

# Complejidad, redes neuronales artificiales y simulación computacional en la investigación científica.

Autor:

**Rubio, Erick Manuel**

Tutor:

**Barberis, Sergio Daniel**

**2022**

Tesis presentada con el fin de cumplimentar con los requisitos finales para la obtención del título Doctor de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires en Filosofía.

Posgrado



**FILO:UBA**  
Facultad de Filosofía y Letras

Universidad Nacional de Buenos Aires

Facultad de Filosofía y Letras

Tesis Doctoral

***Complejidad, redes neuronales artificiales y simulación  
computacional en la investigación científica***

Autor: Erick Manuel Rubio

Director: Dr. Sergio Daniel Barberis

Codirector: Dr. Hernán Miguel

Julio de 2022

El presente trabajo ha sido realizado con una Beca en el marco del proyecto “Desarrollo del Programa de Formación de Alto Nivel para el Departamento del Atlántico” financiado con recursos del Fondo de Ciencias, Tecnología e Innovación del Sistema General de Regalía (2014); asimismo, con un Apoyo Institucional por parte de la Universidad del Atlántico para cursar estudios de postgrado (2014-1).

## AGRADECIMIENTOS

Muy especialmente a Sergio Barberis y a Hernán Miguel, principalmente por permitirme hacer parte de una labor creativa conjunta. Además de la enorme paciencia y comprensión que sostuvieron durante dicha labor.

A Olimpia Lombardi, Guillermo Folguera y Sebastián Fortín. Sin duda alguna sus valiosos aportes iniciales representaron las condiciones de posibilidad para el desarrollo de este trabajo.

Al grupo de '*Sistemas complejos e IA*' (SADAF). Resulta difícil pensar en un mejor escenario posible donde haber presentado mi motivación de conectar la complejidad, las redes neuronales artificiales y las simulaciones computacionales.

Al grupo '*Holosapiens*' de la Universidad del Atlántico. El apoyo brindado en todas las dimensiones fue sin duda invaluable.

A la Universidad del Atlántico y a la Universidad de Buenos Aires, por la defensa de lo público.

A mi madre, Alicia Esther Rubio, a quien le debo absolutamente todo y mucho más.

Por último, pero no menos importante, a mis queridos amigos por el apoyo y la confianza inmensa que me han brindado a lo largo de los años.

# DEDICATORIA

*A mi amigo y mentor David L. Dahmen*

## Contenido

<b>INTRODUCCIÓN</b> .....	6
<b>CAPÍTULO I</b> .....	19
<b>ALGUNOS RASGOS DISTINTIVOS DE LOS SISTEMAS COMPLEJOS</b> .....	19
<i>¿Es posible una teoría unificada de la complejidad?</i> .....	19
I.1. <i>Emergencia</i> .....	20
I.2. <i>Comportamientos colectivos</i> .....	33
I.3. <i>Autoorganización</i> .....	39
I.4. <i>Adaptación</i> .....	44
I.5. <i>Estructura jerárquica</i> .....	47
I.6. <i>Feedback (retroalimentación)</i> .....	53
<b>CAPÍTULO II</b> .....	59
<b>APROXIMACIÓN TEÓRICA A LOS SISTEMAS COMPLEJOS</b> .....	59
<i>¿Es posible una métrica para la complejidad?</i> .....	59
II.1. <i>Dinámica no lineal y caos</i> .....	60
II.2. <i>Evolución</i> .....	76
II.2.1. <i>Los orígenes de la idea de evolución</i> .....	76
II.2.2. <i>Selección natural, adaptación y gradualismo</i> .....	78
II.2.3. <i>Teoría sintética de la evolución</i> .....	80
II.2.4. <i>¿Cambio o direccionalidad? Progreso panglosiano y el reto de Hull</i> .....	82
II.3. <i>Estadística y probabilidad</i> .....	85
II.3.1. <i>Aleatoriedad y complejidad algorítmica</i> .....	85
II.3.2. <i>Tres tipos de azar</i> .....	88
II.3.3. <i>Orden estadístico: teorema del límite central</i> .....	94
II. 4. <i>Complejidad modelo paramétrica (C *)</i> .....	98
II.4.1. <i>Propiedad emergente modelo paramétrica (E *)</i> .....	103
<b>CAPÍTULO III</b> .....	108
<b>REDES NEURONALES ARTIFICIALES Y EL PROBLEMA DEL PLEGAMIENTO DE LAS PROTEÍNAS</b> .....	108
<i>¿Se está produciendo un giro hacia la predicción en ciencia?</i> .....	108
III.1. <i>Paradigmas de aprendizaje dentro del campo del 'Machine learning'</i> .....	109
III.1.1. <i>Aprendizaje supervisado</i> .....	110
III.1.2 <i>Aprendizaje no supervisado</i> .....	111

III.1.3 <i>Aprendizaje reforzado</i> .....	113
III.2 <i>Redes neuronales artificiales de aprendizaje profundo</i> .....	114
III.2.1 <i>'Backpropagation' y 'gradient descent' (Retropropagación del error y el descenso del gradiente)</i> .....	125
III.3 <i>El problema del plegamiento de las proteínas y su asombrosa resolución por parte de AlphaFold</i> .....	134
III.4 <i>¿Se está produciendo un giro en ciencias hacia la predicción?</i> .....	151
<b>CAPÍTULO IV</b> .....	154
<b>ALGUNOS RETOS EPISTEMOLÓGICOS QUE REPRESENTA EL ESTUDIO DE SISTEMAS COMPLEJOS Y SU RELACIÓN CON LOS MODELOS Y SIMULACIONES COMPUTACIONALES</b> .....	154
<i>¿Cuál es la novedad epistémica de las simulaciones computacionales?</i> .....	154
IV.1 <i>¿Qué se suele entender por simulación computacional?</i> .....	155
VI.1.1 <i>Simulación computacional en un sentido estrecho</i> .....	155
IV.1.2 <i>Simulación computacional en un sentido amplio</i> .....	156
IV.1.3 <i>Un punto de vista alternativo sobre las simulaciones computacionales</i> .....	158
IV.2 <i>Métodos de simulación</i> .....	159
IV.2.1 <i>Método de Monte Carlo</i> .....	159
IV.2.2 <i>Simulaciones basadas en ecuaciones</i> .....	162
IV.2.3 <i>Simulaciones basadas en agentes</i> .....	163
IV.2.4 <i>Simulaciones multiescalas</i> .....	164
IV.3 <i>Algunos posibles aspectos epistemológicos novedosos en las simulaciones computacionales</i> .....	165
IV.3.1 <i>Visualización y manipulación</i> .....	174
IV. 3.2 <i>Experimentación</i> .....	181
IV.3.3 <i>Opacidad epistémica y consideraciones prácticas en las simulaciones computacionales</i> .....	185
<b>CONCLUSIONES</b> .....	190

# INTRODUCCIÓN

## Abstract

En este capítulo se presenta el problema filosófico de definir el fenómeno de la complejidad. Se introduce también la estrecha relación entre los estudios de la complejidad y las simulaciones computacionales. Por último, se especifica el plan de la exposición.

En el 2004, en el marco anual de los cursos de verano que se realizan en el prestigioso *Santa Fe Institute* para el estudio de sistemas complejos, Melanie Mitchell, profesora de ciencias de la computación adscrita al mismo, con razón de los veinte años de fundado el Instituto, organizó un panel conmemorativo con varios de sus más prominentes miembros (Jim Crutchfield, Bob Eisenstein, Doyne Farmer, Eric Smith, Alfred Hübler, Stephanie Forrest, entre otros) con el objetivo de brindarle a los asistentes del curso la oportunidad de plantearles a estos reconocidos científicos sus más variadas inquietudes sobre el estudio de los sistemas complejos.

La primera pregunta que realizaron los estudiantes fue, obviamente, “¿cómo se define la complejidad?” La pregunta causó en un principio gracia entre los asistentes. Lo que luego quizá no haya causado gracia a los estudiantes es que, estando en uno de los centros más reconocidos en el estudio de los sistemas complejos, no pudiesen obtener, por tentativa que fuese, una definición de complejidad. Cada panelista ofreció una definición distinta; dándose el caso incluso que algunas iban en contra de la que propusiera algún(a) colega.

“Si la facultad del Instituto Santa Fe –la institución más famosa en el mundo dedicada a la investigación de sistemas complejos– no puede ponerse de acuerdo sobre

lo que se entiende por complejidad, entonces, ¿cómo puede incluso llegar a haber una ciencia de la complejidad?”<sup>1</sup> (Mitchell, 2009, p. 94)

No resulta sorprendente pues, que, a partir de una situación como esta, se reafirme un escenario de sospecha en torno a la noción de complejidad: “¿Será acaso que el término ‘*complejidad*’ es un término con un significado poco relevante en la investigación científica que se sostiene y promueve más bien por un valor de relaciones públicas?” Esta pregunta había sido planteada por un reconocido periodista científico pocos años antes de dicho evento. John Horgan había logrado incomodar a varios investigadores dedicados al estudio de la complejidad cuando, en 1995, publicó un artículo (*‘From complexity to perplexity’*) en la *Scientific American*, donde de manera despectiva llamaba ‘*complejólogos*’ (*complexologists*) a los reputados miembros del Instituto Santa Fe por su insistencia en observar complejidad por donde quiera se mire. Para los complejólogos prácticamente todo caería dentro del dominio de la complejidad. Lo llamativo, quizás incluso sospechoso, sostenía Horgan, es que esta observación la tuviesen sin siquiera haberse puesto de acuerdo en qué es la complejidad; presentando únicamente una gran proclividad al uso de *computadores* como factor cohesivo en sus investigaciones.

En un apartado del mencionado artículo, titulado ‘*The poetry of artificial life*’, Horgan llegó a insinuar que los fondos destinados para el estudio de estos sistemas, bajo la suposición de que la complejidad es un aspecto interesante, podría ser una de las razones por la que se sostenía este término, muy a pesar de que entre los mismos investigadores se habían cuestionado su significatividad.

---

<sup>1</sup> *If the faculty of the Santa Fe Institute— the most famous institution in the world devoted to research on complex systems—could not agree on what was meant by complexity, then how can there even begin to be a science of complexity?* (Mitchell, 2009, p. 94)

“En varios momentos, los investigadores han debatido si la complejidad ha llegado a perder tanto significado que debería abandonarse, pero invariablemente concluyen que el término tiene demasiado valor en las relaciones públicas. Los complejólogos a menudo emplean ‘interesante’ como sinónimo de ‘complejo’. Pero ¿qué agencia gubernamental proporcionaría fondos para la investigación sobre una ‘teoría unificada de cosas interesantes’?”

Los complejólogos pueden no estar de acuerdo sobre lo que están estudiando, pero la mayoría está de acuerdo en cómo deberían estudiarlo: con *computadoras*<sup>2</sup>.<sup>3</sup>  
(Horgan, 1995, p. 106)

La crítica de Horgan puede que haya sido un tanto radical. Horgan consideraba en su escrito que el campo de estudios constituido alrededor de la noción de complejidad era un campo condenado a ser un campo de estudios blando (*soft*) (Horgan, 1995, p. 107). Sostenía que era poco probable que el campo de estudio de los sistemas complejos lograra llegar a encontrar algunos principios generales significativos para su entendimiento. (Mitchell, 2009, p. 291) Además, consideraba que la gran predominancia del modelado computacional en el estudio de los sistemas estimados complejos podría ocasionar que en dicho estudio la artificialidad sea la que predomine. La predominancia del modelado computacional en el estudio de la complejidad llevaría a este a ser un estudio artificial; un estudio que se quedaría

---

<sup>2</sup> Cursiva introducida aquí.

<sup>3</sup> *At various times, researchers have debated whether complexity has become so meaningless that it should be abandoned, but they invariably conclude that the term has too much public-relations value. Complexologists often employ “interesting” as a synonym for “complex”. But what government agency would supply funds for research on a “unified theory of interesting things”? (The Santa Fe Institute, incidentally, will receive about half its \$5-million 1995 budget from the federal government and the rest from private benefactors.)*

*Complexologists may disagree on what they are studying, but most concur on how they should study it: with computers. (Horgan, 1995, p. 106)*

en lo computacional careciendo de hechos. (“*fact-free science*” Cf. Horgan, p. 105; Mitchell, 2009, p. 291)

Pero, más allá de una intención provocadora con la manera de referirse a algunos miembros del Instituto y sus campos de trabajo, y de sus reservas a la predominancia de los modelos computacionales para llevar a cabo ciertos estudios –reservas por demás muy difíciles de sostener hoy día, dado que el modelado computacional es la herramienta de investigación más efectiva en muchos casos (Cf., por ejemplo, Winsberg, 2010) –, no se puede desconocer que en la crítica de Horgan había, e incluso hay todavía, cierto fundamento. La proliferación de conceptos con los que medir ‘complejidad’, muchos de ellos desvinculados entre sí, sin que haya un significado representativo de la misma, es una situación que puede generar sospechas. Pues puede, por ejemplo, que se esté etiquetando indiscriminadamente de ‘complejo’ todo aquello que en estos días nos genere algún tipo de interés y asombro. O puede ser, asimismo, como él insistía, que los fondos destinados al desarrollo de los estudios ligados a la complejidad estén de alguna manera ayudando a sostener una idea global de la misma, aun cuando las nociones que se estiman o consideran no guardan entre sí una completa relación de coherencia.

En un corto artículo de 2002, publicado en la revista *Nature* bajo el título ‘*The bigger picture*’, el físico húngaro Tamás Vicsek expresaba un interrogante similar al planteado por Horgan: “¿por qué y cómo esta noción tan vaga llegó a ser un asunto tan central en la ciencia moderna...?”<sup>4</sup> (Vicsek, 2002, p. 131)

La preocupación de Tamás Vicsek, en cuanto al concepto de complejidad, iba igualmente por la posibilidad de un uso abusivo del mismo.

---

<sup>4</sup> *Why and how did this vague notion become such a central motif in modern science? Is it only a fashion, a kind of sociological phenomenon...?* (Vicsek, 2002, p. 131)

“Si un concepto no está bien definido, se puede abusar de él. Esto es particularmente cierto respecto a la complejidad, un concepto inherentemente interdisciplinario que ha penetrado en una variedad de campos intelectuales, desde la física hasta la lingüística, pero sin una teoría unificada subyacente.

La complejidad se ha convertido en una palabra de moda que se utiliza con la esperanza de llamar la atención o de obtener fondos; los institutos y las redes de investigación asociadas con sistemas complejos crecen como hongos.”<sup>5</sup> (Vicsek, 2002, p. 131)

Lo anterior entonces pone de manifiesto que el concepto de complejidad, dado su uso cada vez más extensivo y el sostenimiento de sus cada vez más variadas acepciones en simultáneo, ha llegado a despertar en algunos una sensación de sospecha en torno a su significatividad. No ha generado muy buenas sensaciones el hecho de que se sostengan nociones que parecen inconexas cuando de la complejidad se trata. Sin embargo, está la posibilidad, viendo el problema de la falta de criterios de unificación en torno a la noción de complejidad desde una perspectiva de la investigación científica un poco más optimista, que lo que esté sucediendo sea algo muy similar a lo sucedido en relación con el electromagnetismo previo a la década del 50 del siglo XIX.

