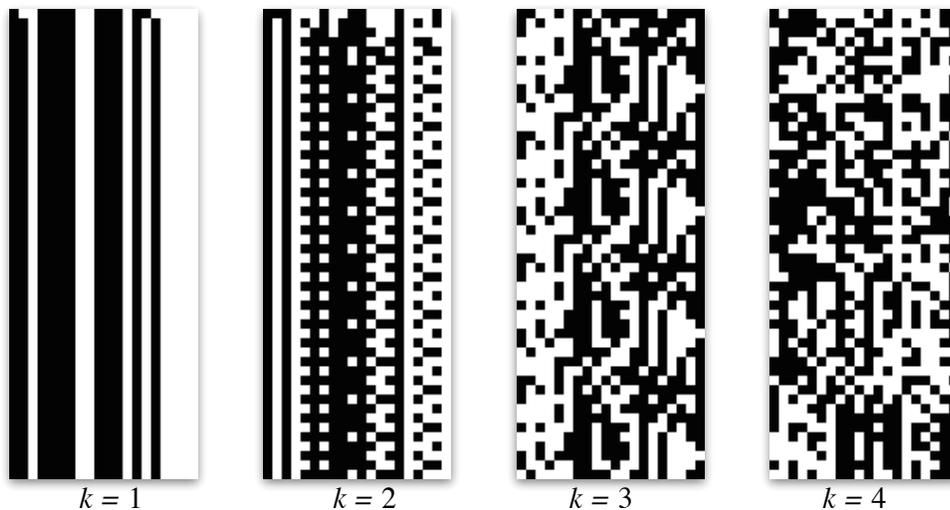


Figura 3-6: Trayectorias a través del espacio de estados de RBAs en diferentes fases. Un *bit* representa el estado de un nodo,  $n = 20$ . Los estados iniciales están hasta arriba, el tiempo corre hacia abajo.

## Capítulo 4

# Redes Difusas Aleatorias

No hay modelo que pueda decirse efectivo si fundamentalmente es más complejo que lo que intenta modelar. Por otro lado, ir al extremo en la simplificación puede potencialmente ignorar factores relevantes pero no primarios, o incluso dejar a un lado la posibilidad de representar observaciones que por su naturaleza son ambiguas o poco claras. Es por esto, que obtener un balance entre modelos burdos y detallados es una cuestión fundamental para el desarrollo tanto de la ingeniería como de la ciencia.

Durante décadas recientes, las redes Boleanas aleatorias (RBAs) se han convertido en populares modelos utilizados en la biología de sistemas para tratar principalmente de entender los mecanismos subyacentes en las redes de regulación genética. Esta popularidad puede estar asociada al hecho de que son modelos muy generales. Al momento de construirlas, ninguna funcionalidad o estructura en particular es asumida. Sin embargo, la *idealización* booleana ha recibido fuertes críticas, basadas en el supuesto de que el uso de variables booleanas restringe las capacidades del modelo para recrear los procesos observados en experimentos formales.

En este capítulo nos dedicamos a estudiar la borrosa línea divisoria entre la complejificación y la utilidad de estos modelos. Para hacer esto, *difusificamos* el espacio de redes dinámicas aleatorias, presentando un modelo difuso que puede ir de Booleano a continuo, y analizamos hasta que punto, el incremento gradual del número de valores que un nodo puede tomar, afecta el comportamiento dinámico global en las mismas redes.

## 4.1. Modelos Booleanos y Continuos

Generar de manera aleatoria una red es muy conveniente para modelar sistemas cuyos mecanismos subyacentes no han sido aún explícitamente reconocidos [Gershenson, 2004a]. Las *redes dinámicas* (RDs) son una clase de sistemas dinámicos discretos que representan el estado de cada variable individual de un sistema como los nodos de una red. Las variables del modelo toman valores en conjuntos finitos y se desarrollan a lo largo del tiempo en pasos discretos. En cada instante, el valor de una variable es determinado por una *regla de actualización* que depende determinísticamente de los valores de otras variables en un instante anterior.

Una gran variedad de sistemas dinámicos de este tipo han sido profundamente estudiados durante las últimas décadas [Kauffman, 1969; Derrida y Pomeau, 1986; Bagley y Glass, 1996; Harvey y Bossomaier, 1997; Nichitiu y Rémila, 1998; Wuensche, 1998; Solé *et al.*, 2000; Wolfram, 2002; Aldana *et al.*, 2003; Kappler *et al.*, 2003; Gershenson, 2004a,b,c; Tosić, 2005; Yingjun *et al.*, 2007; Wittmann *et al.*, 2009; Cheng *et al.*, 2011; Poblanno-Balp y Gershenson, 2011; Wittmann y Theis, 2011]. Las RBAs con sus diferentes esquemas de actualización [Gershenson, 2002], tanto sincrónicos como asincrónicos, son todas instancias de RDs .

### 4.1.1. Método Sintético

Somos seres finitos y todo lo que podemos experimentar y comprender es por lo tanto finito. Por otro lado, una cantidad astronómica, en la práctica, es virtualmente infinita. Esto presenta un dilema a la hora de intentar construir modelos precisos de la variedad de fenómenos que ocurren en la naturaleza. Richard Feynman describió, en un artículo seminal para la computación cuántica [Feynman, 1982], su sentimiento frente a este dilema de la siguiente manera: “... *va a ser necesario que todo lo que ocurre en un volumen finito de espacio y tiempo tenga que ser exactamente analizable con un número finito de operaciones lógicas.*”.

Claramente, la simulación más precisa de un hecho es la reconstrucción del fenómeno mismo. Dicho de otra forma, siguiendo una postura similar a la de la *cibernética* [Rosenblueth y Wiener, 1945], el conocimiento científico puede ser descrito como una serie de modelos abstractos, preferiblemente formales pero a veces materiales, que cuando se aproximan asintóticamente a la complejidad de la situación original, tienden a ser idénticos con el sistema original, y en el

límite, se convierten en el sistema mismo.

La abstracción, entonces, se vuelve un proceso necesario para lograr cierto entendimiento en el análisis y control de cualquier parte significativa del universo. Dada una observación descrita en una forma sistémica, un modelo más complejo que el sistema mismo no es un modelo en lo absoluto, es simplemente una mala descripción del fenómeno estudiado. Por otro lado, la falta de consideración de características fundamentales para que el modelo recree, en cierta medida, la realidad observada, refleja la inutilidad de la descripción trivializada sobre las observaciones hechas.

Medir la *reducción de complejidad* [Kolmogorov, 1965; Chaitin, 1969] que un modelo puede lograr es importante para comprender hasta que punto podemos extender la utilización de dicho modelo. Fue esta la intención del trabajo presente. Notamos que en nuestra descripción de la complejidad y la utilidad de un modelo no se toma en cuenta la capacidad de predicción de este, debido a que el método que utilizamos toma un enfoque *sintético* [Gershenson *et al.*, 2014]. En este, un modelo es generado con factores abstraídos de su contexto. Después, a partir del modelo se construye un sistema artificial que recrea, sin pretender igualar, la estructura y funcionalidad del sistema original. La verificación del modelo es a través de la reproducción de los hechos observados.

#### 4.1.2. Idealización Booleana

Se pueden encontrar en la literatura diversas críticas, tanto a la idealización booleana [Yingjun *et al.*, 2007; Wittmann *et al.*, 2009], como a la discretización temporal [Bagley y Glass, 1996; Kappler *et al.*, 2003] de las RBAs. La irreducibilidad de la complejidad computacional [Wolfram, 2002; Tosić, 2005; Zhao, 2005] que representa encontrar propiedades emergentes que simulen sistemas vivientes en una RBA, ha sido un factor determinante que continúa limitando tanto el tamaño como la forma en la que estos modelos pueden ser estudiados.