Para aquella época, se contaba con un esquema que permitía describir cómo las cargas eléctricas generaban un campo magnético (Ley de Coulomb); se contaba además con un esquema cuya interpretación nos decía que no existían verdaderas cargas magnéticas

---

<sup>5</sup> *If a concept is not well defined, it can be abused. This is particularly true of complexity, an inherently interdisciplinary concept that has penetrated a range of intellectual fields from physics to linguistics, but with no underlying, unified theory.*

*Complexity has become a popular buzzword that is used in the hope of gaining attention or funding — institutes and research networks associated with complex systems grow like mushrooms.* (Vicsek, 2002, p. 131)

(Conjetura de Ampère); asimismo, de un esquema que describía cómo se generaba una corriente eléctrica a partir de un campo magnético variable (Ley de Faraday) y de un esquema que describía cómo las corrientes eléctricas generaban un campo magnético (Ley de Ampère) (Gell-Mann, 1998, p. 100).

Para la década del 50, del siglo XIX, con el desarrollo de las famosas ecuaciones de Maxwell, se logró integrar en un solo esquema estos fenómenos que se observaban previamente como actuantes de manera separada. El físico escocés, mediante el empleo de *ecuaciones diferenciales parciales*<sup>6</sup>, logró ofrecer una descripción novedosa de los fenómenos eléctricos y magnéticos bajo la idea de campo electromagnético. Es decir, con la teoría de Maxwell se logró, a partir de unas ecuaciones que podían expresar relaciones entre varios parámetros a la vez, interpretar fenómenos eléctricos y magnéticos -estimados de manera previa como fenómenos separados- bajo la clasificación a partir de ese momento de fenómenos electromagnéticos.

Lo que quiere decir que es posible, como insinúa Seth Lloyd (2001), que la situación problemática con los fenómenos que caen bajo la clasificación de complejos, sin haber aún un esquema unificador significativo mediante el cual podamos describirlos, se deba a que aún no ha llegado un personaje como Maxwell al campo de estudio de la complejidad. Restaría esperar entonces que un genio como el físico escocés logre, mediante unas ecuaciones o un esquema más general que los que se han propuesto, integrar todos estos fenómenos asociados e interactuantes de la complejidad. Dice al respecto Seth Lloyd:

---

<sup>6</sup> Las ecuaciones diferenciales parciales (EDP) representan una poderosísima herramienta matemática para el desarrollo, por ejemplo, de abstracciones de un cierto nivel de profundidad, como ocurre con el concepto de 'campo'. Abstracciones muy difíciles de lograr mediante el uso de ecuaciones diferenciales ordinarias (EDO), que son más bien una herramienta matemática muy útil para la descripción de trayectorias. Las EDP permiten el tratamiento matemático de varios parámetros interactuando simultáneamente.

“Un análogo histórico del problema de la medición de la complejidad es el problema de la descripción del electromagnetismo antes de las ecuaciones de Maxwell. En el caso del electromagnetismo, cantidades tales como las fuerzas eléctricas y magnéticas que surgieron en diferentes contextos experimentales fueron originalmente consideradas como fundamentalmente diferentes. Eventualmente quedó claro que la electricidad y el magnetismo eran de hecho aspectos estrechamente relacionados de la misma cantidad fundamental, el campo electromagnético. Similarmente, los investigadores contemporáneos en arquitectura, biología, informática, sistemas dinámicos, ingeniería, finanzas, teoría de juegos, etc., han definido diferentes medidas de complejidad para cada campo.”<sup>7</sup> (Lloyd, 2001, p. 7)

Según este enfoque, insinuado por Lloyd, podría considerarse que los estudios de la complejidad conformarían un campo científico en un estadio de desarrollo pre-paradigmático, en la terminología de Kuhn (1962/1970), en la espera de una revolución científica inicial que unifique la diversidad de escuelas o tendencias dentro del campo.

Según otro enfoque, que no ve a los estudios de la complejidad como conformando un campo científico pre-paradigmático, todavía en búsqueda de su unificación, sino como una manera radicalmente novedosa de hacer ciencia, la complejidad en sí misma es un fenómeno inaprensible desde un único esquema teórico o explicativo. Esto se debería en gran medida a que la complejidad se suele estimar de objetos (sistemas) inmersos en una multiplicidad de

---

<sup>7</sup> *An historical analog to the problem of measuring complexity is the problem of describing electromagnetism before Maxwell's equations. In the case of electromagnetism, quantities such as electric and magnetic forces that arose in different experimental contexts were originally regarded as fundamentally different. Eventually it became clear that electricity and magnetism were in fact closely related aspects of the same fundamental quantity, the electromagnetic field. Similarly, contemporary researchers in architecture, biology, computer science, dynamical systems, engineering, finance, game theory, etc., have defined different measures of complexity for each field.* (Lloyd, 2001, p. 7)

procesos de transformación en los que el *azar*, entre otros factores, juega ciertamente un papel no menor, puesto que en ocasiones contribuye a una expansión del potencial, bien sea combinatorio o asociativo, entre las (muchísimas) partes que integran dichos objetos (sistemas), permitiendo esto que se abra un abanico de nuevas posibilidades de aspectos sobre los que estimar complejidad (Cf., por ejemplo, Simon, 1962). Con la expansión del potencial, bien sea combinatorio o asociativo, de los componentes que integren el sistema, se abre paso a la posibilidad, dada sus dinámicas de transformación, de que la complejidad se manifieste de nuevas maneras.

Esto dejaría rezagados a los esquemas previos con los que se habían logrado capturar o encapsular, en cierta medida, aspectos parciales de sus posibles múltiples manifestaciones. La expansión de los aspectos sobre los que estimar complejidad, dada la multiplicidad de los procesos de transformación de los objetos (sistemas) en los que se suele dirigir la mirada para realizar dichas estimaciones, hace que los esquemas iniciales queden rezagados en el intento de especificación de un posible contenido global del concepto de complejidad.

Sin embargo, los ‘complejólogos’, tan duramente criticados por Horgan, han sido perseverantes en el desarrollo de métodos con los que intentan promover alternativas creativas a los retos que esta situación de rezago plantea. Con sus métodos intentan revelar algunos aspectos importantes de estos objetos (sistemas) -que se van dando en sus procesos de transformación- aun cuando predicciones precisas de cómo todo va a terminar no sean posibles. Acá, efectivamente, el uso de computadores se constituye como una de las herramientas más útiles de las que disponen en este intento, básicamente porque estos resultan indispensables para la exploración, por medio de *simulaciones*, de posibles situaciones futuras dada la dinámica particular en la que estén inmersos los objetos de estudio en consideración.

Cabe destacar que, a diferencia de Horgan, Tamás Vicsek, bajo la pregunta que planteaba de si las investigaciones acerca de la complejidad responden a una suerte de moda o fenómeno social, también dejaba abierta la puerta a que “[l]o que estamos presenciando en este contexto sea un cambio de paradigma en los intentos de comprender nuestro mundo mientras nos damos cuenta de que las leyes del todo no pueden deducirse cavando más profundo en los detalles.”<sup>8</sup> (Vicsek, 2002, p. 131) Cambio que viene en gran medida posibilitado por el desarrollo de ordenadores.

“En cierto modo, este cambio ha sido forjado por el desarrollo de instrumentos. Tradicionalmente, mejores microscopios o telescopios más grandes son construidos para obtener una mejor comprensión de problemas particulares. Pero las computadoras han permitido nuevas formas de aprendizaje. Al modelar directamente un sistema compuesto por muchas unidades, uno puede *observar, manipular*<sup>9</sup> y comprender el comportamiento de todo el sistema mucho mejor que antes, como en los casos de los modelos de redes neuronales y de las subastas virtuales por agentes inteligentes, por ejemplo. En este sentido, una computadora es una herramienta que no mejora nuestra vista (como lo hace el microscopio o el telescopio), sino más bien nuestra visión de los mecanismos subyacentes de los sistemas complejos.”<sup>10</sup> (Vicsek, 2002, p. 131)<sup>11</sup>

---

<sup>8</sup> *What we are witnessing in this context is a change of paradigm in attempts to understand our world as we realize that the laws of the whole cannot be deduced by digging deeper into the details.* (Vicsek, 2002, p. 131)

<sup>9</sup> Cursivas introducidas aquí.

<sup>10</sup> *In a way, this change has been wrought by the development of instruments. Traditionally, improved microscopes or bigger telescopes are built to gain a better understanding of particular problems. But computers have allowed new ways of learning. By directly modelling a system made of many units, one can observe, manipulate and understand the behavior of the whole system much better than before, as in the cases of networks of model neurons and virtual auctions by intelligent agents, for example. In this sense, a computer is a tool that improves not our sight (as does the microscope or telescope), but rather our insight into mechanisms within complex systems.* (Vicsek, 2002, p. 131)

<sup>11</sup> La idea de que el progreso en una disciplina puede estar impulsado por revoluciones técnicas también está defendida, respecto de la neurociencia, en Bickle 2016 (*Revolutions in neuroscience*).

Lo que Vicsek nos estaría planteando entonces sería la consideración de la posibilidad de una suerte de abordaje metodológico científico de tipo «*manipulacionista*», en el que, por ejemplo, técnicas de experimentación computacionales como las redes neuronales artificiales constituyan un tipo de herramienta fructífera para el abordaje del fenómeno de la complejidad, pues estas permiten una suerte de “manipulación” -vía *simulación computacional*- de esta clase de sistemas -vistos como un todo- de una manera no antes posible, y, sobre todo, permiten *visualizar* los posibles efectos en la dinámica de estos sistemas de tales manipulaciones. Mediante modelos computacionales de redes neuronales artificiales podrían ofrecerse aproximaciones no analíticas de fenómenos complejos (Lenhard, 2016). Se habla de aproximaciones no analíticas, principalmente por los métodos de resolución regularmente empleados para las ecuaciones diferenciales parciales con las que se suelen estudiar la dinámica de estos sistemas.

Debido a sus características, los sistemas complejos difícilmente pueden ser tratados ni satisfactoria ni exclusivamente por métodos analíticos tradicionales, pues los mismos solo tienen un alcance limitado y en el mejor de los casos, éxito muy parcial. El tipo de aproximación que ha logrado tratar con éxito abrumador sistemas complejos está basado en el uso intensivo e imprescindible de métodos computacionales y representación gráfica de enormes cantidades de información. Un análisis de lo anterior arroja que se debe renunciar al menos parcialmente a las pretensiones epistémicas normativas (y de comprensión) para poder tratar adecuadamente estos fenómenos. ¿Qué se entiende aquí por “éxito”? No se trata de transparencia epistémica ni de entendimiento en términos del establecimiento de teorías generales. En un primer momento, se entiende por éxito la capacidad de resolver un problema concreto respecto de la identificación de un patrón de comportamiento del sistema complejo,

y la representación científica del mismo mediante simulación computacional de un aspecto particular de ese comportamiento.

Argumentos a favor de lo anterior se ofrecerán en los **capítulos I y II**, guiados respectivamente por las preguntas *¿Es posible una teoría unificada de la complejidad? ¿Es posible una métrica para la complejidad?* Describir los rasgos que se han logrado identificar como distintivos de los sistemas complejos [cap.1] y los intentos teóricos para abordar algunos de estos rasgos [cap.2] nos permitirá plantear que pretensiones epistemológicas unificacionistas resultan fuertemente frustradas cuando del fenómeno de la complejidad se trata. Sin embargo, esto no constituye una invitación a que se dejen de proponer por completo alternativas teóricas para un abordaje general de la misma.

Intentos teóricos como el de Murray Gell-Mann, con la noción de *complejidad efectiva*; el de Brian Goodwin, con la noción de *complejidad morfológica* -mediante el número de simetrías rotas durante el proceso de desarrollo-; o el de Miguel Fuentes, con la noción de la *complejidad* y la *emergencia modelo paramétricas*, por mencionar sólo algunos, representan avances significativos para su abordaje. Particularmente, para un abordaje de carácter *métrico*, y no propiamente uno de carácter definitorio.

Una vez establecida una justificación para lo anterior, nos volcaremos en el **capítulo III** a la revisión de un ejemplo particular del tipo de herramienta general más exitosa en la actualidad para tratar con la complejidad: las simulaciones computacionales, y su aplicación al estudio del plegamiento de proteínas. En este capítulo, guiado por la pregunta *¿Se está produciendo un giro hacia la predicción en ciencia?*, se resalta el éxito abrumador obtenido por una red neuronal artificial de aprendizaje profundo en el abordaje de una clara manifestación de la complejidad. El problema del plegamiento de las proteínas, en el que se había dedicado grandes esfuerzos y grandes sumas de dinero en su intento de resolución, fue

resuelto por un laboratorio de inteligencia artificial mediante una red neuronal a la que llamaron *AlphaFold*.

Lo que esta red consigue es *predecir* la estructura tridimensional específica de una extensa cantidad de proteínas, tomando como *inputs* las secuencias de aminoácidos que las constituyen. De modo que en este capítulo no sólo se intenta evidenciar la gran relevancia que cobran las simulaciones computacionales en el tratamiento de la complejidad, sino que además se intenta evidenciar que efectivamente se está produciendo un giro en ciencias, debido en gran medida al auge de las simulaciones computacionales, que privilegia la predicción por sobre otros valores epistémicos más tradicionales como la explicación o la transparencia epistémica. Idea que, como lo hacen saber, por ejemplo M. Chirimuuta (2018) y Daniel Weiskopf (2022), se viene discutiendo recientemente en filosofía de la ciencia.

Esto nos lleva a plantear la pregunta sobre *¿Cuál es la novedad epistémica de las simulaciones computacionales?* Esta pregunta será la que guíe el **capítulo IV**, en el que intentaremos plantear que, pese a las consideraciones importantes de Roman Frigg y Julian Reiss (2009), quienes sostienen que las simulaciones computacionales no introducen estrictamente aspectos novedosos en la discusión filosófica sobre modelos, *la opacidad epistémica*, junto con el manipulacionismo ligado a las poderosas técnicas de visualización, y la *experimentación iterada* sí introducen novedades epistemológicas; novedades que se ligan positivamente con el marcado giro predictivo que se viene dando en ciencias en tiempos recientes. En apoyo de lo anterior, nos valdremos principalmente de las propuestas de Paul Humphreys (1995; 2004; 2008), Johannes Lenhard (2016; 2017) y Peter Galison (1996).

Una vez resueltas estas cuatro preguntas, podemos ocuparnos de establecer algunas conclusiones sobre cuáles son las relaciones entre epistemología, los sistemas complejos y las simulaciones computacionales desarrolladas a través del uso de redes neuronales

artificiales. Se verá aquí, con mayor claridad, una vez hecho el tránsito del capítulo I al capítulo IV, por qué es necesaria la renuncia a la pretensión normativa, por qué hay que asumir cierta opacidad epistémica y por qué es imposible prescindir de recursos heurísticos para aproximarnos a una comprensión de los fenómenos complejos. Esta renuncia parcial a la normatividad epistémica no es, sin embargo, una derrota para la epistemología. Solamente supone modificar los fines de esta y su enriquecimiento con nuevas herramientas teórico-prácticas.

# CAPÍTULO I

## ALGUNOS RASGOS DISTINTIVOS DE LOS SISTEMAS COMPLEJOS

### *¿Es posible una teoría unificada de la complejidad?*

#### Abstract

En este capítulo se presentan algunas de las características distintivas del fenómeno de la complejidad. Se otorga un papel fundamental a la emergencia, identificando sistemas complejos con sistemas emergentes. Además, se resaltan el comportamiento colectivo, la autoorganización, la adaptación, la estructura jerárquica y la retroalimentación.

En este capítulo se presentarán algunas de las características estimadas distintivas de los sistemas complejos. Si bien parece no haber consenso aún sobre cómo definir a estos sistemas, sí se han logrado identificar ciertos rasgos distintivos; destacándose entre ellos la *emergencia*. Un sistema complejo es, entre otras cosas, un sistema en el que se dan procesos y propiedades emergentes (Mitchell, 2009; Bunge, 2004; Gell-Mann, 1998; Taylor, 2001). De ahí que este capítulo inicie justamente con una discusión de la emergencia.

Si bien se acepta que la emergencia es quizás una de las características más distintivas de la complejidad, cierto es también que un sistema complejo es uno constituido de muchísimas partes que interactúan de manera *no lineal* (Mitchell, 2009, p. 22), dando lugar a *comportamientos colectivos* (ibid., 2009, p. 4), *autoorganizados* (ibid., 2009, p. 12), con los que intentan, entre otras cosas, *adaptarse* al medio (ibid., p. 13). Además, presentan *estructuras jerárquicas* (ibid., p. 109), y entre sus distintos niveles de organización hay *feedback* (ibid., p. 296), o retroalimentación. A continuación, analizaremos en orden cada una de estas características.

## II. *Emergencia*

Algunos ejemplares paradigmáticos de sistemas complejos/emergentes son los siguientes:

### *Colonias de insectos*

“Una colonia de hormigas, por ejemplo, puede constar de cientos a millones de hormigas individuales, cada una de las cuales es una criatura bastante simple que obedece a sus imperativos genéticos de buscar comida, responder de manera simple a las señales químicas de otras hormigas en su colonia, luchar contra intrusos, etcétera. Sin embargo, como puede atestiguar cualquier observador casual al aire libre, las hormigas en una colonia, cada una realizando sus propias acciones relativamente simples, trabajan juntas para construir estructuras asombrosamente complejas que son claramente de gran importancia para la supervivencia de la colonia en su conjunto.”<sup>12</sup>

(Mitchell, 2009, p. 4)

### *Las economías*

“Las economías son sistemas complejos en los que los componentes ‘simples y microscópicos’ vendrían a ser personas (o empresas) que compran y venden bienes, y el comportamiento colectivo vendría a ser el comportamiento complejo y difícil de predecir de los mercados en su conjunto, como los cambios del precio de las viviendas en el mercado en diferentes zonas del país o las fluctuaciones en los precios de las acciones.”<sup>13</sup> (ibid., pp. 9, 10)

---

<sup>12</sup> *An ant colony, for instance, can consist of hundreds to millions of individual ants, each one a rather simple creature that obeys its genetic imperatives to seek out food, respond in simple ways to the chemical signals of other ants in its colony, fight intruders, and so forth. However, as any casual observer of the outdoors can attest, the ants in a colony, each performing its own relatively simple actions, work together to build astoundingly complex structures that are clearly of great importance for the survival of the colony as a whole.* (Mitchell, 2009, p. 4)

<sup>13</sup> *Economies are complex systems in which the “simple, microscopic” components consist of people (or companies) buying and selling goods, and the collective behavior is the complex, hard-to-predict behavior of*

### *Los sistemas inmunológicos*

“El sistema inmune es otro ejemplo de un sistema en el que componentes relativamente simples dan lugar colectivamente a un comportamiento muy complejo que involucra señalización y control, y en el que la adaptación se produce con el tiempo.”<sup>14</sup> (ibid., p. 6)

### *El cerebro*

“El científico cognitivo Douglas Hofstadter, en su libro ‘*Gödel, Escher, Bach*’, hace una extensa analogía entre colonias de hormigas y cerebros, siendo ambos sistemas complejos, en los que, componentes relativamente simples, con sólo una limitada comunicación entre sí, dan lugar colectivamente a un sistema que presenta un comportamiento ‘global’ complejo y sofisticado. En el cerebro, los componentes simples son células llamadas neuronas. El cerebro está formado por muchos tipos diferentes de células además de las neuronas, pero la mayoría de los neurocientíficos creen que las acciones de las neuronas y los patrones de conexión entre grupos de neuronas son lo que causa la percepción, el pensamiento, los sentimientos, la conciencia y las otras actividades cerebrales a gran escala que son importantes.”<sup>15</sup> (ibid., pp. 5, 6)

---

*markets as a whole, such as changes in the price of housing in different areas of the country or fluctuations in stock prices. (ibid., p. 9, 10)*

<sup>14</sup> *The immune system is another example of a system in which relatively simple components collectively give rise to very complex behavior involving signaling and control, and in which adaptation occurs over time. (ibid., p. 6)*

<sup>15</sup> *The cognitive scientist Douglas Hofstadter, in his book *Gödel, Escher, Bach*, makes an extended analogy between ant colonies and brains, both being complex systems in which relatively simple components with only limited communication among themselves collectively give rise to complicated and sophisticated system-wide (“global”) behavior. In the brain, the simple components are cells called neurons. The brain is made up of many different types of cells in addition to neurons, but most brain scientists believe that the actions of neurons and the patterns of connections among groups of neurons are what cause perception, thought, feelings, consciousness, and the other important large-scale brain activities. (ibid., pp. 5, 6)*

Los anteriores son algunos ejemplos con los que se intenta poner de manifiesto un fenómeno sumamente interesante que se da con algunos agregados de elementos o ítems relativamente comunes, y es que presentan ciertas características no detectables si a estos elementos o ítems constitutivos se los intenta analizar de manera aislada; son características presentes en una colectividad mas no en sus elementos o ítems constitutivos. Son características emergentes; características por asociación o combinación de sus elementos o ítems constitutivos.

En los ejemplos presentados se puede observar además que son características presentes en colectividades extensamente numerosas, pues un rasgo distintivo de los sistemas complejos es que están constituidos de muchísimas partes. A continuación, nos detendremos en un ejemplo en particular, la superconductividad.

La licuefacción (*liquefaction*), o licuación de los gases, es un proceso que permite el cambio de una sustancia del estado gaseoso al estado líquido. En la licuefacción de los gases lo que ocurre básicamente es una sobrepresión elevada, a bajas temperaturas. Esto, debido a un proceso de compresión isotérmica, que aumenta su presión; y, a un proceso de expansión adiabática, que disminuye su temperatura. En 1908, el físico holandés, Heike Kamerlingh Onnes, logró la licuación del helio-4; un isótopo no radioactivo y ligero del helio. Esto le significó la obtención del premio nobel de física en el año 1913. La licuación de los gases es un experimento que, entre otras cosas, permite disminuir considerablemente la temperatura a la que se encuentre un material; posibilidad de gran importancia para la época, pues, una pregunta que atraía la atención de varios físicos en aquel momento era: ¿qué pasaría con la conducción de una corriente eléctrica si el metal conductor era sometido a una disminución de temperatura? (Slichter, 2007). Haciendo experimentos en esta dirección, el mismo Heike Kamerlingh Onnes, junto con un grupo de estudiantes, en 1911, dio con una respuesta

sumamente sorprendente. Examinando las bondades conductoras del mercurio a bajas temperaturas, observó algo muy llamativo. Ese fenómeno llamativo que estaba tras la observación de Kamerlingh era el fenómeno de la superconductividad (Slichter, 2007).

“Él y sus alumnos se pusieron manos a la obra para estudiar qué sucedía con varias propiedades de los materiales cuando eran enfriados. Uno de sus alumnos estaba estudiando la resistencia eléctrica de los cables. Descubrió que a medida que enfriaba el cable de mercurio, la resistencia eléctrica del cable descendía abruptamente cuando llegaba a unos 3,6 grados por encima del cero absoluto. La caída fue enorme: la resistencia se hizo al menos veinte mil veces menor. La caída se produjo en un intervalo de temperatura demasiado pequeño para medirlo. Pero por lo que lo que pudieron notar, la resistencia eléctrica había desaparecido por completo.”<sup>16</sup> (ibid., 2007)

Que la resistencia eléctrica del metal se haga cero es un fenómeno que representa una situación, podría decirse, conflictiva con el tratamiento regular de los metales en cuanto a su característica conductora. Previo a la observación de este fenómeno, un metal era observado como una red de iones, rodeada por una capa, o nube de electrones, que viajan con cierto grado de libertad a través de esa red cuando se le induce una diferencia de potencial (ibid., 2007). Lo que se observa bajo esta interpretación es que los electrones pierden algo de energía en este fluir de un extremo a otro de la red por choques que se producen entre estos y algunos iones. Pérdida de energía que se transforma en calor (efecto Joule).

---

<sup>16</sup> *He and his students set to work to study what happened to various properties of materials when they were that cold. One of his students was studying the electrical resistance of wires. He found that as he cooled mercury wire the electrical resistance of the wire took a precipitous drop when he got to about 3.6 degrees above absolute zero. The drop was enormous - the resistance became at least twenty thousand times smaller. The drop took place over a temperature interval too small for them to measure. As far as they could tell, the electrical resistance completely vanished.* (Slichter, 2007)

En los superconductores ocurre una transición de fase, como la que ocurre cuando el hielo se derrite o cuando el agua se congela (ibid., 2007). Sin embargo, en los superconductores el aspecto del material no cambia. Lo que cambia es el comportamiento de los electrones. Bajo cierta temperatura crítica, los electrones del material se organizan en estados ligados de dos electrones (pares de Cooper) debido a sus interacciones vía fonones (nodos que se forman por la vibración de los iones del metal). El comportamiento de los electrones en este estado difiere del que presentan en estado libre. Los pares de Cooper se comportan en cierto sentido como bosones. Lo que hace que se puedan formar estados coordinados (coherentes). Estas parejas en efecto se coordinan para formar un estado cuántico colectivo macroscópico (el estado superconductor, sin resistencia eléctrica).

Al tener la misma carga los electrones se repelen, pero se encuentran a su vez interactuando con los iones positivos del material (que están vibrando). Es el acoplamiento entre los electrones y las vibraciones de los iones lo que da lugar a una interacción atractiva neta entre los electrones (que forman pares). Gracias a la teoría BCS (Bardeen, Cooper, Schrieffer), podemos dar cuenta de lo que se conoce como superconductores convencionales.

En 1933, gracias a los trabajos de Walther Meissner y Robert Ochsenfeld se pudo observar otra propiedad importante que aparece en ciertos metales cuando son enfriados por debajo de una temperatura crítica. Lo que observaron estos dos físicos es que por debajo de una temperatura crítica (relativa al material en consideración) el campo magnético interno se hace cero. Los superconductores repelen el campo magnético a su alrededor (efecto Meissner) (ibid., 2007). Lo que quiere decir entonces que, ciertos metales, enfriados por debajo de su temperatura crítica, presentan nuevas propiedades. A saber, resistencia eléctrica nula, o casi nula, y campo magnético interno nulo, o casi nulo. Por debajo de una temperatura crítica el comportamiento de los electrones varía de una manera considerable. Este nuevo tipo de

comportamiento requiere un nuevo dominio de análisis para poder dar cuenta de él. Desde las observaciones de Kamerlingh Onnes hasta la teoría BCS, transcurrieron varios años. Y en medio, varios desarrollos teóricos y experimentales para poder dar cuenta del fenómeno de la superconductividad. Entre ellos, el desarrollo teórico de la condensación Bose-Einstein; la estadística de Fermi-Dirac; los desarrollos que lograron Walther Meissner y Robert Ochsenfeld; la teoría de Ginzburg-Landau; entre otros (ibid., p. 2007).

El conocimiento que se tenía del comportamiento de los electrones cuando se observó el fenómeno de la superconductividad no era suficiente para poder dar cuenta de él; cosa que era apenas natural, pues era un comportamiento totalmente nuevo el que presentaban bajo esas condiciones. A partir del conocimiento de las propiedades de los electrones, a temperaturas no más bajas de la temperatura crítica del metal en consideración, no era posible dar cuenta de la propiedad superconductor de ciertos metales cuando se traspasa dicha temperatura crítica. La razón es que es una propiedad enteramente nueva que *emerge* de un tipo especial de interacción de estos con los iones positivos del metal. El estar bajo cierta temperatura conlleva un tipo de organización o disposición de los electrones, que da lugar a nuevos tipos de interacción. Estos nuevos tipos de interacción pueden dar lugar a nuevas propiedades, como es el caso de la superconductividad; que, como se mencionó, es un estado cuántico macroscópico.

A principios de 1920, un físico de la India de nombre Satyendra Nath Bose, pensando en la teoría de los fotones con la que Einstein diera explicación al efecto fotoeléctrico, propuso un tipo de tratamiento estadístico especial para este tipo de partículas indistinguibles. La idea de Bose era que, a partir de cierta temperatura crítica extremadamente baja, la longitud de onda de de Broglie se hace bastante grande, y las partículas empiezan a solaparse, formando así un macroestado cuántico coherente; como ocurre con el fenómeno de la

superconductividad. Lo que describe en términos generales la longitud de onda de de Broglie es el tamaño de la región en la que está distribuida la función de onda asociada a una partícula. En este tratamiento estadístico de partículas indistinguibles dos consideraciones surgen de manera principal. La primera es que estas partículas presentan estados cuánticos excluyentes. La otra es que estas partículas comparten estados cuánticos; es decir, que pueden estar en un mismo tipo de estado cuántico. A temperaturas por debajo de una temperatura crítica ciertas sustancias presentan una reorganización de sus partículas formando un macroestado. Un macroestado en el que la mayoría de sus partículas se solapan en el estado de mínima energía; se solapan en el estado fundamental. En este estado fundamental se dice de una sustancia que es un superfluido por no presentar viscosidad.

En un *paper* de 1972 titulado '*More is Different*', Philip Anderson, premio nobel de física, nos planteaba que el fenómeno de la superconductividad era un fenómeno que nos invitaba a replantearnos algunas consideraciones epistemológicas en relación con la explicación científica, permeadas fuertemente por hipótesis reduccionistas (explicaciones mecanicistas), en el sentido que pretendían dar cuenta de un fenómeno o proceso a partir de sus partes constitutivas (Suárez, 2005, p. 228); y, añadimos acá, deterministas. Determinismo en sentido ontológico; esto es: aplicado a sistemas concretos bajo la consideración que, "dado su estado en un cierto instante, su evolución para todo instante posterior resulta físicamente necesaria; en otras palabras, si el sistema se encuentra en el estado  $e_1$  en el instante  $t_1$ , las leyes físicas (en tanto regularidades ontológicas) hacen imposible que se encuentre en un estado diferente de  $e_2$  en  $t_2$ ." (Lombardi y López, 2015, p. 42)

“Resulta que el comportamiento de grandes y complejos agregados de partículas elementales no debe ser entendido en términos de una simple extrapolación de las propiedades de unas cuantas partículas.”<sup>17</sup> (Anderson, 1972, p. 393)

“El ejemplo sobresaliente de esto es el cristal. Construido a partir de un sustrato de átomos y espacio de acuerdo con las leyes que expresan la perfecta homogeneidad de los espacios, el cristal muestra de forma repentina e impredecible una simetría completamente nueva y muy hermosa. La regla general, sin embargo, incluso en el caso del cristal, es que un sistema grande es menos simétrico de lo que sugeriría su estructura subyacente. Simétrico como lo es, un cristal no llega a presentar una perfecta homogeneidad.