El uso del modelo Booleano para cuantificar datos de expresión genética representados con variables que pueden tomar valores reales, ha sido justificado [Shmulevich y Zhang, 2002] desarrollando algoritmos que discretizan las medidas continuas, evaluando la información recolectada en comparación con un umbral determinado. Si una observación se encuentra por arriba de dicho umbral, entonces se le asigna el valor 1; si este no es el caso, la medición es

representada con el valor 0. Otros estudios [Karlebach, 2013] han probado la efectividad de métodos para inferir un formalismo Booleano a partir de información continua.

Los enfoques multivaluados [Nichitiu y Rémila, 1998; Wuensche, 1998; Solé *et al.*, 2000; Wuensche, 2008; Wittmann y Theis, 2011] emergieron de la consideración de que en algunos sistemas naturales, las tareas que se realizan a niveles individuales están claramente definidas en el tiempo, por lo que la descripción discreta de la dinámica es ciertamente plausible, a diferencia de las transiciones entre un estado y otro, que pueden no ser tan claras como para ser representadas únicamente con variables booleanas.

El nuevo modelo que presentamos a continuación está inspirado en estas propuestas, sin embargo, extiende aún más las capacidades de los modelos multivaluados actuales en el sentido en el que una red tiene la posibilidad, dependiendo de un grado en una escala infinita (parcialmente ordenada), de incrementar la cantidad de estados que pueden tomar sus nodos en cada instante.

## 4.2. Redes Difusas Aleatorias

Las redes difusas aleatorias (RDAs) son modelos que extienden a las RBAs clásicas propuestas por Kauffman [Kauffman, 1969]. Una RDA es una red dinámica constituida por un conjunto de  $n$  nodos  $\{x_1, x_2, \dots, x_n\}$  que representan el comportamiento de cada variable individual de un sistema. El estado de cada nodo  $x_i$  está determinado por  $k_i$  nodos escogidos aleatoriamente  $\{x_{i_1}, \dots, x_{i_{k_i}}\}$ , y a los que se encuentra conectado. Estas conexiones representan la interdependencia de los elementos dentro del sistema. Como en el modelo clásico de RBAs [Kauffman, 1969; Wuensche, 1998; Aldana *et al.*, 2003; Gershenson, 2004a],  $k_i = k_j, \forall i, j$ , por lo que cada nodo tiene exactamente  $k$  conexiones.

### 4.2.1. Estructura

La forma en la que el estado de los  $k_i$  elementos determinan el valor de  $x_i$  es a través de una función lógica  $f_i(x_{i_1}, \dots, x_{i_{k_i}})$ , que para cada combinación de entradas ( $2^{k_i}$  posibilidades) regresa  $f_i = 1$  con probabilidad  $p$  y  $f_i = 0$  con probabilidad  $1 - p$ . Una vez que las conexiones y funciones son escogidas aleatoriamente, la estructura de la red no cambia. El comportamiento

$x_{i_1}(t)$	$x_{i_2}(t)$	$\dots$	$x_{i_k}(t)$	$x_i(t+1)$
1	1	$\dots$	1	1
1	1	$\dots$	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	0	$\dots$	1	0
1	0	$\dots$	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
0	0	$\dots$	1	0
0	0	$\dots$	0	1

Tabla 4-1: Tabla de verdad para el nodo  $x_i$ . Hay una columna por cada uno de los  $k$  nodos vecinos de  $x_i$ . La columna de hasta la derecha representa el estado de  $x_i$  en el instante  $t+1$ .

dinámico es generado por medio del esquema de actualización del modelo clásico, donde todos los nodos son actualizados de manera sincrónica:

$$f_i(t+1) = f_i(x_{i_1}(t), \dots, x_{i_{k_i}}(t)). \quad (4-1)$$

Así, cada  $f_i$  puede ser representada como una *tabla de verdad*, donde la columna de hasta la derecha representa el valor de  $x_i(t+1)$  determinado por  $k_i$  entradas representadas en el resto de las columnas combinadas exhaustivamente en  $2^{k_i}$  filas (ver tabla 4-1).

#### 4.2.2. Difusificación

En una RDA la difusificación del espacio de estados es controlada con un parámetro que llamamos  $\phi \in [0, 1]$ . Si  $\phi = 0$ , tenemos el caso Booleano donde el conjunto de posible valores que un nodo puede tomar es idéntico el ordinal<sup>1</sup>  $2 = \{0, 1\}$ . Por otro lado, si  $\phi = 1$ , los estados varían sobre el conjunto  $\{r \in \mathbb{Q} : 0 \leq r \leq 1\}$ , de todos los números racionales entre 0 y 1, y que claramente es un conjunto de cardinal<sup>2</sup>  $\omega$ .

Denotamos a la *base* de la red por el cardinal  $\mathfrak{b}$ , definido para  $0 \leq \phi < 1$  específicamente como el número

$$\mathfrak{b} = \mathfrak{b}(\phi) = 1 + \left\lceil \frac{1}{1 - \phi} \right\rceil,$$

<sup>1</sup> Concebimos a los ordinales de tal manera que cada ordinal es el conjunto de todos los ordinales más pequeños, y los ordinales finitos son idénticos a los números naturales. El tipo de orden de los números naturales es  $\omega$  y, por lo anterior, es el conjunto de todos los números naturales.

<sup>2</sup> Un ordinal es un *ordinal sucesor* si es de la forma  $\xi + 1$ . Consideramos que los cardinales son idénticos a los ordinales iniciales. Así,  $\omega$  es también el cardinal de los números naturales.

donde  $\lceil x \rceil$  es el menor entero que es mayor o igual que  $x$ , i.e.  $\lceil x \rceil = \min\{z \in \mathbb{Z} : z \geq x\}$ . De esta manera  $\mathfrak{b}$  representa el número de estados que cada nodo puede tomar dependiendo del parámetro  $\phi$ . Es fácil ver que  $\mathfrak{b} \rightarrow \omega$  cuando dejamos que  $\phi \rightarrow 1$ .

Dado  $0 \leq \phi \leq 1$ , interpretamos cada nodo de la red como una variable individual del sistema expresándolo como una variable lógica que puede tomar valores a lo largo del tiempo en el conjunto

$$S_\phi = \left\{ 1, \frac{\mathfrak{b}(1-\phi)-2}{\mathfrak{b}(1-\phi)-1}, \frac{\mathfrak{b}(1-\phi)-3}{\mathfrak{b}(1-\phi)-1}, \dots, 0 \right\}.$$

Esto lo representamos como una función  $x_i : \omega \rightarrow S_\phi$ , donde  $x_i(t) \in S_\phi$  con  $i = 1, 2, \dots, n$ . De esta forma, para cada  $\phi \in [0, 1]$ , el estado de un nodo  $x_i$  en el instante  $t$  puede tomar  $\mathfrak{b}$  valores diferentes. Por lo tanto una RDA con  $0 \leq \phi \leq 1$  es un modelo  $\mathfrak{b}$ -valuado donde  $2 \leq \mathfrak{b} \leq \omega$ . Claramente una RDA con  $\phi = 0$  es una RBA clásica.

También, es claro que dada la sucesión  $\phi_1 > \phi_2 > \phi_3 \dots$  se tiene que  $S_{\phi_1} \subseteq S_{\phi_2} \subseteq S_{\phi_3} \subseteq \dots$ , por lo que una misma RDA tiene potencialmente una infinidad de conjuntos  $S_\phi$ , todos anidados y contenidos en el intervalo  $[0, 1]$ , en donde sus nodos pueden tomar valores. De esta manera, definimos el *conjunto de estados* de una red difusa como el conjunto difuso  $S$  representado por la familia

$$\mathbf{S} = \{S_\phi : \phi \in [0, 1]\},$$

cuya función de pertenencia está determinada de forma única (cf. 1-2) por la fórmula

$$S(x_i(t)) = \sup \{\phi : x_i(t) \in S_\phi\}. \quad (4-2)$$

Cada función lógica  $f_i$  en una RBA (cf. 2.1) puede ser difusificada, representándola en términos de combinaciones de los operadores  $\wedge, \vee$  y  $\neg$ , y usando los operadores de Zadeh (cf. Nota 6)  $\min, \max$  y  $1 - x$  respectivamente. De esta manera, dependiendo únicamente de la diversidad de estados iniciales, la misma RDA puede ir de booleana a continua sin que las tablas de verdad de sus funciones crezcan descontroladamente, como es común en previos modelos multivaluados [Wuensche, 1998; Solé *et al.*, 2000; Wittmann y Theis, 2011].

### 4.3. Familias de Atractores

Para analizar el comportamiento dinámico de una RDA, definimos un término llamado la *familia* de un atractor, que intuitivamente representa la forma en la que se da la transición de los estados dentro del atractor. Este concepto captura de alguna manera la dinámica en el espacio de estados de las RDAs, sin importar la  $\phi$  escogida o los cambios en la base, ya que no considera el valor específico de los nodos. Solamente representa en que *dirección* cambian sus estados individuales. Por practicidad, presentaremos los conceptos de espacio de estados, atractor y familia de atractores como conjuntos de información multi-dimensional (cf. 2.2).

#### 4.3.1. Espacio de Estados

Una RDA es una red dinámica cuyo *espacio de estados* puede ser definido como el conjunto

$$\mathcal{X} = \{x_{i_1, i_2, \dots, i_n} : S(i_j) = \phi, j = 1, 2, \dots, n\},$$

donde a cada  $x(t) \in \mathcal{X}$  lo llamamos un *estado global* de la red en el instante  $t$ .

Por lo general no cabe duda cuando estamos hablando sobre el estado individual de un nodo o sobre un estado global de la red por lo que simplemente hablaremos de estados, a menos de que el contexto no lo deje claro.

#### 4.3.2. Atractores

Un *atractor* de período  $p$  es un subconjunto del espacio de estados

$$A^p = \{a_{i_1, i_2, \dots, i_n}^l : i_j \in S_\phi, j = 1, \dots, n, l = 1, \dots, p\} \subseteq \mathcal{X},$$

tal que si  $a \in A^p$ , entonces  $a(t) = a(t+p)$  para cada  $t = 0, 1, \dots$ . Denotamos por  $\mathcal{A} \in \{0, 1\}^{\mathcal{X}}$  al conjunto de todos los atractores de una RDA<sup>1</sup>. Cuando hablemos de un *p-atractor* nos referimos a un conjunto de  $p$  estados  $A^p \in \mathcal{A}$ , para algún  $p$  específico.

---

<sup>1</sup>En la práctica, encontrar este conjunto para una RBA es un problema NP-completo [Tosić, 2005; Zhao, 2005], por lo que también lo será para una RDA.

### 4.3.3. Familias

Sea  $\mathcal{T} = \{\rho_{i_1, i_2, \dots, i_n} : i_j \in \tau, j = 1, \dots, n\}$  donde  $\tau = \{-1, 0, 1\}$ . Decimos que  $\Phi$  es una familia de  $p$ -atractores si es un conjunto formado por elementos del espacio de transiciones  $\mathcal{T}$ , es decir,

$$\Phi = \{\rho^\ell \in \mathcal{T} : 0 \leq \ell \leq p-1\}.$$

A los elementos de  $\mathcal{T}$  los llamamos *esquemas de transición*. Claramente  $|\mathcal{T}| = 3^n$ .

Dado un atractor  $A_i^p \in \mathcal{A}$ , podemos describirlo en el instante  $t$  de forma vectorial como

$$\begin{aligned} A_i^p(t) = & [(x_{i_1}(t), x_{i_2}(t), \dots, x_{i_n}(t)), \\ & (x_{i_1}(t+1), x_{i_2}(t+1), \dots, x_{i_n}(t+1)), \dots, \\ & (x_{i_1}(t+p-1), x_{i_2}(t+p-1), \dots, x_{i_n}(t+p-1))], \end{aligned} \quad (4-3)$$

donde  $x_{i_j}(t+\ell) \in S_\phi$  para alguna  $\phi \in [0, 1]$ , con  $j = 1, \dots, n$ , y  $\ell = 0, 1, \dots, p-1$ .

Decimos que el atractor  $A_i^p$  pertenece a la familia  $\Phi$ , si el valor de cada índice del esquema de transición  $\rho^\ell \in \Phi$  con  $0 \leq \ell \leq p-1$  satisface que

$$i_j = \begin{cases} 1, & \text{si } x_{i_j}(t) < x_{i_j}(t+1), \\ 0, & \text{si } x_{i_j}(t) = x_{i_j}(t+1), \\ -1, & \text{si } x_{i_j}(t) > x_{i_j}(t+1). \end{cases} \quad (4-4)$$

Varios atractores pueden pertenecer a la misma familia. En particular todos los 1-atractores pertenecen a la familia  $\Phi = \{\rho_{i_1, i_2, \dots, i_n} : i_j = 0, j = 1, \dots, n\} = \{\rho_{00\dots 0}\}$ . La figura 4-1 ilustra el concepto.

**Ejemplo 33** Sea  $\mathcal{R}$  una RDA con  $n = 3$ ,  $k = 2$  cuya dinámica está dada por el sistema de ecuaciones lógicas

$$\begin{cases} f_1(t+1) = f_1(x_{1_1}(t), x_{1_2}(t)), \\ f_2(t+1) = f_2(x_{2_1}(t), x_{2_2}(t)), \\ f_3(t+1) = f_3(x_{3_1}(t), x_{3_2}(t)), \end{cases} \quad (4-5)$$

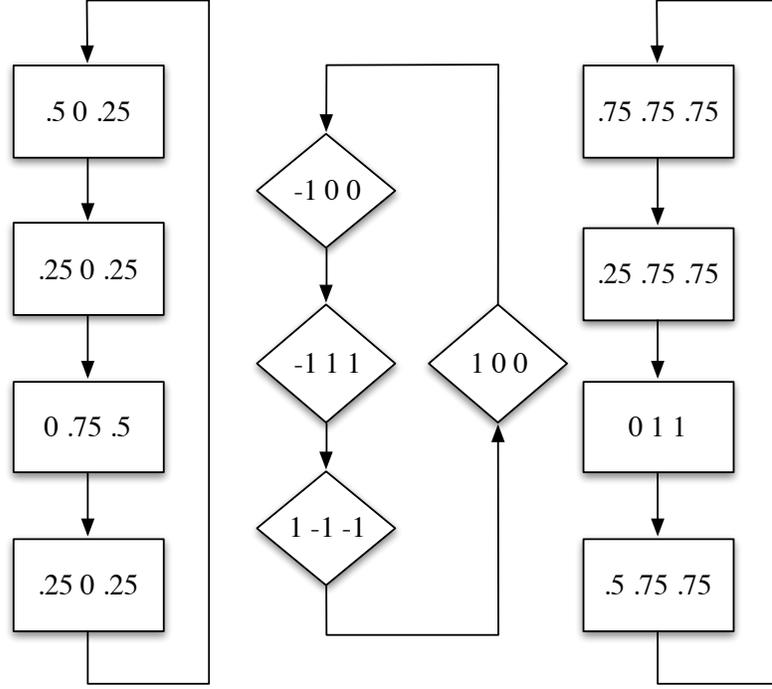


Figura 4-1: Dos 4-atractores de una RDA de  $n = 3$  con base 5 que pertenecen a la misma familia (e.g.,  $\Phi = \{\rho_{100}^0, \rho_{-100}^1, \rho_{-111}^2, \rho_{1-1-1}^3\}$ ). Los estados están representados por rectángulos mientras que los esquemas de transición los representamos con rombos.