Quizás en el caso de los cristales esto parece ser simplemente un ejercicio de confusión. La regularidad de los cristales podía deducirse semiempíricamente a mediados del siglo XIX sin ningún razonamiento complicado. Pero a veces, como en el caso de la superconductividad, la nueva simetría –ahora llamada *simetría rota*<sup>18</sup> porque la simetría original ya no es evidente– puede ser de un tipo completamente inesperado y extremadamente difícil de visualizar. En el caso de la superconductividad transcurrieron treinta años entre el momento en que los físicos estaban en posesión de todas las leyes fundamentales necesarias para explicarla y el momento en que realmente lo lograron.

---

<sup>17</sup> *The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles.* (Anderson, 1972, p. 393)

<sup>18</sup> Cursiva introducida aquí.

[...] En este caso podemos ver cómo el todo no sólo va más allá que la suma de sus partes, sino que es muy diferente de las partes.”<sup>19</sup> (Anderson, 1972, p. 395)

En un interesante libro de 2005 titulado ‘*A different universe. Reinventing physics from the bottom down*’, el también premio nobel de física Robert B. Laughlin, quien se dice acostumbraba a pedirles a sus mejores estudiantes deducir a partir de primeros principios las leyes de la *superfluididad* a sabiendas de que es imposible (Sanjuán 2007, p. 891), siguiendo el llamado de Philip Anderson, en cuanto a lo que fenómenos como la superconductividad nos planteaba, nos decía que:

“Creo que fenómenos de organización primarios, como el clima, tienen algo de importancia duradera para contarnos sobre fenómenos más complejos, incluido nosotros mismos. Su carácter primario nos permite demostrar con certeza que están regidos por leyes microscópicas, pero también, paradójicamente, que algunos de sus más sofisticados aspectos son insensibles a los detalles de estas leyes. En otras palabras, podemos demostrar en estos casos simples que la organización puede adquirir significado y vida propia, y comenzar a trascender las partes de las que está hecha.

---

<sup>19</sup> *The outstanding example of this is the crystal: Built from a substrate of atoms and space according to laws which express the perfect homogeneity of spaces, the crystal suddenly and unpredictably displays an entirely new and very beautiful symmetry. The general rule, however, even in the case of the crystal, is that the large system is less symmetrical than the underlying structure would suggest: Symmetrical as it is, a crystal is less symmetrical than perfect homogeneity.*

*Perhaps in the case of crystals this appears to be merely an exercise in confusion. The regularity of crystals could be deduced semiempirically in the mid-19th century without any complicated reasoning at all. But sometimes, as in the case of superconductivity, the new symmetry -now called broken symmetry because the original symmetry is no longer evident- may be of an entirely unexpected kind and extremely difficult to visualize. In the case of superconductivity, 30 years elapsed between the time when physicists were in possession of every fundamental law necessary for explaining it and the time when it was actually done.*

[...] *In this case we can see how the whole becomes not only more than but very different from the sum of its parts.* (Anderson, 2007, p. 395)

Lo que la ciencia física tiene entonces que decirnos sobre este punto es que ‘*el todo es más que la suma de sus partes*’ no es simplemente un asunto conceptual, sino realmente un fenómeno físico.

[...] Estoy cada vez más convencido de que toda ley física que conocemos tiene orígenes colectivos, no solo una parte de ellas. En otras palabras, la distinción entre las leyes fundamentales y las leyes que descenden de ellas es un mito, como lo es la idea del dominio del universo únicamente a través de las matemáticas. Una ley física no puede generalmente ser anticipada por el pensamiento puro, sino que debe descubrirse experimentalmente, porque el control de la naturaleza se logra sólo cuando la naturaleza lo permite a través de un principio de organización. Uno podría denominar esta tesis como el fin del reduccionismo (la creencia de que las cosas necesariamente serán aclaradas al dividir las partes cada vez más pequeñas), pero eso no sería del todo exacto. Todos los físicos son reduccionistas de corazón, incluido yo mismo. No deseo impugnar el reduccionismo tanto como establecer su lugar apropiado en el gran esquema de las cosas.”<sup>20</sup> (Laughlin, 2005, p. 15)

---

<sup>20</sup> *I think primitive organizational phenomena such as weather have something of lasting importance to tell us about more complex ones, including ourselves: their primitiveness enables us to demonstrate with certainty that they are ruled by microscopic laws but also, paradoxically, that some of their more sophisticated aspects are insensitive to details of these laws. In other words, we are able to prove in these simple cases that the organization can acquire meaning and life of its own and begin to transcend the parts from which it is made.*

*What physical science thus has to tell us is that the whole being more than the sum of its parts is not merely a concept but a physical phenomenon.*

*[...] I am increasingly persuaded that all physical law we know about has collective origins, not just some of it. In other words, the distinction between fundamental laws and the laws descending from them is a myth, as is the idea of mastery of the universe through mathematics alone. Physical law cannot generally be anticipated by pure thought, but must be discovered experimentally, because control of nature is achieved only when nature allows this through a principle of organization. One might subtitle this thesis the end of reductionism (the belief that things will necessarily be clarified when they are divided into smaller and smaller component parts), but that would not be quite accurate. All physicists are reductionists at heart, myself included. I do not wish to impugn reductionism so much as establish its proper place in the grand scheme of things. (Laughlin, 2005, p. xv)*

Fenómenos como la superconductividad y la superfluidez, entre muchísimos otros, son fenómenos, sobre todo, de organización cambiante, y cada organización adquiere una nueva significación, un nuevo contenido de información efectiva o complejidad efectiva (Cf. pp. 45 y 45), por lo que describirlos desde un único esquema general resulta la mayor parte de las veces sumamente difícil, ofreciéndonos, en no pocas ocasiones, «*aproximaciones no analíticas*» como única posibilidad; aproximaciones de las que hablaremos más adelante. Como afirma Laughlin: son fenómenos que presentan aspectos con cierta insensibilidad a las leyes que rigen sobre los elementos o ítems involucrados en ellos; son fenómenos cuyas explicaciones cuesta organizar dentro de una teoría o incluso un marco conceptual específico.

Si bien en el caso de la superconductividad y la superfluidez se ha logrado una aproximación explicativa analítica (al menos en superconductores convencionales), no siempre es este el caso. El fenómeno de la complejidad suele ser visto como un fenómeno fuertemente asociado al de la emergencia (Mitchell, 2009; Bunge, 2004; Gell-Mann, 1998; Taylor, 2001); fenómeno descrito como uno en el que surgen propiedades *globales* irreductibles explicativamente a las propiedades de los componentes (considerados bajo cierto nivel de organización) de un sistema, que dan lugar a ellas mediante sus múltiples interacciones (combinándose dichos elementos o asociándose).

Es decir, son propiedades de un objeto (sistema) complejo que ni sus constituyentes ni sus precursores poseen (Bunge, 2004, p. 20). Lo que no quiere decir que sean propiedades que no se puedan abordar a partir de los componentes del sistema. Más bien, no se pueden explicar a partir únicamente de las propiedades de sus partes constitutivas, puesto que se requiere conocer el tipo de interacción que están teniendo dichas partes, y, sobre todo, el conjunto de valores paramétricos que bajo estas interacciones presentan algunas variables consideradas relevantes en la dinámica del sistema complejo en estudio.

Y es ahí donde radica la dificultad: en disponer de este conjunto de interacciones y de los valores paramétricos que bajo tales interacciones presentan ciertas variables implicadas en la dinámica del sistema, ya que estos suelen ser sistemas dinámicos no deterministas, que, bajo ciertas condiciones, algunas veces favorecidas por el azar (del que hablaremos en el segundo capítulo), presentan nuevas asociaciones o combinaciones de sus elementos, lo que conlleva a que el conjunto de interacciones varíe. Esta variación en sus interacciones viene acompañada regularmente de nuevos comportamientos; comportamientos que suelen dar lugar a nuevas propiedades del sistema. Lo que podría llevarnos a considerar que el quid de las propiedades emergentes está en las interacciones y en los valores de ciertos parámetros que bajo tales interacciones presentan ciertas variables ligadas a los elementos constituyentes de estos sistemas complejos; propiedades que, insistimos, no son distributivas sino globales. (Bunge, 2004, p. 29)

Estas propiedades globales pueden ser de tipo *funcional* (Rescher, 1998, p. 9) – como la que se desprende de la estructura tridimensional de las proteínas-, *estructural* (ibid., p. 9) – como el procesamiento de información de tipo jerárquico en el neocórtex- u *organizacional* (ibid., p. 9) – como la que presenta una bandada de estorninos (*Sturnus vulgaris*) en busca de alimento o en defensa de otras aves rapaces. El conocimiento, aun detallado y preciso de las propiedades de los componentes de un sistema vistos de manera aislada, nos resulta insuficiente para poder dar cuenta de estas propiedades en sentido global. Esto, pese a considerar que es a partir de estos componentes (considerados bajo cierto nivel de organización), combinados o asociados, que dichas propiedades tienen lugar, y que es a partir de la interacción de estos que se pueden definir casos específicos de dichas propiedades globales.

Estas propiedades, bien sean de tipo funcional, estructural u organizacional, requieren de un gran conocimiento adicional al que se tenga de las propiedades de dichos componentes para intentar dar cuenta de ellas; además del desarrollo de técnicas y estrategias no analíticas, sino más bien aproximativas, implementadas iterativamente y retroalimentadas mediante muchísimo ensayo y error.

Si bien es cierto entonces que la complejidad no se agota en el fenómeno de la emergencia, la fuerte vinculación que hay entre ellas nos puede llevar a observar a la complejidad como justamente la generación de propiedades emergentes. Lo que querría decir que por complejidad nos estaríamos refiriendo intuitivamente al fenómeno de la emergencia. De ahí que llamemos ‘complejo’ a un sistema en el que se dan propiedades emergentes; considerando intuitivamente que un sistema se hace más complejo en tanto las propiedades emergentes que se dan en él se “distancien” más de las propiedades de sus componentes.

Fenómenos como la superconductividad y la superfluidez, entre otros, ejemplifican situaciones en las que surgen propiedades sistémicas enteramente nuevas a partir de nuevas relaciones entre sus elementos constitutivos. Estas nuevas relaciones pueden ser, a grandes rasgos, de dos tipos: asociativas y combinatorias (Bunge, 2004). Siendo las relaciones de tipo asociativa las más comunes, pues son por ejemplo las menos exigentes desde un tipo de vista energético (Bunge, 2004). Además, en este tipo de relación la naturaleza de los elementos no se modifica. En la combinación, por contraposición, los constituyentes sí se modifican. La emergencia, característica notoria de la complejidad, siendo principalmente un fenómeno ontológico, representa serias dificultades epistemológicas, particularmente en sentido metodológico, para abordarla predictiva y explicativamente.

De manera deductiva, desde el cuerpo de conocimientos inicial que se tenga del sistema y de las propiedades de sus componentes, no es posible dar con estas propiedades. Se requiere

un cuerpo de conocimiento adicional al que se tenía inicialmente para retroalimentar el esquema o los esquemas con los que se teorizaba sobre los elementos del sistema; en esta retroalimentación se suelen considerar nuevos comportamientos de dichos elementos, quizás producto de nuevas asociaciones o combinaciones. La retroalimentación se suele dar por una relación teórica bidireccional (*bottom up* y *top down*) entre el esquema con que se da cuenta del sistema en sentido global, incorporando la propiedad emergente en consideración, y el esquema con el que se da cuenta de los elementos constituyentes del sistema.

Los diferentes tipos de elementos que dan lugar a propiedades emergentes mediante, básicamente, sus interacciones; los diferentes tipos de leyes que actúan sobre esta variedad de elementos; los diferentes tipos de dominio a los que pertenecen; los diferentes tipos de procesos en ellos involucrados: todo esto representa una enorme dificultad en el intento de aprehensión de la emergencia bajo un único esquema general; esquema que pueda dar cuenta de fenómenos no solo como el de la superconductividad y el de la superfluidez, sino que pueda dar cuenta de otro tipo de fenómenos emergentes en otros tipos de sistemas complejos.

### I.2. *Comportamientos colectivos*

Fenómenos como el de la complejidad, donde intervienen múltiples variables independientes, que ejercen interacciones simultáneas entre sí, nos invitan a considerar enfoques como el que diera Boltzmann (mecánica estadística) a gases ideales mediante el concepto de entropía. El tratamiento estadístico de Boltzmann para sistemas con múltiples componentes se ha evidenciado fructífero. Sin embargo, un gas ideal no es un sistema complejo. Pese a estar compuestos de muchísimas partes (miles de moléculas), estos sistemas no exhiben una característica notoria de los sistemas complejos, y es la de presentar *interacciones no lineales*. Los componentes de un sistema complejo (bajo cierto nivel de

organización) interactúan de manera no lineal. Situación que no ocurre en los sistemas tratados por Boltzmann (gases ideales). En estos sistemas (por ejemplo, en equilibrio) una única molécula podría ser representativa de todo el sistema.

Pensemos en el siguiente ejemplo:

“En 1915 Bridges describió una curiosa mutación en la *Drosophila*: la mosca mutante tenía cuatro alas en vez de dos. Una triple mutación en el gen *Ultrabithorax* produce la sustitución de unos pequeños apéndices para estabilizar el vuelo, llamados halterios, por unas alas. Estas transformaciones resultaban fascinantes, tanto por la magnitud repentina del cambio, como porque la estructura desarrollada está bien formada [...], pero aparece en un lugar equivocado. La mutación en un solo gen produce una gran transformación fenotípica, porque cambia de forma decisiva la ruta de desarrollo. La explicación es la existencia de genes selectores.” (Cadevall, 2005, p. 184)

Asimismo, en ocasiones, de manera aleatoria, ocurren “error[es] en la replicación del ADN, error[es] en el montaje de los fragmentos, accidentes de sobrecruzamiento en la recombinación” (ibid., 2005, p. 187), que pueden conllevar a una duplicación de genes.

“La duplicación en un primer momento crea redundancia, por lo que a veces los genes duplicados son eliminados sin consecuencias. Pero, cuando persisten, suelen sufrir cambios. Los cambios pueden afectar a la secuencia codificadora, que permite sintetizar las proteínas, o a los elementos reguladores de otras partes del genoma. También pueden afectar al tiempo o a la pauta de expresión de los genes. Estos cambios permiten una evolución funcional en los genes duplicados, ya sea repartiendo las funciones entre ellos o incorporando elementos que permiten nuevas pautas de expresión.” (ibid., p. 188)

Con este ejemplo podemos observar claramente que: *more is different*. Genes duplicados pueden llegar a hacer una diferencia considerable en el desarrollo de una especie.

“En la historia de la vida animal no sólo ha habido un aumento del número de genes *Hox* y con ello un incremento de la complejidad de las pautas de desarrollo, sino que en el caso de los vertebrados ha habido una duplicación a gran escala con el consiguiente incremento de complejidad.” (ibid., p. 188).

Además, podemos observar un caso claro de no linealidad en un sistema. Entendiendo aquí por sistema lineal uno en el que es posible lograr cierto entendimiento de él a partir de una especie de unión de los análisis que se hagan por separado de sus elementos o ítems constitutivos; elementos estos considerados, claro está, bajo cierto nivel de organización. Es decir, un sistema lineal es un sistema en el que se aplica la máxima: ‘*el todo es igual a la suma de sus partes*’. Por el contrario, un sistema no lineal es uno en el que bajo ningún motivo vale eso de que el todo es igual a la suma de sus partes. La información genética de un organismo, por ejemplo, no puede ser estimada a partir de las informaciones parciales que la constituyen; informaciones parciales suministradas, digamos, principalmente, por la expresión o activación de sus genes.