Si  $\phi = 0$  como  $S_0 = \{0, 1\}$  entonces  $\mathcal{R}$  tiene base 2, por lo que es una RBA clásica y su espacio de estados es  $\mathcal{X} = \{x_{i_1, i_2, i_3} : S(i_j) = 0, j = 1, 2, 3\}$  donde  $|\mathcal{X}| = 2^3$ . Dadas las funciones  $f_i : S_\phi^2 \rightarrow S_\phi$  con  $i = 1, 2, 3$ , definidas como

$$\begin{aligned}
 f_1(x_2, x_3) &= \bar{x}_2 \vee x_3, \\
 f_2(x_1, x_2) &= x_1 \wedge x_2, \\
 f_3(x_3, x_2) &= \overline{x_3 \vee x_2},
 \end{aligned}
 \tag{4-6}$$

la estructura de  $\mathcal{R}$  queda completamente determinada.

Cambiando la elección del parámetro  $\phi = 0.6$ , la base de la red incrementa ( $\mathbf{b} = 4$ ), y consecuentemente el espacio  $\mathcal{X} = \{x_{i_1, i_2, i_3} : S(i_j) = 0.6, j = 1, 2, 3\}$  pasa a estar conformado por 64 estados. Cada nodo toma ahora valores sobre el conjunto  $S_\phi = \{0, 0.25, 0.75, 1\}$ , y las

funciones lógicas (4-6) son difusificadas de la siguiente manera:

$$\begin{aligned}
 f_1(x_2, x_3) &= \text{máx}\{1 - x_2, x_3\}, \\
 f_2(x_1, x_2) &= \text{mín}\{x_1, x_2\}, \\
 f_3(x_3, x_2) &= 1 - \text{máx}\{x_3, x_2\}.
 \end{aligned}
 \tag{4-7}$$

Es claro que para toda RDA el espacio de transiciones  $\mathcal{T}$  es invariante a los cambios de  $\phi$ , es decir,  $|\mathcal{T}| = 3^n$  para cada  $\phi \in [0, 1]$ . En el caso límite, cuando  $\phi = 1$ , decimos que la red difusa  $\mathcal{R}$  es una red *continua* ya que su conjunto (difuso) de estados  $S$  está conformado por una cantidad infinita de conjuntos clásicos  $S_\phi$ , uno por cada  $\phi$  que podamos escoger en el intervalo  $[0, 1]$ . Como  $|[0, 1]| = 2^\omega$  y es bien sabido en la teoría clásica de conjuntos que el cardinal del continuo  $\mathfrak{c} = 2^\omega$ , entonces podemos decir que  $|S| = \mathfrak{c}$  y por lo tanto que  $\mathcal{R}$  es una red propiamente continua.

El concepto de familia de atractores nos ayuda a estudiar cuanta información puede ser comprimida en una RDA mientras variamos su parámetro  $\phi$ . Decimos que los estados  $x, y \in \mathcal{X}$  tienen la misma familia  $\Phi$ , si los atractores a los que pertenecen  $A_x, A_y \in \mathcal{A}$  respectivamente, son elementos de la misma familia  $\Phi$ . Dados cualesquiera dos atractores  $A_1, A_2 \subseteq \mathcal{X}$ , definimos la relación  $A_1 \sim_\Phi A_2 \Leftrightarrow A_1, A_2 \in \Phi$ , donde  $\Phi$  es una familia de atractores. Es claro que  $\sim_\Phi$  es reflexiva, transitiva y simétrica, por lo que cada  $\Phi$  induce una partición del espacio de estados  $\mathcal{X}$  constituida por las clases de equivalencia  $[x] = \{y \in \mathcal{X} : A_x \sim_\Phi A_y\}$

De esta forma, podemos definir la *compresibilidad de información* de una RDA como una medida inversamente proporcional al número de familias en las que el espacio de estados  $\mathcal{X}$  es dividido. Así, podemos analizar cuanta información se pierde (i.e. no llega a ser comprimida) al cambiar la base del modelo utilizado.

## Capítulo 5

# Simulaciones y Resultados

Para analizar el comportamiento dinámico de las RDAs, hicimos simulaciones en el laboratorio virtual de código abierto RBNLab [Gershenson, 2005]. Utilizamos ensambles de 50 redes para cada valor diferente de  $k$ . Escogimos cinco valores diferentes para el parámetro  $0 \leq \phi \leq 1$  con la intención de que la base de la red ( $\mathfrak{b}$ ) tomara los valores 2, 4, 8, 16 y  $\omega$  respectivamente.

Tres diferentes tamaños de redes fueron escogidos ( $n = 10, 20, 100$ ). Para detectar atractores en el espacio de estados de una red, computacionalmente se puede hacer una búsqueda exhaustiva sobre dicho espacio, con la limitación de restringir el tamaño de la red ( $n \leq 10$ ). Para redes más grandes, es difícil tomar este enfoque debido a que el tamaño del espacio de estados ( $2^n$  para el caso Booleano) crece exponencialmente con  $n$ . Lo que se hace en este caso es restringir el número de posibles estados iniciales, esperando que estos nos ayuden a realizar una “buena” búsqueda, pero esto puede llevar a que sólo una pequeña parte del espacio de estados sea explorada.

### 5.1. Algoritmos

Una atractor puede ser visto como una parte del espacio de estados donde es muy probable que la dinámica de la red quede atrapada. Diferentes condiciones iniciales pueden converger a un mismo atractor. El algoritmo que utilizamos para encontrarlos es el siguiente: Dejamos correr a la red a partir de un estado inicial para que llegue a un atractor. Más de 10000 pasos son observados. Después, si el estado en el que estamos es parte de un atractor encontrado

previamente, reiniciamos la búsqueda para un estado inicial diferente. Si no, creamos un nuevo atractor formado por ese estado, y los estados obtenidos siguiendo la dinámica de la red son añadidos al atractor, omitiendo los estados que ya fueron visitados. Esta búsqueda de estados en un atractor se lleva a cabo hasta que se alcanza un período máximo (se usaron 3000 iteraciones). Cuando termina este proceso, se escoge otro estado inicial y el algoritmo se vuelve a correr.

La forma de explorar el espacio de estados, para  $\mathbf{b} = 2$  es generando aleatoriamente un estado inicial (Booleano), mientras que para  $\mathbf{b} > 2$  difusificamos los estados individuales de este estado de la siguiente manera: Escogemos aleatoriamente un número  $r_{i_0} \in \{0, 1, \dots, \lceil \mathbf{b}/2 \rceil - 1\}$  con  $i = 1, \dots, n$ , y si es el  $i$ -ésimo estado individual es 1, definimos el número  $r_{i_1} = 1 + r_{i_0}$ . Entonces, sustituimos cada estado individual por el valor de  $r_{i_j}/(\mathbf{b} - 1)$ , donde  $j$  es el valor original del  $i$ -ésimo estado individual Booleano. De esta forma, considerando 2048 estados iniciales, consideramos que se hace una buena exploración del espacio de estados.

## 5.2. Atractores

Analizamos el comportamiento dinámico de 50 redes para cada juego de parámetros ( $\phi$ ,  $k$ ). En las figuras 5-1, 5-2 y 5-3 se muestran los resultados obtenidos en nuestras simulaciones. Estas muestran la cantidad promedio de atractores  $\langle att \rangle$  encontrados en RDAs con  $n = 10$ ,  $n = 20$  y  $n = 100$  respectivamente, para diferentes valores de  $k \in \{1, \dots, 5\}$ , mientras variamos el parámetro  $\phi$  desde 0 hasta 1. Los resultados están graficados en escala logarítmica.

Como el tamaño del espacio de estados ( $2^n$ ) crece rápidamente al variar el tamaño de la red ( $n$ ) hacer estudios estadísticos completos es posible sólo para redes muy pequeñas. Por ejemplo, para  $n = 20$  hay más de un millón de estados iniciales. Por lo tanto, estos estudios se concentran en redes pequeñas, o toman en cuenta solamente algunos cuantos estados iniciales. En el primer caso, algunas propiedades de redes grandes no serán observadas, mientras que en el segundo, algunos atractores no ser encontrados, especialmente si su cuenca de atracción consiste en muy pocos estados.

Sin embargo, incluso cuando pudiera haber potencialmente muchos más atractores de los que fueron encontrados limitando nuestra muestra, creemos que en la práctica se obtendrían los mismos resultados, ya que la naturaleza no agota todas las posibles configuraciones en las

que podría evolucionar (un argumento de esto ha sido planteado por Kauffman cf. 3.2.2). Es decir, podrían haber más tipos celulares para un número dado de genes, pero *desarrollarse* en estos es imposible debido a que su cuenca e atracción es muy pequeña.

Para redes con  $n = 10$  no observamos ningún comportamiento que pueda ser considerado altamente relevante. Para el caso Booleano, los datos obtenidos son congruentes con estudios previos [Gershenson, 2002] sobre la cantidad y longitud promedio de atractores  $\langle \ell(a) \rangle$ . Por lo tanto, no consideramos significativas las estadísticas sobre ensambles de redes de dicho tamaño. Para redes de tamaño  $n = 100$ , muy probablemente no conseguimos realizar una buena exploración del espacio de estados ya que su tamaño crece descontroladamente ( $2^{100} > 10^{30}$  para el caso Booleano), por lo que tampoco consideramos significantes las estadísticas sobre ensambles de redes de 100 nodos.

Obviamente al aumentar la base de una RDA, y consecuentemente la cantidad de estados posibles en los que puede evolucionar el sistema, el número de atractores potenciales (subconjuntos de estados) crece considerablemente, por lo que los resultados mostrados son consistentes con nuestra intuición.

### 5.3. Familias

Claramente la cantidad de estados crece exponencialmente al incrementar la base de una RDA. El espacio de estados ( $\mathcal{X}$ ) de una red difusa es un conjunto dinámico que varía dependiendo de  $\phi$  y que tiene cardinalidad  $\mathfrak{b}(\phi)^n$ . Diferentemente, el espacio de transiciones  $\mathcal{S}$ , por la forma en la que fue definido se mantiene constante para cualquier  $\phi \in [0, 1]$ , por lo tanto el número de familias posibles también queda fijo para todos los conjuntos  $S_\phi$  contenidos en el conjunto de estados  $S$ .

De esta manera, clasificamos a los atractores en el espacio de estados  $\mathcal{X}$  de acuerdo con la forma en que se dan las transiciones entre los estados dentro de los ellos. Así, podemos agrupar con el concepto de familia, conjuntos de atractores que tienen un comportamiento dinámico similar, y analizar cuanta información dinámica se pierde en el proceso de discretización<sup>1</sup> o *reducción en la complejidad* del sistema.

---

<sup>1</sup> Entendemos por discretización el proceso contrario a la difusificación del conjunto de estados individuales, siendo el caso Booleano el extremo de este proceso.

Las figuras 5-4, 5-5 y 5-6 muestran la cantidad promedio de familias  $\langle fam \rangle$  detectadas en los atractores encontrados de acuerdo a la descripción de la sección anterior. Aunque la cantidad de atractores ( $\langle att \rangle$ ) crece exponencialmente con la base de la red ( $b$ ), el número de familias de atractores ( $\langle fam \rangle$ ) en un principio también crece rápidamente, pero a continuación se mantiene casi constante para incrementos adicionales en la base. Utilizamos escala logaítmica para detectar el posible crecimiento exponencial.

Estos resultados sugieren que, como es bien sabido, los modelos Booleanos desempeñan una gran reducción en la complejidad de un sistema, pero esto no está relacionado con una pérdida de información, como es usualmente pensado. De esta forma, el presente trabajo sirve como soporte para la justificación del uso de redes booleanas aleatorias en la simulación y modelación de sistemas sintéticos vivientes.

## 5.4. Dinámica

Utilizamos los mismos algoritmos para calcular la longitud promedio de los atractores  $\langle \ell(a) \rangle$  y la longitud promedio de las familias  $\langle \ell(\rho) \rangle$  que encontramos en un ensamble de 100 redes con  $n = 20$  para diferentes valores de  $\phi$  y  $k$ . Los resultados de este análisis para atractores y familias están representados en las figuras 5-7 y 5-8, respectivamente. Estos también están graficados en escala logarítmica.

Observamos, en algunos casos, un decremento tanto en  $\langle \ell(a) \rangle$  como en  $\langle \ell(\rho) \rangle$ . Lo que esto significa es que los *nuevos* atractores encontrados al incrementar la base, junto con sus familias asociadas, son puntuales o de período 1.

Las desviaciones estándar se muestran sólo para el caso crítico  $k = 2$ . Para  $\langle att \rangle$ , las desviaciones estándar más grandes las encontramos en ensambles continuos ( $\phi = 1$ ), mientras que para  $\langle fam \rangle$  este caso se dio cuando  $k = 5$ .

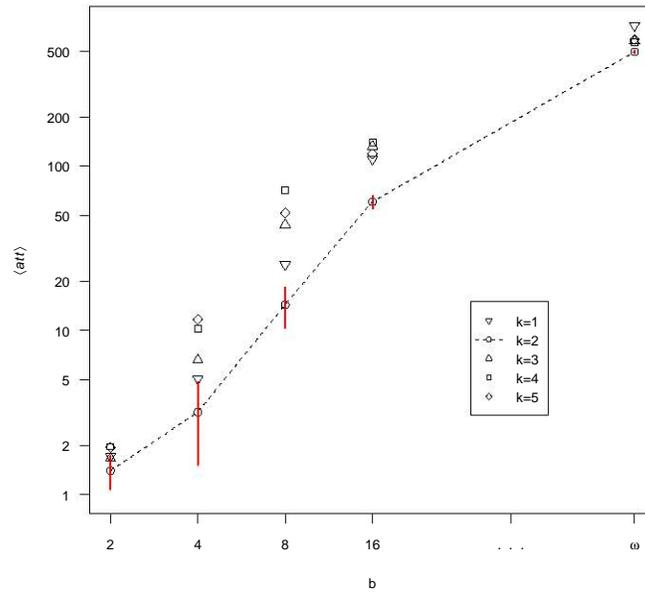


Figura 5-1: Número promedio de atractores  $\langle att \rangle$  para  $n = 10$  (50 redes).