En el ejemplo se pone de manifiesto que modificaciones -sea por azar o por cualquier otra fuente de cambio- de genes selectores, dentro de los que destacan los famosos genes *Hox*, suelen tener mayores repercusiones en la ruta de desarrollo de un organismo que las modificaciones de otros genes. Se puede observar igualmente que la actividad de genes por separado no da lugar al desarrollo de un organismo; en cambio, la actividad conjunta de todos ellos en el proceso de desarrollo sí da lugar a que el organismo en consideración se constituya.

Otro aspecto distintivo de estos agregados extensos que presentan comportamientos distintos a los que presentan sus elementos o ítems constitutivos por separado, es que dichos comportamientos son *descentralizados*. Estos comportamientos colectivos no se dan bajo la dirección de una central de mando o de algún elemento que guíe y coordine los estados del resto de elementos o ítems que constituyen el sistema. Pensemos en una bandada de estorninos.



© Daniel Biber, Germany, Shortlist, Professional competition, Natural Word & Wildlife, 2008 Sony World Photography Award



© Daniel Biber, Germany, Shortlist, Professional competition, Natural World & Wildlife, 2008 Sony World Photography Award



© Daniel Biber, Germany, Shortlist, Professional competition, Natural World & Wildlife, 2008 Sony World Photography Award.

En las imágenes se pueden observar bellísimas formas constituidas a partir de la asociación de cientos de miles de aves, dando algunas además la impresión de que se ha formado un “superorganismo”; un superorganismo con una forma muy similar a la de sus elementos o ítems constitutivos. Este “superorganismo” presenta, podría considerarse, un comportamiento propio, “capaz” de defenderse de aves rapaces suficientemente bien; situación no tan esperable en caso de considerar a cada uno de sus elementos constitutivos por separado. Un estornino por sí solo difícilmente saldría bien librado de un ataque de un halcón, por ejemplo.

Pero, más allá de intentar ilustrar de una muy bella manera un caso concreto de la famosa máxima ‘*la unión hace la fuerza*’, se intenta, sobre todo, destacar el carácter descentralizado de estas asombrosas organizaciones. A partir de unas pocas reglas simples (como sucede por ejemplo con el famoso *Game of Life* de John Horton Conway), y sin la supervisión ni control de miembro alguno de la bandada, tiene lugar una asombrosa organización coherente y cohesiva que se mantiene estable, si bien mediante vínculos relativamente débiles en el sentido que resultan fáciles de romper. A diferencia de las características emergentes que surgen por combinación, las que se dan por asociación tienen en promedio una duración menos prolongada. En el caso de los estorninos, el comportamiento colectivo que surge a partir de su asombrosa organización se mantiene entre tanto la distancia que los separa no supere el metro. En la interacción que sostienen no hay un intercambio significativo de energía, y al cesar la interacción sus características no se ven modificadas.

En el famoso *Game of life*, como se había mencionado, con dos simples reglas se pueden generar prácticamente todo tipo de organizaciones, en este caso bidimensional, a partir de los dos estados (vivas o muertas) que experimentan sus elementos (células). Dos son también por su parte básicamente los parámetros que habría que considerar en las asombrosas

organizaciones de los estorninos: la luz y el sonido. Los estorninos se guían por la intensidad del sonido de sus vecinos y por las sombras que se generan entre ellos. Con esto se está queriendo decir que lo complejo en muchas ocasiones surge de lo simple. De lo simple repetido hasta la saciedad.

### I.3 Autoorganización

Una caracterización que se ha ofrecido de los sistemas complejos es que son sistemas que exhiben características emergentes no triviales y comportamientos autoorganizados (Mitchell, 2009, p. 13)

“Sistemas en los que comportamientos organizados surgen sin un regulador o líder interno o externo son algunas veces llamados sistemas *autoorganizados*. Debido a que simples reglas producen comportamientos complejos en cierta manera difíciles de predecir, el comportamiento macroscópico de estos sistemas es algunas veces llamado *emergente*.”<sup>21</sup> (Mitchell, 2009, p. 13).

Una pregunta que podría surgir cuando de autoorganización se trata es qué se está entendiendo por *orden*, pues intuitivamente cuando se habla de sistemas autoorganizados se está pensando en sistemas que adquieren cierto orden sin el control o la dirección de central de mando alguna. El concepto de orden no es un concepto fácil de caracterizar; sin embargo, una manera de acercarnos a él puede ser mediante el concepto de *información*.

‘*Información*’, en términos bastante amplios, suele concebirse como un ítem vinculado a cualquier medio que presente u ofrezca conocimiento: periódicos, libros, la Internet, etc. En un sentido un poco más técnico suele “ser usado para describir una amplia gama de

---

<sup>21</sup> *Systems in which organized behavior arises without an internal or external controller or leader are sometimes called self-organizing. Since simple rules produce complex behavior in hard-to-predict ways, the macroscopic behavior of such systems is sometimes called emergent.* (Mitchell, 2009, p. 13)

fenómenos que van desde la transmisión por fibra óptica que constituyen las señales de un computador a otro en la internet, hasta las pequeñas moléculas que las neuronas usan para comunicarse entre sí en el cerebro”<sup>22</sup> (Mitchell, 2009, p. 41)

La noción de información que se desprende de la teoría de Claude Shannon (1948), incluso la noción misma de entropía con la que Shannon también caracteriza a su concepto métrico (información de Shannon o entropía de Shannon), se distancian un poco, tanto de la noción intuitiva de información, como del concepto de entropía empleado por Boltzmann para el tratamiento de sus gases ideales.

“Boltzmann definió la entropía de un macroestado como una función del número de microestados que podrían dar lugar a dicho macroestado”<sup>23</sup> (Mitchell, 2009, pp. 50-51). Un macroestado es una colección de microestados que viene dada por una distribución de probabilidad sobre estos últimos; que vendrían a ser descripciones detalladas del sistema en el espacio de fases. Esto último, pensado en sistemas termodinámicos bajo el tratamiento de la mecánica estadística. Sin embargo, el análisis en términos de microestados y macroestados es extensible no sólo a esta clase de sistemas. El concepto de microestado es aplicable para la especificación de cómo están configuradas las partes de un sistema en un tiempo dado, y el de macroestado para clasificar esos microestados.

Piénsese a manera de ejemplo en un sistema compuesto por dos gases distintos confinados en una habitación. Un macroestado sería una colección de descripciones del sistema bajo un conjunto de distribuciones de probabilidad de estos dos componentes. Por ejemplo: “más del

---

<sup>22</sup> [I]t is used to describe a vast array of phenomena ranging from the fiber-optic transmissions that constitute signals from one computer to another on the Internet to the tiny molecules that neurons use to communicate with one another in the brain. (Mitchell, 2009, p. 41)

<sup>23</sup> Boltzmann defined the entropy of a macrostate as a function of the number of microstates that could give rise to that macrostate. (Mitchell pp. 50-51)

60% del *gas I* concentrado en cierto lado de la habitación”. En ella se encuentran todas aquellas descripciones del sistema en las que el *gas I* se halla bajo esas condiciones: “*el gas I* se encuentra en un 65% del lado izquierdo de la habitación”, por mencionar una; suponiendo que la habitación esté dividida verticalmente en 2 partes iguales.

Piénsese como otro ejemplo en un sistema con 125 configuraciones posibles, que consisten en secuencias de tres letras, habiendo cinco con las que formar la secuencia. Se está pensado aquí en el juego de tragamonedas en el que ganas si obtienes una secuencia de tres letras iguales, entendiendo por letra la imagen de una fruta. Hay cinco microestados en los que ganas: “limón-limón-limón”; “pera-pera-pera”; “manzana-manzana-manzana”; “naranja-naranja-naranja”; “cereza-cereza-cereza”. Un macroestado entonces de este sistema sería: “todas las imágenes iguales – tú ganas”, compuesto de cinco microestados; mientras que el macroestado: “No todas las imágenes iguales –tu pierdes” estaría compuesto de 120 microestados (Mitchell, 2009. p. 50).

El concepto de entropía entonces, desde la física estadística, puede caracterizarse simplemente como una medida de la multiplicidad de estados equivalentes, y, aunque comúnmente está asociado a la idea de desorden, no es exactamente una medida del grado de desorden de un sistema. La entropía puede perfectamente incrementarse sin un consecuente incremento de desorden. (Hidalgo, 2015, p. 23)

“Información y entropía, en el lenguaje de Shannon, son funcionalmente equivalentes, ya que el número de bits necesarios para especificar un mensaje (información de Shannon) es una función del número de posibles mensajes que podrían ser transmitidos (la multiplicidad de estados que conocemos como entropía) [...] Como ha insistido el premio Nobel de química Manfred Eigen: ‘Entropía refiere a un

promedio de estados (estados físicos); mientras que información refiere a un estado (físico) particular.”<sup>24</sup> (ibid., 2015, p. 23)

‘Información’, en los sistemas que acá nos interesa, y en general en los sistemas físicos, es opuesto a ‘entropía’ (ibid., p. 24). En ese sentido diremos que los sistemas complejos, centrándonos en este momento en el orden que exhiben, son, tanto peculiares, como poco comunes. El orden que exhiben estos sistemas es, tanto peculiar, como poco común. (ibid., p. 23). Para ilustrar un poco mejor esta idea pensemos en el siguiente ejemplo dado por César Hidalgo en su interesante libro ‘*Why Information Grows. The Evolution of Order, from Atoms to Economies*’ (2015).

Leyendo la noticia de portada de la sección de negocios de un periódico chileno, Hidalgo quedó sorprendido con el altísimo valor pagado por un auto, justamente por un compatriota suyo. El auto era, nada más y nada menos, que un Bugatti Veyron, cuyo costo superó los 2.5 millones de dólares. En un estimativo bastante rudimentario, podría decirse que el precio por kilo de este auto fue de alrededor \$1.300. Lo que, nos dice Hidalgo, sobrepasa el precio por kilo del metal plata (Ag). Ahora, ¿qué pasaría con ese valor de más de 2.5 millones de dólares si el auto desafortunadamente fuera chocado? La respuesta, nos dice Hidalgo, es obvia:

“El costo del auto (en dólares) se evaporaría en el segundo en que el auto es chocado [por ejemplo] contra un muro, pero el peso seguiría siendo el mismo. ¿Hacia dónde se fue ese valor? El valor del auto en dólares se evaporó al momento del choque, no porque el choque haya destruido los átomos de los que está constituido el Bugatti, sino

---

<sup>24</sup> So in Shannon’s language, information and entropy are functionally equivalent because the number of bits you need to specify a message (Shannon’s information) is a function of the number of possible messages that could be transmitted (the multiplicity of states, which we know as entropy). Yet this does not make entropy and information the same thing. As Manfred Eigen, the winner of the 1967 Nobel Prize in Chemistry, remarked: “Entropy refers to an average of (physical) states, information to a particular (physical) state. (Hidalgo, 2015, p. 23).

porque modificó la manera en la que estos se encontraban; [modificó cierta disposición física]. Esa disposición física es lo que se entiende por información.”<sup>25</sup> (Hidalgo, 2015, p. 20)

Decimos entonces que los sistemas complejos son sistemas con una organización peculiar y poco común porque sus elementos constituyentes presentan disposiciones físicas con un altísimo grado de información, cuyas rutas de acceso para llegar a ellas, ateniéndose a las constricciones, básicamente fisicoquímicas, que operan sobre tales elementos, son extremadamente limitadas. Estos estados ricos informacionalmente se dan en extremadamente muy pocos microestados, por lo que la entropía de estos estados es bastante baja.

Pensemos, por ejemplo, en una proteína. Sabemos que es una configuración específica de su secuencia de aminoácidos la que la hace funcional. El espectro de configuraciones o plegamientos posibles, dadas las constricciones fisicoquímicas que operan sobre los aminoácidos constitutivos de una proteína, es extremadamente amplio. Sin embargo, de todas esas posibles configuraciones, o de todas esas posibles maneras en las que podrían plegarse los aminoácidos, la funcionalidad de la proteína está reducida a una configuración particular. Decimos entonces que es esa configuración particular la que contiene un alto contenido de información. Es rica en información.

Acá, el concepto de *complejidad efectiva* se nos presenta como uno de gran utilidad. Murray Gell-Mann define el concepto de complejidad efectiva como “la longitud de una

---

<sup>25</sup> *The dollar value of the car evaporated in the seconds it took you to crash it against that wall, but its weight did not. So where did the value go? The car's dollar value evaporated in the crash not because the crash destroyed the atoms that made up the Bugatti but because the crash changed the way in which these were arranged. As the parts that made the Bugatti were pulled apart and twisted, the information that was embodied in the Bugatti was largely destroyed. This is another way of saying that the \$2.5 million worth of value was stored not in the car's atoms but in the way those atoms were arranged. That arrangement is information.* (Hidalgo, 2015, p. 20)

descripción concisa de las *regularidades* [de un sistema]” (Gell-Mann, 1998, p. 66). Es decir, se está pensando aquí en una descripción comprimida de las relaciones conspicuas que los elementos constituyentes del sistema presentan en esos estados ricos informacionalmente. Se está pensando en la descripción más concisa del conjunto de patrones identificables entre las interacciones que estos elementos presenten.

Así, con lo dicho hasta acá, podemos decir que los sistemas que exhiben comportamientos autoorganizados son sistemas que presentan una disposición de sus elementos constitutivos con alta complejidad efectiva y baja entropía; son sistemas cuyos elementos constituyentes adquieren cierta disposición física de manera espontánea con un alto contenido de información. El comportamiento autoorganizado de sistemas complejos es uno en cuya disposición de los elementos que dan lugar a él podemos encontrar ciertos patrones, pero también cierta dosis de aleatoriedad. Curiosamente, estos estados se suelen dar en lo que se conoce como el borde del caos (*edge of chaos*). Y curiosamente también, ahí encuentran una suerte de equilibrio entre la flexibilidad y la rigidez de sus comportamientos que les resultan indispensables para adaptarse.

#### I.4. *Adaptación*

En la tercera parte de *‘El Quark y el Jaguar’* (1998), Gell-Mann se dedica a describir aquellos eventos que permiten la emergencia de nuevos niveles de organización. El biofísico Harold Morowitz ha introducido el término *‘sucesos umbral’* para hacer referencia a esta clase de eventos, cuya característica más notoria es que ocurren al azar; entendido principalmente en términos de coincidencia - *concoirs de circonstances* (véase la sección 3.2 del Cap. II). Según la definición convencional, un suceso umbral es el tipo de evento cuya ocurrencia produce el surgimiento de procesos o sistemas emergentes (Gell-Mann, 1998, p.

258). Los sucesos umbrales de más interés son los que posibilitan la emergencia de patrones generales (o regularidades nómicas), y Gell-Mann se refiere a tales sucesos umbrales bajo el rótulo de '*accidentes congelados*'.