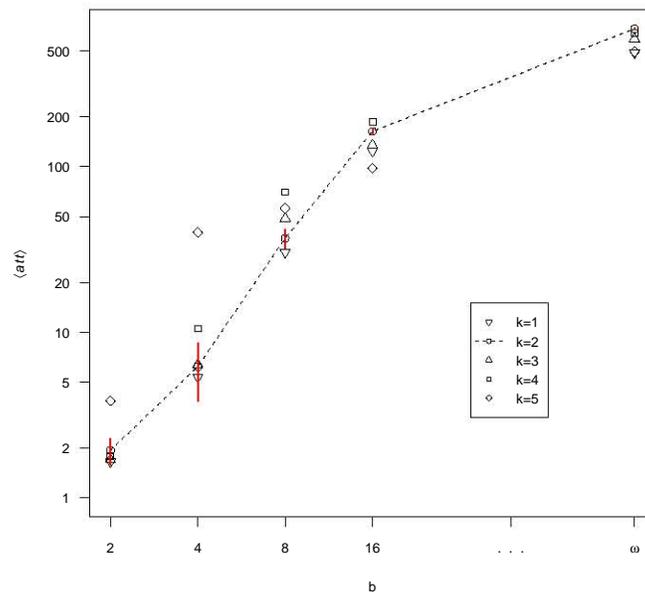


Figura 5-2: Número promedio de atractores  $\langle att \rangle$  para  $n = 20$  (50 redes).

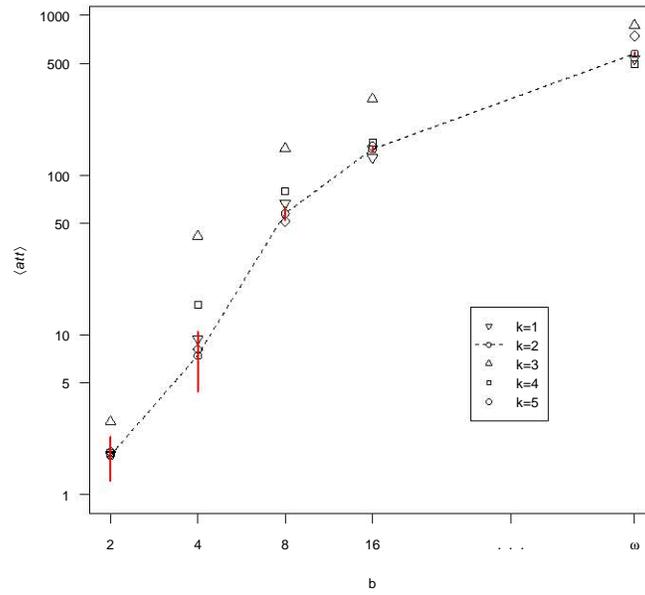


Figura 5-3: Número promedio de atractores  $\langle att \rangle$  para  $n = 100$  (50 redes).

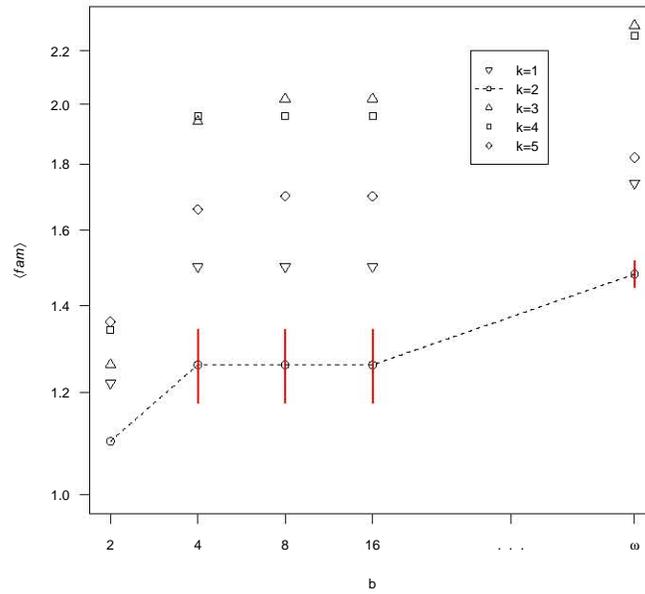


Figura 5-4: Número promedio de familias  $\langle fam \rangle$  para  $n = 10$  (50 redes).

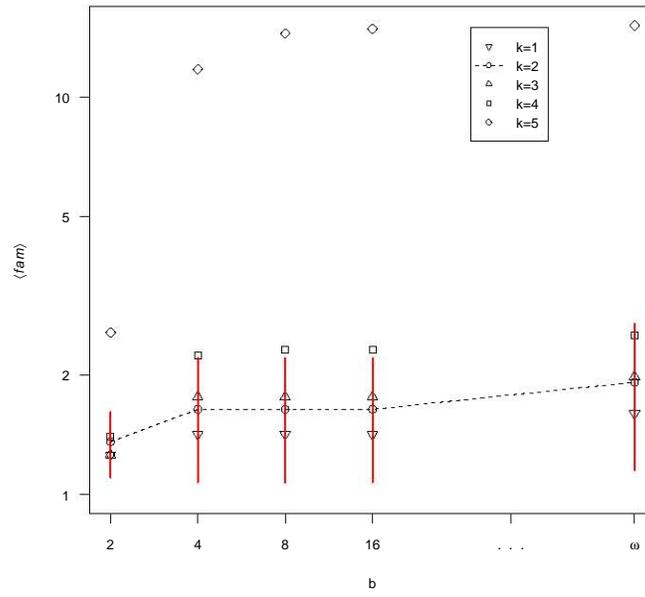


Figura 5-5: Número promedio de familias  $\langle fam \rangle$  para  $n = 20$  (50 redes).

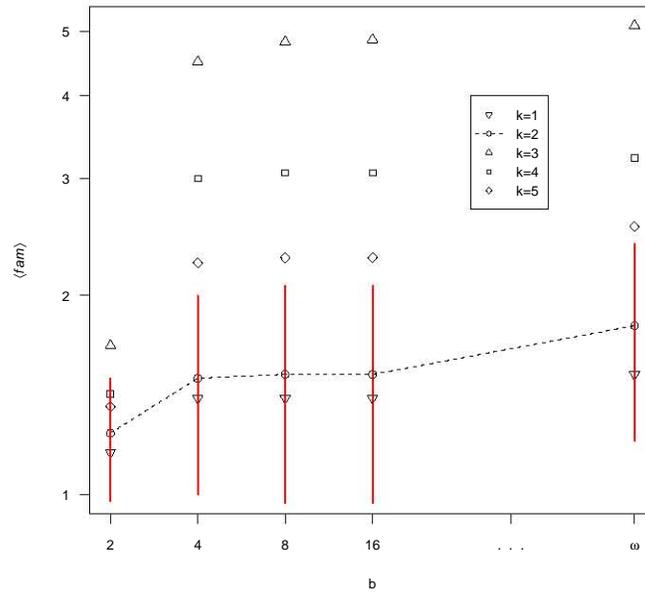


Figura 5-6: Número promedio de familias  $\langle fam \rangle$  para  $n = 100$  (50 redes).

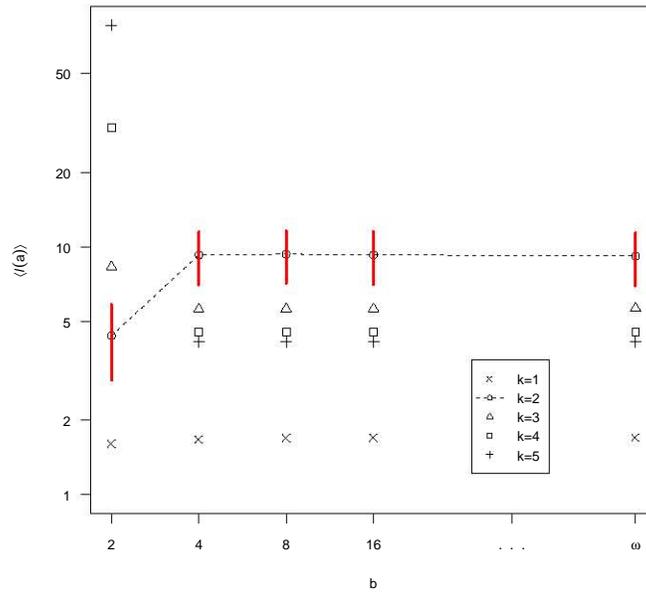


Figura 5-7: Longitud promedio de los atractores encontrados  $\langle \ell(a) \rangle$  para  $n = 20$  (100 redes).

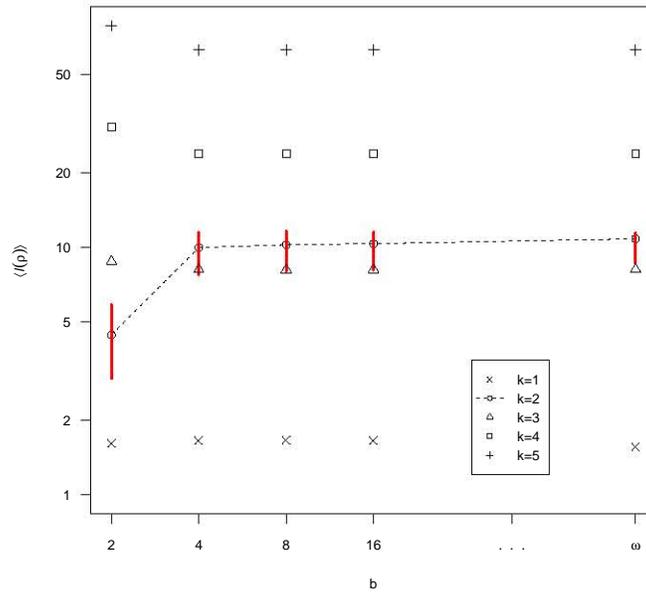


Figura 5-8: Longitud promedio de las familias encontradas  $\langle \ell(\rho) \rangle$  para  $n = 20$  (100 redes).

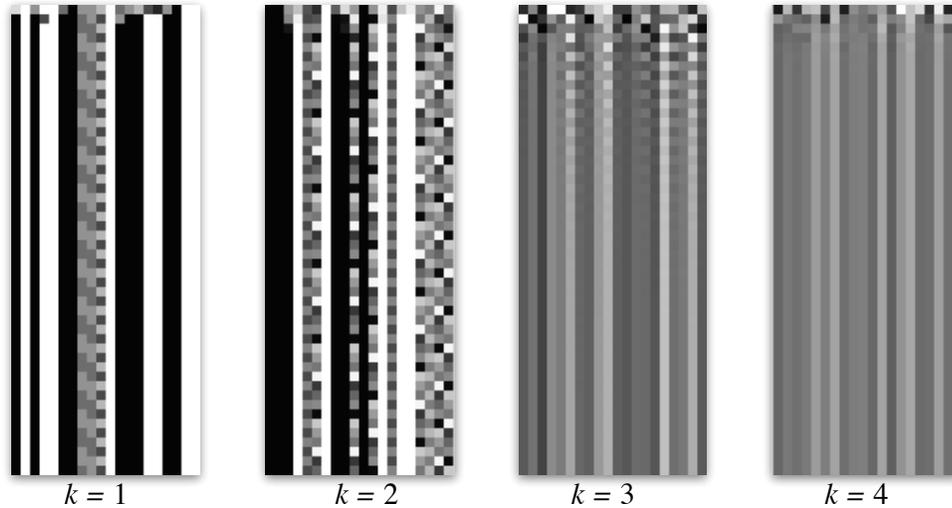


Figura 5-9: Evolución temporal (50 pasos) de cuatro RDAs con  $n = 20$  y base  $\mathfrak{b} = 8$  para diferentes valores de  $k$ . El color de cada *bit* representa un grado en la escala  $[0, 1]$ , donde *negro* representa 0 y *blanco* representa 1. Se pueden observar los diferentes regímenes dinámicos dependiendo de la  $k$  escogida. Las columnas representan las trayectorias de cada nodo en el espacio de estados. Las filas representan el estado global de la red en el instante  $t$ . La dinámica fluye hacia abajo.

En la figura 5-9 podemos observar como al tener toda una escala de posibilidades, cada nodo tiene potencialmente una infinidad de atractores (e.g. cuando  $\phi = 1$  y  $\mathfrak{b} = \omega$ ) por lo que es más *fácil* alcanzar uno. De esta forma, podemos apreciar los diferentes regímenes dinámicos que encontramos en redes difusas con  $n = 20$  y  $\phi = 0.84$  para diferentes valores de  $k$  con  $t = 0, 1, \dots, 49$ .

Como observamos en las figuras anteriores, el número promedio de familias  $\langle fam \rangle$ , crece al parecer de forma lineal con la base  $\mathfrak{b}$  de la red. Es decir, no encontramos un incremento significativo en  $\langle fam \rangle$  cuando la representación de la información va de booleana a continua. Este resultado sugiere que en el proceso de discretización se pierde sólo una cantidad pequeña de información (familias), y entonces podemos afirmar que las redes difusas son capaces de lograr un grado significativo de reducción en la complejidad –a través de su parámetro  $\phi$ . Concluimos diciendo que el uso de modelos booleanos para el estudio de la dinámica discreta en observaciones reales queda, por lo tanto, más justificado.

## Capítulo 6

# Conclusiones

Propusimos un modelo que extiende a las redes booleanas aleatorias clásicas y utilizando lógica difusa logramos representar qué tan continua (o booleana) puede ser una red dinámica generada por nuestro modelo. Cuantificamos el grado de complejidad de una red con el parámetro  $\phi$  en la escala  $[0, 1]$ . Definimos el concepto de familias de atractores para analizar el comportamiento dinámico de redes difusas aleatorias.

Realizando simulaciones en el laboratorio virtual RBNLab [[Gershenson, 2005](#)] no detectamos un cambio significativo en el número promedio de familias de atractores, al pasar las redes de booleanas a continuas, lo cual sugiere que no hay una correlación fuerte al reducir la complejidad de estos modelos con la compresibilidad de información dinámica.

Estos resultados justifican la *idealización booleana* en el estudio de sistemas complejos que presentan un comportamiento semejante al observado en organismos vivientes, ya que apoya la hipótesis de que la interacción de componentes relativamente simples puede consecuentemente producir la emergencia de patrones altamente complejos imposibles de predecir mediante la observación detallada de los niveles individuales del sistema.

Un interesante trabajo futuro podría consistir en obtener resultados analíticos y experimentales acerca de las transiciones de fase, así como la cuantificación de las medidas de información, emergencia, autoorganización y homeostasis propuestas en [Gershenson y Fernandez \[2012\]](#), para analizar de forma teórica la capacidad para comprimir información y, de esta forma, la complejidad del modelo difuso propuesto en el presente trabajo que aquí en este momento termina.