En el curso de la evolución biológica podemos encontrar varios ejemplos de accidentes congelados. Uno muy difundido es el del origen de las células eucariotas. Hoy se acepta que en su origen un suceso muy importante (suceso umbral) fue la unión simbiótica de dos o más organismos primitivos. Antes de que esta unión simbiótica se diese, estos organismos primitivos podían replicarse de manera independiente. Una vez se dio la unión, se congeló la situación, con una enorme contribución de la selección natural, y ahora sólo pueden hacerlo cuando lo hace la célula entera. (Smith y Szathamary, 2001, p. 40). Otro ejemplo de accidente congelado en el proceso evolutivo es el del origen del sexo. Después de su origen, “los individuos sólo pueden reproducirse como miembros de una población sexual, mientras que antes podían reproducirse asexualmente, por su cuenta” (ibid., 2001, p. 40)

Estos accidentes congelados no sólo permitieron el origen de ciertas estructuras y organismos, sino que han contribuido a que las especies biológicas aumenten su complejidad efectiva. Lo que resulta de esto es que las presiones selectivas del proceso evolutivo han favorecido a los organismos más complejos. Sin embargo, ¿es esta magnitud suficiente para determinar la complejidad de una especie? La respuesta que ofrece Gell-Mann es que no. Esta magnitud debe ser complementada con la noción de complejidad potencial; cuyo caso más representativo se encuentra quizás en el caso de los humanos. Se considera que “los cambios genéticos que [han permitido] a una criatura simiesca desarrollar lenguajes, pensamiento avanzado y culturas elaboradas, manifestaciones todas de gran complejidad efectiva, son relativamente pequeños, pero de gran significación” (Gell-Mann, 1998, p. 87).

Pequeñas alteraciones en su esquema han permitido el despliegue de un amplio abanico de posibilidades para el aumento de su complejidad efectiva en un intervalo de tiempo dado. Quiere decir entonces que su material genético ha sido lo suficientemente potente para aprovechar las posibilidades que se le han presentado para el aumento de su complejidad efectiva, permitiéndole ser quizás el organismo que más control ejerce sobre su medio, bien sea para bien o para mal.

Pero un accidente congelado no sólo genera el surgimiento de regularidades nómicas en los niveles de organización a los que da cabida, sino que posibilita las condiciones adecuadas para el desarrollo de nuevos niveles de organización. En otros términos: los niveles que emergen de un accidente congelado pueden ser el sustrato a partir del cual emerjan sistemas pertenecientes a nuevos niveles de organización, tal como en el caso de la emergencia de las células eucariotas a partir de dos células primitivas. De acuerdo con Gell-Mann, en casos como el de la emergencia de las eucariotas se puede hablar de la aparición de un nivel de organización superior, lo que plantea la complicada cuestión de cómo definir la superioridad relativa de un sistema. Gell-Mann define la superioridad relativa de cualquier nivel de organización en términos de la complejidad efectiva: cuanto mayor sea la complejidad efectiva tanto mayor será la superioridad relativa.

Los sistemas biológicos, en virtud de la dinámica del proceso evolutivo, son en gran medida el resultado de una combinación de azar (en forma de accidentes congelados) y causalidad. Por esta razón no debe causar sorpresa que Gell-Mann coincida con Maynard Smith y Szathmáry en aquello de que la selección natural tiende a favorecer un incremento en la complejidad efectiva de los sistemas biológicos. Si estos autores tienen razón, no existen motivos para rechazar de antemano la posibilidad de que la evolución exhiba tendencias progresivas. Sobre esto se discutirá un poco más en la sección 2 del capítulo II.

### I.5. Estructura jerárquica

En el caso de sistemas con sistema nervioso central principalmente, la información se procesa de manera jerárquica, en el sentido de que se da de manera “ascendente” a través de una estructura de niveles; estructura en la que, si se observa en términos netamente formales, se da una relación de dominación o de subordinación entre los niveles (subsistemas) que conforman la estructura; donde cada subsistema viene constituido por elementos equivalentes bajo algún respecto, es decir, que comparten sus propiedades básicas y están sujetos a las mismas leyes. (Bunge, 1969, p. 18). Una característica además fundamental de los sistemas complejos es que cada nivel que conforma la estructura de la jerarquía involucra emergencia o novedad cualitativa (ibid., 1969, p. 18). A estos niveles se les suele llamar también *niveles de organización*.

En un ya clásico artículo sobre los sistemas complejos: ‘*The Architecture of Complexity*’ (1962), Herbert Simon insiste en la importancia de la arquitectura jerárquica para el surgimiento, sostenimiento y aumento de complejidad en ciertos sistemas. Los sistemas que disponen de esta arquitectura obtienen, permítaseme la expresión, cierta rentabilidad, sobre todo energética e informacional, de los recursos empleados en sus subsistemas. Quizás el quid del asunto esté en los tipos de interacciones que se dan, *dentro y fuera*, de los subsistemas (niveles). Las interacciones que se dan entre los elementos o ítems que conforman un subsistema son mucho más fuertes e intensas que las que se dan entre subsistemas. Claramente los subsistemas se hallan en interacción constante y de estas depende en gran medida la estabilidad global del sistema. Pero estas interacciones reposan, o se cimentan, en estados de equilibrio de la dinámica que cada subsistema experimenta. En el caso por ejemplo del procesamiento de información que ocurre en el neocórtex de los

mamíferos esta se da por capas o niveles, donde cada capa o nivel genera una especie de *output* estable que constituirá el *input* de la capa subsiguiente.

Los estados de equilibrio de cada subsistema, o nivel, representan, sobre todo en materia de optimización de procesos (por ejemplo, reducción de error en el envío de un mensaje; aprovechamiento de recursos energéticos; entre otros) una suerte de pista sobre qué rutas en el espacio de estados “transitar”. Sobre este punto, un ejemplo muy básico, que podría ilustrar la situación antes descrita, puede ser el siguiente:

“Supongamos que se desea abrir una caja fuerte cuya cerradura tiene 10 ruedecillas o diales, cada uno de estos con cien posiciones o ajustes posibles, numerados del 0 al 99 ¿Cuánto tiempo tardaríamos en abrir la caja fuerte buscando ciegamente la combinación correcta por ensayo y error? Puesto que de entrada existen  $100^{10}$  combinaciones posibles, podemos estimar que en promedio tendríamos que comprobar la mitad de ellas antes de encontrar la combinación correcta, esto es, cincuenta trillones de combinaciones distintas. Supongamos, sin embargo, que la caja fuerte es defectuosa, de modo que podemos oír un *click* cuando cada una de las ruedecillas llega a su posición correcta. En este caso, cada una de las ruedecillas se puede ajustar de manera independiente y no hay necesidad de volverlas a tocar mientras se ajustan las demás. El número total de ajustes o combinaciones que tienen que intentarse es solo  $50 \times 10$ , es decir, 500. La tarea de abrir la caja fuerte ha pasado, por las pistas que los *clicks* proporcionan, de ser prácticamente imposible a ser una trivial.”<sup>26</sup> (Simon, 1962, p. 472)

---

<sup>26</sup> *Suppose that the task is to open a safe whose lock has ten dials, each with one hundred possible settings, numbered from 0 to 99. How long will it take to open the safe by a blind trial-and-error search for the correct setting? Since there are  $100^{10}$  possible settings, we may expect to examine about one-half of these, on the average, before finding the correct one—that is, fifty billion billion settings. Suppose, however, that the safe is defective, so that a click can be heard when any one dial is turned to the correct setting. Now each dial can be adjusted independently, and does not need to be touched again while the others are being set. The total number*

Lo que se está intentando mostrar es que una arquitectura jerárquica, favorecida por la selección natural, ha representado una ventaja de ciertos sistemas frente a otros en materia de optimización de procesos. Podría decirse además que la complejidad es producto, entre otras cosas, de una explotación, favorecida por los mecanismos de evolución biológica, de las distintas posibilidades de combinación y disposición de los subsistemas de ciertos sistemas. La evolución biológica juega con lo ya existente para generar nuevas formas. La complejidad es, hasta cierto punto, un producto de nuevas disposiciones y combinaciones de estructuras ya existentes.

“Un ejemplo familiar son las proteínas, cuya enorme variedad surge de la recombinación de sólo veinte aminoácidos distintos. De manera similar, los noventa y tantos elementos proporcionan todos los bloques de construcción necesarios para constituir la infinita variedad de moléculas. Por tanto, podemos construir nuestra descripción a partir de un alfabeto limitado de términos elementales que corresponden al conjunto básico de subsistemas elementales a partir de los cuales se forman los sistemas complejos.”<sup>27</sup> (ibid., p. 478)

Con relación, por ejemplo, al procesamiento de información que se da en nuestro neocórtex, Vernon B. Mountcastle, en un artículo publicado en 1978 bajo el título “*An Organizing Principle for Cerebral Function*”, nos planteaba que:

“el neocórtex es notablemente uniforme, tanto en apariencia como en estructura.

Presenta las mismas capas, tipos de células y conexiones. Las diferencias son tan sutiles

---

*of settings that has to be tried is only 10 X 50, or five hundred. The task of opening the safe has been altered, by the cues the clicks provide, from a practically impossible one to a trivial one. (Simon, 1962, p. 472)*

<sup>27</sup> *A familiar example is the proteins, their multitudinous variety arising from arrangements of only twenty different amino acids. Similarly, the ninety-odd elements provide all the kinds of building blocks needed for an infinite variety of molecules. Hence, we can construct our description from a restricted alphabet of elementary terms corresponding to the basic set of elementary subsystems from which the complex system is generated. (Simon, 1962, p. 478)*

que ni los anatomistas entrenados logran ponerse de acuerdo respecto a ellas [...] Todas las regiones de la corteza realizan la misma operación. Lo que hace [, por ejemplo, que cierta área de la corteza sea considerada como un] área de visión y otra como un área motora, es la manera en la que están conectadas entre sí y con otras regiones del sistema nervioso central. [...] Existe una función común, un algoritmo común que es desplegado por todas las áreas de la corteza.”<sup>28</sup> (Hawkins y Blakeslee, 2004, pp. 35, 36)

El procesamiento de información que se da en el proceso de visión no difiere del que se da en el proceso de audición, que tampoco difiere del que se da a nivel motor. El algoritmo es el mismo en todas las áreas de la corteza. (ibid., 2004, p. 36).

“Pensemos en esto por un momento. Para mí, la vista, el oído y el tacto resultan muy diferentes. Tienen cualidades fundamentalmente diferentes. La vista implica color, textura, perfil, profundidad y forma. La audición tiene tono, ritmo y timbre. Se sienten muy diferentes. ¿Cómo pueden ser iguales? Mountcastle dice que no son lo mismo, pero la forma en que la corteza procesa las señales del oído es la misma que la forma en que procesa las señales de los ojos. Continúa diciendo que el control motor también funciona con el mismo principio.”<sup>29</sup> (ibid., p. 36)

---

<sup>28</sup> *In a field of anatomists looking for minute differences in cortical regions, he shows that despite the differences, the neocortex is remarkably uniform. The same layers, cell types, and connections exist throughout. It looks like the six business cards everywhere. The differences are often so subtle that trained anatomists can't agree on them. Therefore, Mountcastle argues, all regions of the cortex are performing the same operation. The thing that makes the vision area visual and the motor area motoric is how the regions of cortex are connected to each other and to other parts of the central nervous system.*

*He concludes that there is a common function, a common algorithm, that is performed by all the cortical regions.* (Hawkins y Blakeslee, 2004, pp. 35, 36).

<sup>29</sup> *Let's think about this for a moment. To me, sight, hearing, and touch seem very different. They have fundamentally different qualities. Sight involves color, texture, shape, depth, and form. Hearing has pitch, rhythm, and timbre. They feel very different. How can they be the same? Mountcastle says they aren't the same, but the way the cortex processes signals from the ear is the same as the way it processes signals from the eyes. He goes on to say that motor control works on the same principle, too.* (Hawkins y Blakeslee, 2004, p. 36)

Este algoritmo, que es desplegado en todas las áreas de la corteza, es uno en el que la información se procesa de manera jerárquica. Sobre este punto Hawkins y Blakeslee nos plantean que, en un sistema jerárquico, algunos elementos, en sentido abstracto, se encuentran “arriba” o “abajo” con relación a otros elementos, pero que esto nada tiene que ver con su posición física. Nada tiene que ver con la disposición física en el cerebro. Todas las áreas funcionales de la corteza se ubican en la misma lamina cortical ultra plegada. Lo que hace que un área se considere por “encima” o por “debajo” de otra en la jerarquía es la manera en la que se conectan con otras áreas. En la corteza, las áreas que se estiman “bajas”, alimentan de información a las áreas consideradas “superiores”. Esto, a través de cierto patrón de conectividad. (ibid., p. 32)

“Las regiones funcionales más bajas, las áreas sensoriales primarias, son aquellas donde la información sensorial llega por primera vez a la corteza. Estas regiones procesan la información en su nivel más básico y crudo. Por ejemplo, la información visual ingresa a la corteza a través del área visual primaria, llamada V1, para abreviar. V1 se ocupa de características visuales de bajo nivel, como pequeños segmentos de borde, componentes de movimiento a pequeña escala, disparidad binocular (para visión estereoscópica) e información básica de color y contraste. V1 alimenta informacionalmente a otras áreas, como V2, V4 e IT, además de muchas otras áreas. Cada una de estas áreas se ocupa de aspectos más especializados o abstractos de la información. Por ejemplo, las células en V4 responden a objetos de complejidad media, como formas de estrellas en diferentes colores, como rojo o azul. Otra área llamada MT se especializa en los movimientos de objetos. En los niveles superiores de la corteza visual hay áreas que representan sus recuerdos visuales de todo tipo de objetos, como rostros, animales, herramientas, partes del cuerpo, etc.

[Nuestros] otros sentidos tienen jerarquías similares. [La] corteza tiene un área auditiva primaria llamada A1, y una jerarquía de regiones auditivas por encima de ella; asimismo, tiene un área somatosensorial primaria (sentido corporal) llamada S1, y una jerarquía de regiones somatosensoriales por encima de ella. Finalmente, la información sensorial pasa a “áreas de asociación”, que es el nombre que a veces se usa para las regiones de la corteza que reciben entradas de más de un sentido. Por ejemplo, [la] corteza tiene áreas que reciben información tanto de la visión como del tacto. Es gracias a las regiones de asociación que [podemos] ser consciente de que la vista de una mosca trepando por [nuestro] brazo y la sensación de cosquilleo que allí se siente comparten la misma causa. La mayoría de estas áreas reciben información altamente procesada de varios sentidos, y sus funciones siguen sin estar claras.”<sup>30</sup>(Hawkins y Blakeslee, 2004, pp. 32, 33)

Con lo anterior queda evidenciado que el procesamiento de información que ocurre en nuestro neocórtex se da de manera jerárquica. Pero si queremos dar una descripción un poco más detalladas de dicho procesamiento debemos, sin duda alguna, destacar su carácter

---

<sup>30</sup> *The lowest of the functional regions, the primary sensory areas, are where sensory information first arrives in the cortex. These regions process the information at its rawest, most basic level. For example, visual information enters the cortex through the primary visual area, called V1 for short. V1 is concerned with low-level visual features such as tiny edge-segments, small-scale components of motion, binocular disparity (for stereo-vision), and basic color and contrast information. V1 feeds information up to other areas, such as V2, V4, and IT (we'll have more to say about them later), and to a bunch of other areas besides. Each of these areas is concerned with more specialized or abstract aspects of the information. For example, cells in V4 respond to objects of medium complexity such as star shapes in different colors like red or blue. Another area called MT specializes in the motions of objects. In the higher echelons of the visual cortex are areas that represent your visual memories of all sorts of objects like faces, animals, tools, body parts, and so on.*

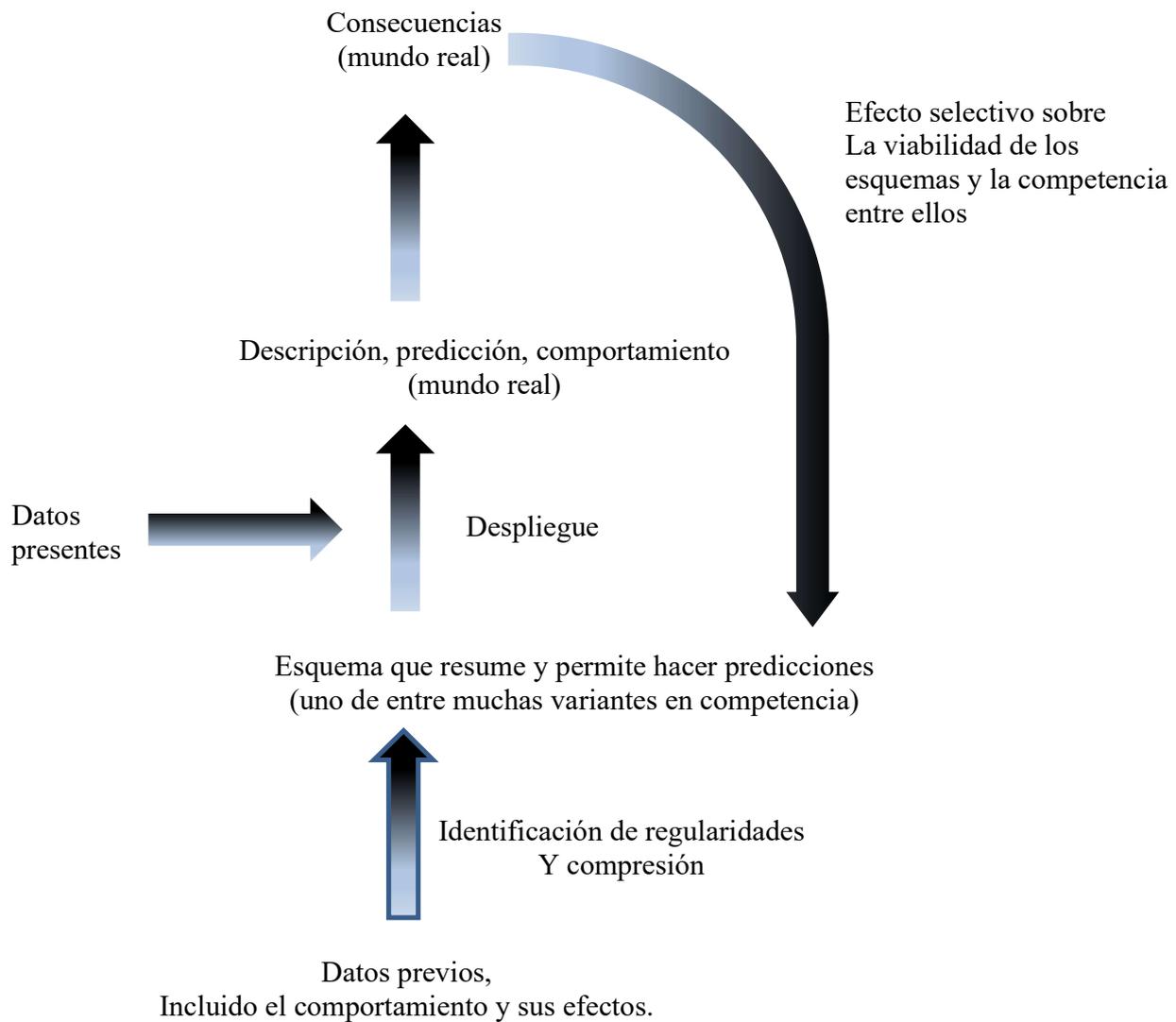
*Your other senses have similar hierarchies. Your cortex has a primary auditory area called A1 and a hierarchy of auditory regions above it, and it has a primary somatosensory (body sense) area called S1 and a hierarchy of somatosensory regions above that. Eventually, sensory information passes into “association areas”, which is the name sometimes used for the regions of the cortex that receive inputs from more than one sense. For example, your cortex has areas that receive input from both vision and touch. It is thanks to association regions that you are able to be aware that the sight of a fly crawling up your arm and the tickling sensation you feel there share the same cause. Most of these areas receive highly processed input from several senses, and their functions remain unclear. (Hawkins & Blakeslee, 2004, pp. 32, 33)*

retroalimentativo. La retroalimentación (*feedback*) es una característica distintiva, tanto de los sistemas estimados complejos en general, como de nuestro cerebro, en particular.

#### I.6. *Feedback (retroalimentación)*

La retroalimentación es otro de los aspectos distintivos de los sistemas estimados complejos. En la sección de '*Adaptación*' se había planteado que los sistemas complejos son sistemas que procesan información, tanto interna como del medio que los rodea, a fin de poder sortear sus presiones selectivas. Este procesamiento de información consiste principalmente en la identificación de patrones o regularidades a distintos niveles, tanto internos como externos. Patrones que se encapsulan en una suerte de esquemas o representación simbólica, puestos a prueba en dicho intento por sortear las presiones selectivas que operen sobre el sistema en consideración. Estos esquemas, o representaciones simbólicas, en los que se encapsulan las regularidades que el sistema logra percibir de su entorno, se encuentran en constante cambio, sea a corto, mediano o largo plazo. Si el sistema logra solventar sus presiones selectivas, el esquema, o los esquemas desplegados, se alimentan positivamente; en el sentido que ingresa en él, o en ellos, información general acerca de ellos mismos, que refuerza su despliegue en situaciones análogas a las presentadas en el o los despliegues previos. Cabe adicionar, como hecho importante, que en estos esquemas ocurren también cambios producto de factores de azar; además de ajustes por parte del propio sistema y por parte también de mecanismos evolutivos.

A continuación, diagramaremos el funcionamiento de un sistema complejo adaptativo según Murray Gell-Mann (1998, p. 41)



Como se puede observar, las consecuencias del despliegue de los esquemas en el entorno circundante del sistema generan un efecto retroactivo, sea positivo o negativo, sobre dichos esquemas. Quiere decir esto que información saliente (*output*) pasa a ser, en alguna etapa del proceso, información entrante (*input*), generando así una especie de ciclo, que deja una huella informacional que se puede ir haciendo cada vez más pronunciada en caso de que el efecto retroactivo sea positivo. La retroalimentación puede verse como un regreso por el mismo sendero informacional siguiendo una especie de huellas o marcas dejadas metafóricamente

en el camino de ida. Al recorrer el mismo camino informacional de vuelta la ruta transitada se remarca. Esta especie de huella o ruta remarcada podría verse como una especie de atractor informacional para rutas de procesamiento aledañas. Pensemos, por ejemplo, en las redes neuronales del cerebro humano.

“En el nivel celular, el cerebro es una red gigantesca de células nerviosas interconectadas por pasajes microscópicos de axones y dendritas. Un destello de la actividad cerebral -pensar una palabra, elaborar un concepto, desbrozar la sintaxis de la frase que estamos leyendo- dispara una serie de circuitos neuronales como rutas de tránsito en el mapa de la mente.

Cada nueva actividad mental desencadena una serie nueva, y una cantidad inimaginable de posibles circuitos jamás llegará a realizarse en el transcurso de una vida humana [...] Pero bajo toda esta aparente diversidad, ciertos circuitos se repiten una y otra vez. Una de las hipótesis más provocadoras de la neurociencia actual es que la base celular del aprendizaje reside en la repetición de esos circuitos. [...] Cualquier circuito podría asociarse inicialmente con la imagen de un sándwich o con la forma de un triángulo isósceles; con la suficiente repetición, un circuito específico ocuparía un espacio fijo en el cerebro y desde ese momento pasaría a ser parte de nuestro vocabulario mental.

¿Por qué ocurren estos circuitos de retroalimentación y de reverberación? Porque las redes neuronales del cerebro están densamente interconectadas: cada neurona individual contiene vínculos -en la forma de axones y sinapsis- con otras miles de neuronas. Cuando una neurona es estimulada transmite su descarga a todas esas otras células, quienes a su vez, dadas determinadas condiciones, descargan su estímulo sobre

sus conexiones celulares, y así sucesivamente. Si cada neurona estuviera vinculada a una o dos neuronas vecinas, las posibilidades de establecer de reverberación se verían enormemente reducidas. Pero dado que las neuronas transmiten su estímulo simultáneamente en tantas direcciones, es muy probable que cuando una neurona dispara la transmisión, esta vuelve a la fuente original y el proceso comience de nuevo. La ocurrencia de un circuito de retroalimentación se relaciona directamente con la interconexión general del sistema.” (Johnson, 2003, p. 120)

Se puede observar entonces que la retroalimentación juega un papel importante en el procesamiento de información que ocurre en nuestro cerebro. Los sistemas descentralizados, como algunos de los que hemos presentado hasta el momento, dependen de la retroalimentación, “tanto para su crecimiento como para su regulación” (ibid., 2003, p. 119). De hecho, en el caso de nuestro cerebro, un sistema altamente descentralizado, la retroalimentación es altamente preponderante.

“Los neuroanatomistas han sabido durante mucho tiempo que el cerebro está saturado de conexiones de retroalimentación. Por ejemplo, en el circuito entre el neocórtex y una estructura inferior llamada tálamo, ¡las conexiones que van hacia atrás (hacia la entrada) exceden las conexiones que salen en casi un factor de diez! Es decir, por cada fibra (nervio) de alimentación de información hacia el neocórtex, hay diez fibras que alimentan información hacia los sentidos. La retroalimentación también domina la mayoría de las conexiones a lo largo del neocórtex. Nadie entendió el papel preciso de esta retroalimentación, pero de la investigación publicada quedó claro que

existía en todas partes. Pensé que debía ser importante.”<sup>31</sup> (Hawkins y Blakeslee, 2005, p. 18)

Terminamos este capítulo habiendo presentado algunas de las características que se estiman distintivas de los sistemas complejos. Dar cuenta de estas resulta, en la mayoría de los casos, actualmente posible desde uno o más esquemas conceptuales. Así, por ejemplo, frente a la característica sumamente distintiva de estos sistemas de estar constituidos de muchísimas partes, se tiene la *estadística y la teoría de la probabilidad* (Cap.II.3). Se había dicho que los elementos constitutivos de estos sistemas interactúan de manera no lineal; por lo que, con *dinámica no lineal y teoría del caos* (Cap. II.1) se puede dar cuenta de este aspecto representativo de los sistemas complejos. Con *teoría sintética de la evolución* (Cap.II.2.3) podría darse cuenta de la adaptación. Por su parte, la teoría de fenómenos críticos podría ayudarnos a dar cuenta de los comportamientos colectivos. La teoría de control nos permitiría abordar la retroalimentación. La teoría de redes podría ayudarnos a abordar las estructuras jerárquicas.

Ninguno de estos esquemas conceptuales o teorías tiene como pretensión expresa dar cuenta de la complejidad en sus múltiples manifestaciones y rasgos distintivos. Son más bien esquemas o teorías concretas que apuntan inicialmente hacia algunos rasgos particulares que se han logrado identificar hasta el momento (Mitchell, 2009) como distintivos de los sistemas

---

<sup>31</sup> *Neuroanatomists have known for a long time that the brain is saturated with feedback connections. For example, in the circuit between the neocortex and a lower structure called the thalamus, connections going backward (toward the input) exceed the connections going forward by almost a factor of ten! That is, for every fiber feeding information forward into the neocortex, there are ten fibers feeding information back toward the senses. Feedback dominates most connections throughout the neocortex as well. No one understood the precise role of this feedback, but it was clear from published research that it existed everywhere. I figured it must be important.* (Hawkins & Blakeslee, 2005, p. 18)

complejos<sup>32</sup>. Sin embargo, podrían considerarse ciertamente como una suerte de aproximaciones teóricas acerca de esta clase de sistemas, pues a lo menos nos ayudan a comprender algunas de sus características particulares; si bien, de manera aislada, no integrada. De ahí que le llamemos “aproximaciones”. Esto es, son teorías o esquemas que permiten un abordaje parcial de la complejidad.

Justamente, en el próximo capítulo, hablaremos sobre algunos de estos marcos. Quizás destacando el marco de la teoría sintética de la evolución, al ser este un proceso donde se ha evaluado la posibilidad de que un aumento de complejidad se haya visto favorecido por su dinámica. Así mismo, se hablará de dinámica no lineal y caos, y teoría de la probabilidad y estadística, pues estos esquemas, además de constituir por sí mismos aproximaciones valiosas, si bien parciales, para el estudio de sistemas complejos, se compaginan en una interesante propuesta de abordaje de las propiedades emergentes.

---

<sup>32</sup> Si se tuviese certeza de que los rasgos presentados como distintivos de los sistemas complejos fuesen solo esos, pues a partir de ellos podría ciertamente aventurarse un acercamiento definitorio de la complejidad. Podría definirse, tentativamente, como la manifestación en su sistema concreto de justamente estas características. La dificultad radica en que aún hay aspectos desconocidos de esta. En el caso, por ejemplo, de nuestro cerebro, el que sea quizás el sistema paradigmático de la complejidad, es claro que aún desconocemos algunos aspectos. Básicamente, lo que se está intentando decir es que no sabemos si la complejidad se agota en estas características.

## CAPÍTULO II

### APROXIMACIÓN TEÓRICA A LOS SISTEMAS COMPLEJOS

#### *¿Es posible una métrica para la complejidad?*

##### Abstract

En este capítulo se presentan algunos de los marcos teóricos propuestos para explicar diferentes aspectos del fenómeno de la complejidad. Se discutirán la dinámica no lineal y caos, la teoría sintética de la evolución y la teoría de la probabilidad y estadística. Se otorga un papel fundamental a la modelización de la complejidad emergente modelo-paramétrica.

En este capítulo se presentarán algunos de los marcos conceptuales o teorías de los que se disponen para dar cuenta de algunas características distintivas de los sistemas complejos. Se hablará entonces de *dinámica no lineal y caos*, de *teoría sintética de la evolución*, de *teoría de la probabilidad y estadística*; pero, sobre todo, se hablará de *complejidad modelo paramétrica*. Ciertamente, disponemos de algún, o algunos marcos conceptuales o teorías con los que abordar aisladamente estas características que se han planteado como propias de los sistemas complejos. Sin embargo, frente a la emergencia, aún no se dispone de un esquema o teoría comúnmente aceptada con la que se pueda dar completa cuenta de ella. Siendo esto así, el escenario parece el ideal para intentar ofrecer alguna alternativa para su abordaje; que es justamente por ejemplo lo que Miguel Fuentes (2018) ha pretendido. El concepto de '*emergencia modelo paramétrica*' representa un gran esfuerzo por presentar una noción con la que abordar, principalmente en términos métricos, la emergencia. Esfuerzo

que, valga la pena resaltar, habrá de intentar también quien escribe estas líneas, mediante el auxilio concreto de las redes neuronales artificiales de aprendizaje profundo.