# Bibliografía

- ALDANA, M., COPPERSMITH, S., Y KADANOFF, L.P. Boolean dynamics with random couplings. En E. Kaplan, J.E. Marsden, y K.R. Sreenivasan (editores), *Perspectives and Problems in Nonlinear Science. A Celebratory Volume in Honor of Lawrence Sirovich*, págs. 23–89. Springer Applied Mathematical Sciences Series, New York (2003)
- BAGLEY, R.J. Y GLASS, L. Counting and classifying attractors in high dimensional dynamical systems. *Journal of Theoretical Biology* **183**:269–284 (1996)
- BAILLY, F. Y LONGO, G. Randomness and determinism in the interplay between the continuum and the discrete. *Mathematical Structures in Computer Science* **17**:289–305 (2007)
- BELL, J.L. Y SLOMSON, A.B. *Models and ultraproducts: an introduction*. North-Holland (1974)
- BELOHLAVEK, R. Y KLIR, G.J. *Concepts and fuzzy logic*. The MIT Press (2011)
- BUIATTI, M. Y LONGO, G. Randomness and multi-level interactions in biology. *Theory of Biosciences* **132**:139–158 (2013)
- CHAITIN, G. On the length of programs for computing finite binary sequences: Statistical considerations. *Journal of the ACM* págs. 145–159 (1969)
- CHENG, Z., HONGSHENG, Q., Y ZHIQIANG, L. *Analysis and control of Boolean networks: a semi-tensor approach*. Communications and Control Engineering. Springer-Verlag London Limited (2011)
- DERRIDA, B. Y POMEAU, Y. Random networks of automata: A simple annealed approximation. *Europhysics Letters* **1 (2)**:45–49 (1986)

- FEYNMAN, R.P. Simulating physics with computers. *International Journal of Theoretical Physics* **21 (6-7)**:467–488 (1982)
- FREGE, G. Über Sinn und Bedeutung (On sense and reference). *Zeitung für Philosophie und philosophische Kritik* **10**:25–50 (1892)
- GERSHENSON, C. Classification of random Boolean networks. *Artificial Life VIII: Proceedings of the Eight International Conference on Artificial Life* págs. 1–8 (2002)
- GERSHENSON, C. Introduction to random Boolean networks. En M. Bedau, P. Husbands, T. Hutton, S. Kumar, y H. Suzuki (editores), *Workshop and Tutorial Proceedings, Ninth International Conference on the Simulation and Synthesis of Living Systems (ALife IX)*, págs. 160–173. Boston, MA (2004a)
- GERSHENSON, C. Phase transitions in random Boolean networks with different updating schemes. *Submitted to Physica D* (2004b)
- GERSHENSON, C. Updating schemes in random Boolean networks: Do they really matter? *Artificial Life IX Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems* págs. 238–243 (2004c)
- GERSHENSON, C. *RBNLab*. <http://rbn.sourceforge.net> (2005)
- GERSHENSON, C. Guiding the self-organization of random Boolean networks. *Theory in Bioscience* págs. 181–191 (2012)
- GERSHENSON, C., CSERMELY, P., ÉRDI, P., KNYAZEVA, H., Y LASZLO, A. The past, present and future of cybernetics and systems research. *Systema: connecting matter, life, culture and technology* **1 (3)**:4–13 (2014)
- GERSHENSON, C. Computing networks: A general framework to contrast neural and swarm architectures. *CoRR* **abs/1001.5244** (2010)
- GERSHENSON, C. Y FERNANDEZ, N. Complexity and Information: Measuring Emergence, Self-organization, and Homeostasis at Multiple Scales. *CoRR* **abs/1205.2026** (2012)

- HARVEY, I. Y BOSSOMAIER, T. Time out of joint: attractors in asynchronous random Boolean networks. *Proceedings of the Fourth European Conference on Artificial Life (ECAL97)* págs. 65–75 (1997)
- KAPPLER, K., EDWARDS, R., Y GLASS, L. Dynamics in high-dimensional model gene networks. *Signal Processing* **83**:789–798 (2003)
- KARLEBACH, G. Inferring Boolean network states from partial information. *EURASIP Journal on Bioinformatics and Systems Biology* **2013** (11) (2013)
- KAUFFMAN, S.A. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology* **22**:437–467 (1969)
- KITANO, H. Systems biology: a brief overview. *Science* **195**:1662–1664 (2002)
- KOLMOGOROV, A. Three approaches to the quantitative definition of information. *Problemy Peredachi Informatsii (Problems of Information Transmission)* **1**:3–11 (1965)
- MENG, M. Y FENG, J. A matrix approach to hypergraph stable set and coloring problems with its application to storing problem. *Journal of Applied Mathematics* (2014)
- NICHITIU, C. Y RÉMILA, E. Simulations of graph automata. En *Proc. MFCS'98 Satellite Workshop on Cellular Automata*, págs. 330–342 (1998)
- POBLANNO-BALP, R. Y GERSHENSON, C. Modular random Boolean networks. *Artificial Life* **17** 4:331–351 (2011)
- ROSENBLUETH, A. Y WIENER, N. The role of models in science. *Philosophy of Science* **12** (4):316–321 (1945)
- SHMULEVICH, I. Y ZHANG, W. Binary analysis and optimization-based normalization of gene expression data. *Bioinformatics* **18** (4):555–565 (2002)
- SOLÉ, R.V., LUQUE, B., Y KAUFFMAN, S. *Phase transition in random networks with multiple states*. Technical Report 00-02-011, Santa Fe Institute (2000)
- SØRENSEN, M.H. Y URYCZYN, P. *Lectures on the Curry-Howard isomorphism, Studies in logic and the foundations of mathematics*, tomo 149. Elsevier (2006)

- TOSIĆ, P.T. *Counting Fixed Points and Gardens of Eden of Sequential Dynamical Systems on Planar Bipartite Graphs*. Electronic Colloquium on Computational Complexity Report No. 91 (2005)
- TURUNEN, E. Algebraic structures in fuzzy logic. *Fuzzy Sets and Systems* **52**:181–188 (1992)
- ČASLAV, B. Quantum causality. *Nature Physics* **10**:259–263 (2014)
- WITTMANN, D.M. Y THEIS, F.J. Dynamic regimes of random fuzzy logic networks. *New Journal of Physics* **13** (2011)
- WITTMANN, D.M., KRUMSIEK, J., SAEZ-RODRIGUEZ, J., LAUFFENBURGER, D.A., KLAMT, S., Y THEIS, F.J. Transforming Boolean models to continuous models: methodology and application to t-cell receptor signaling. *BMC Systems Biology* **3** (2009)
- WOLFRAM, S. *A new kind of science*. Wolfram Media (2002)
- WUENSCHÉ, A. Discrete dynamical networks and their attractor basins. En R. Standish, B. Henry, S. Watt, R. Marks, R. Stocker, D. Green, S. Keen, y T. Bossomaier (editores), *Complex Systems '98*, págs. 3–21. University of New South Wales, Sydney, Australia (1998)
- WUENSCHÉ, A. Discrete dynamics lab: Tools for investigating cellular automata and discrete dynamical networks. En A. Adamatzky y M. Komosinski (editores), *Artificial Life Models in Software*, págs. 263–297. Springer London (2008)
- YINGJUN, C., WANG, P.P., Y RESCONI, G. Reverse engineering of the nk boolean network and its extension – fuzzy logic network. *New Mathematics and Natural Computation* **3** (1):69–87 (2007)
- ZADEH, L.A. Fuzzy sets. *Information and Control* **8**:338–353 (1965)
- ZHAO, Q. A remark on “scalar equations for synchronous Boolean networks with biological applications” by C.Farrow, J.Heidel, J.Maloney, and J.Rogers. *IEEE Transactions On Neural Networks* **16**(6):1715–1716 (2005)
- ZHAO, Y., GAO, X., Y CHENG, D. Semi-tensor product approach to Boolean functions. *Preprint* (2014)