### II.1. *Dinámica no lineal y caos*

En 1643 nació, en la localidad de Woolsthorpe, condado de Lincolnshire, el gran Sir Isaac Newton; padre del cálculo, junto con Leibniz, y quizás el referente más significativo de lo que se ha dado en llamar *Mecánica clásica*. Con Newton se da inicio al área del conocimiento que se conoce como '*Dinámica*'. Newton es el padre de la dinámica.

La mecánica se divide en dos áreas: la cinemática, que describe cómo las cosas se mueven, y la dinámica, que explica por qué las cosas obedecen las leyes de la cinemática. Por ejemplo, las leyes de Kepler son leyes cinemáticas -ellas describen *cómo* los planetas se mueven (en elipses con el sol como uno de los focos), pero no *por qué* se mueven de esa forma particular. Las leyes de Newton son los cimientos de la dinámica: ellas explican el movimiento de los planetas, y todo lo demás, en términos de las nociones básicas de 'fuerza' y 'masa'.

Las tres famosas leyes de Newton son las siguientes:

1. Movimiento constante: Todo objeto no sujeto a una fuerza se mueve con velocidad constante.
2. Masa inercial: Cuando un objeto está sujeto a una fuerza, el cambio resultante en su movimiento es inversamente proporcional a su masa.
3. Fuerzas iguales en sentido opuesto: Si un objeto A ejerce una fuerza sobre el objeto B, entonces el objeto B ejercerá esa misma fuerza en sentido opuesto sobre el objeto A. (Michelle, 2009, p. 19).

Con estas tres leyes, más la ley de gravitación universal, se tiene una teoría que permite describir, explicar y predecir los cambios en el tiempo de ciertos sistemas que exhiben comportamientos macroscópicos, productos de la interacción de sus partes (ibid., 2009, p. 15); dando así una sensación, al menos en principio, de una suerte de control sobre la naturaleza. El nivel de aparente control que se generó a partir de la teoría de sistemas dinámicos propuesta por Newton dio margen a que se instaurase una visión mecanicista del universo. El universo pasó a ser visto como una especie de gran máquina puesta a andar por las leyes de Newton; como un gran reloj al que se le da cuerda con dichas leyes. Sobre este punto es bien conocida la consideración hecha por el gran matemático francés Pierre Simon Laplace, quien consideró que:

Dadas las leyes de Newton, y conociendo la posición y velocidad en un momento dado de cada partícula del universo, era posible, en principio, predecir a partir de ahí todo a cada momento. “Con la invención de las calculadoras electrónicas en 1940, lo que parecía ‘en principio’ [según Laplace], lucía mucho más cerca de ‘en la práctica’.”<sup>33</sup> (ibid., p. 19). Sobre esta distinción entre ‘en principio’ y en ‘la práctica’ se hablará en el capítulo IV, pues juega un papel importante en la discusión filosófica sobre los modelos y simulaciones computacionales.

Teniendo en cuenta que gran parte de las ecuaciones con las que se modelan los distintos comportamientos y cambios en el tiempo de estos sistemas dinámicos suelen ser ecuaciones diferenciales, muchas veces parciales, el poder de cálculo que se lograba con las computadoras electrónicas representaba la herramienta necesaria para alcanzar las consideraciones laplacianas. Pues estos sistemas de ecuaciones generalmente no se resuelven

---

<sup>33</sup> *With the invention of electronic computers in the 1940s, the “in principle” might have seemed closer to “in practice.”*. (Mitchell, 2009 p. 19)

de manera analítica; se resuelven mayormente mediante métodos de aproximación numérica, en los que comúnmente se requiere una iteración<sup>34</sup> considerable de pasos en los que además se suelen considerar varios decimales para que la aproximación sea un poco más precisa. Esta situación representa uno de los ejes principales en la discusión que se plantea acerca de los modelos y simulaciones computacionales en la investigación científica.

Sin el poder de cálculo de dichas computadoras la resolución de estas ecuaciones sería en extremo difícil, por no decir inalcanzables. Y ni hablar, por ejemplo, de la enorme ventaja que significa poder *visualizar*<sup>35</sup> gráficamente el comportamiento de las soluciones, conforme se varíe el valor de los parámetros considerados; o poder lograr una suerte de visualización gráfica, vía el modelo utilizado, del comportamiento del propio sistema dinámico. Sin duda, el panorama se presentaba como uno bastante prometedor. Aunque, no sobra aclarar que una simulación computacional puede no involucrar la visualización. Asimismo, una visualización del modelo de datos puede no involucrar ecuaciones con capacidad de manipulación virtual para los resultados<sup>36</sup>.

---

<sup>34</sup> Es importante resaltar que igualmente es posible tener una simulación que arroje resultados aproximados pero su método numérico no ser iterativo; por ejemplo, utilizando series de Fourier para modelar una curva, estableciendo cuántos términos de la serie se usarán, pero no usando un resultado anterior como insumo para el cálculo siguiente.

<sup>35</sup> La posibilidad de visualización gráfica representa un nuevo tipo de indicador de gran valor epistémico en el marco de los modelos y simulaciones computacionales. El estudio de sistemas complejos y las nuevas herramientas que se vienen desarrollando en los distintos intentos para su abordaje, entre ellas, por ejemplo, las redes neuronales artificiales, nos invitan seriamente a discutir algunas consideraciones filosóficas sobre la investigación científica; sobre todo, en materia de metodología y justificación. Los modelos y simulaciones computacionales rompen de cierta manera con algunas tradiciones filosóficas acerca de la investigación científica. Sobre esto se hablará en el capítulo IV.

<sup>36</sup> Un ejemplo podría ser un mapa de contaminación sonora en una ciudad, que muestra las zonas en rojo con más ruido y las zonas más blancas con menos ruido, pero no es una simulación computacional, sino que se trata de la geolocalización de los datos.

Por el otro lado, el cálculo de la velocidad que adquiere una sonda espacial al pasar entre los anillos de Saturno (sin choques) puede ser el resultado de una simulación computacional sin necesidad de transformar esos cálculos aproximativos en imágenes de la dinámica del sistema que se esté estudiando.

Sin embargo, curiosamente, una computadora electrónica fue protagonista en uno de los episodios que representaron un duro golpe para la visión mecanicista de la naturaleza. Una ‘*Royal McBee LGP-30*’ nos puso abiertamente de manifiesto que hay sistemas en extremo sensibles a pequeñas variaciones en las condiciones iniciales. Cuando Edward Norton Lorenz, en lugar de usar seis decimales en su modelo de sistema meteorológico, como lo había hecho previamente, usó sólo tres, a modo de aproximación o redondeo, no pensó que esta ligera variación en los valores ingresados, como condiciones iniciales a su sistema de ecuaciones, habría de tener mayores repercusiones. Y, en principio, efectivamente fue así. Las gráficas que representaban la evolución del sistema en el tiempo eran muy similares. Sin embargo, a partir de cierto punto, las gráficas se empezaban a diferenciar, y la diferencia se hacía cada vez mayor conforme se contemplase un mayor tiempo de evolución en la dinámica del sistema. Lo que Lorenz observó fue el comportamiento de un sistema caótico. Un sistema extremadamente sensible a las condiciones iniciales. Piénsese por ejemplo en un péndulo simple. A partir de unas ecuaciones relativamente sencillas podemos calcular su posición al cabo de un tiempo  $T$  -llamemos a esa posición  $Y$ -, indistintamente del punto en el que es soltado o lanzado o tirado; llamemos a ese punto  $P$ . Si el péndulo es nuevamente lanzado desde un punto ligeramente distinto de  $P$ , resulta que, con dichas ecuaciones, al cabo del mismo tiempo  $T$ , el péndulo se encontrará en una posición muy cercana a  $Y$ . Lo que quiere decir que el péndulo simple no es un sistema caótico. Caso totalmente contrario a un péndulo doble. A partir de un sistema de ecuaciones, ya no tan sencillas como en el caso del péndulo simple, se puede igualmente determinar en cualquier tiempo  $T$  la posición de ambos péndulos. Sin embargo, si se vuelve a lanzar o tirar desde una posición muy ligeramente diferente de la inicial, la evolución del sistema en ese mismo tiempo  $T$  será una completamente distinta. Su comportamiento será uno completamente distinto. Esto se verá

reflejado en la gráfica con la que se representa dicho comportamiento. Si se comparan ambas gráficas serán unas muy distintas.

El otro episodio que se podría resaltar como uno duro golpe a la visión mecanicista de la naturaleza fue el desarrollo de la mecánica cuántica. Más concretamente, el *principio de indeterminación* que se da dentro de ella; principio también conocido como “principio de incertidumbre”, que nos dice que no es posible medir o calcular con total exactitud la posición y el momento de una partícula en simultáneo. A mayor precisión que se logre en el cálculo del momento, menor precisión se tendrá en el cálculo de la posición. Y viceversa. A mayor precisión en el cálculo de la posición, menor precisión se tendrá en el cálculo del momento. Lo que quiere decir que el requerimiento principal en las consideraciones mecanicistas, recogidas en el planteamiento laplaciano, es inalcanzable; al menos, desde el marco general de la mecánica cuántica. Marco que, obviamente, es susceptible de ser desplazado, como cualquier otro marco, pero que por el momento se presenta como uno de gran aceptación en la academia, y que además ha dado muy buenos resultados.

Pero, más allá de lo inalcanzable o no que resulte el proyecto mecanicista, “a lo Laplace”, a la luz de las observaciones de sistemas dinámicos no lineales y caóticos, y a la luz de la ahora famosa indeterminación observada por Werner Heisenberg en 1927, quizás lo más importante para resaltar sea que estos desarrollos, junto con otros mencionados, irrumpen, como lo habían planteado Anderson y Laughlin -por mencionar algunos de los autores tratados-, como una nueva forma de ver el mundo; esto, posibilitado por el desarrollo de nuevas herramientas matemáticas.

Con el estudio, por ejemplo, llevado a cabo por Poincaré sobre el famoso *three-body problem*, se empezó a vislumbrar un terreno donde lo complicado y lo caótico confluían. Poincaré podría considerarse el pionero más representativo de la teoría moderna de los

sistemas dinámicos. Aunque ya el famoso físico escocés James Clerk Maxwell había previsto en 1873 que existía una clase de fenómenos “afectados por influencias cuyas magnitudes físicas son tan pequeñas para ser tomadas en cuenta por un ente finito, pero que podrían producir resultados de la mayor importancia” (Mitchell, 2009, p. 20); justamente lo que sería el caos. Teniendo la posición inicial y las velocidades de tres partículas, el problema de los tres cuerpos consiste, de manera general, en determinar, aplicando las leyes de Newton, el movimiento de dichas partículas bajo la fuerza de atracción gravitacional que se ejercen cada una entre sí. Poincaré, que en el intento de resolución de dicho problema hubo de desarrollar toda un área de las matemáticas que se conoce como topología algebraica (ibid., 2009, p. 21), notó que había una gran sensibilidad a ligeras variaciones en los valores iniciales en las posibles soluciones del problema; problema que en últimas no pudo resolver, pero en dicho intento, no sólo desarrolló un área importante de las matemáticas, sino que además nos puso de manifiesto una clase particular de fenómenos, que cambia un poco el panorama de nuestra visión de la naturaleza; nos puso de manifiesto el fenómeno del caos.

“Si conociéramos de manera exacta las leyes de la naturaleza y las condiciones del universo en su momento de inicio, podríamos predecir con exactitud la situación del universo en un momento posterior. Pero incluso si fuera el caso de que las leyes de la naturaleza no fuesen más un secreto para nosotros, aun así, sólo podríamos conocer las condiciones iniciales de manera aproximada. Si esto nos permite predecir la situación futura con la misma aproximación, entonces es todo lo que necesitamos, y podríamos decir que el fenómeno ha sido predicho; lo que quiere decir que el fenómeno es gobernado por dichas leyes. Pero no siempre es este el caso; ocurre que ligeras variaciones en las condiciones iniciales dan lugar a grandes diferencias en el fenómeno resultante. Por lo que un muy pequeño error en el cálculo de las condiciones iniciales

produce un error enorme en la predicción. La predicción se hace imposible y tenemos lo que se conoce como un fenómeno fortuito.”<sup>37</sup> (Poincaré, 1914, p. 68)

Es importante resaltar que caos no implica necesariamente complicación. Ni al nivel del fenómeno a explicar, ni al nivel de las matemáticas involucradas en los esquemas explicativos con los que se intenta dar cuenta. De hecho, quizás la que sea la más tradicional de las ecuaciones en el estudio de sistemas no lineales y caóticos es una ecuación iterativa bastante sencilla. Nos referimos a la famosa ecuación logística, o mapa logístico.

$$x_{t+1} = Rx_t(1 - x_t)$$

Si bien es una ecuación sencilla, nos pone de manifiesto, de manera muy clara, las características distintivas del fenómeno del caos. Piénsese por ejemplo en la situación cuando  $R$  es igual a 2 y  $x$  en  $t = 0$  es igual a 0.5. Se tiene que:

$$x_1 = 2 * 0.5(1 - 0.5) = 0.5$$

$$x_2 = 2 * 0.5(1 - 0.5) = 0.5$$

$$x_3 = 2 * 0.5(1 - 0.5) = 0.5$$

⋮

---

<sup>37</sup> *If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation approximately. If that enabled us to predict the succeeding situation with the same approximation, that is all we require, and we should say that the phenomenon has been predicted, that it is governed by laws. But it is not always so; it may happen that small differences in the initial conditions produce very great ones in the final phenomenon. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have a fortuitous phenomenon* (Poincaré, 1914, p. 68)

$$x_n = 2 * 0.5(1 - 0.5) = 0.5$$

⋮

Y así, sucesivamente. Piénsese ahora en los valores arrojados por la ecuación cuando  $R = 2$ , pero ahora  $x_0 = 0.2$ . Los valores arrojados serían los siguientes:

$$x_1 = 0.32$$

$$x_2 = 0.4352$$

$$x_3 = 0.49160192$$

$$x_4 = 0.4998589445$$

$$x_5 = 0.4999999602$$

$$x_6 = 0.5$$

⋮

$$x_n = 0.5$$

⋮

Como se puede observar, acá también los valores arrojados por la ecuación son, a partir de cierto punto, siempre 0.5; sólo que le toma cinco pasos o iteraciones llegar a ese punto. (Mitchell, 2009, p. 29). Veamos qué sucede si  $x_0 = 0.9$ .

$$x_1 = 0.18$$

$$x_2 = 0.2952$$

$$x_3 = 0.41611392$$

$$x_4 = 0.48592625116$$

$$x_5 = 0.49960385918$$

$$x_6 = 0.49999968614$$

$$x_7 = 0.5$$

⋮

$$x_n = 0.5$$

Nuevamente, luego de cierto número de iteraciones, las soluciones que termina arrojando la ecuación son un mismo valor: 0.5. Esto quiere decir que para  $R = 2$ , la ecuación logística presenta lo que se conoce como un *punto fijo* (ibid., 2009, p. 29). A partir de cierto número de iteraciones, salvo en el caso en el que  $x_0 = 0.5$ , que lo alcanza inmediatamente, la ecuación se fija en un punto y no arroja más soluciones distintas; queda, por así decirlo, atrapada en ese punto. Para el caso, por ejemplo, de  $R = 2.5$ , la ecuación queda “atrapada” en otro punto. En este caso, 0.6.

Veamos qué ocurre ahora para el caso  $R = 3.1$  y  $x_0 = 0.2$ .

$$x_1 = 0.496$$

$$x_2 = 0.7749504$$

$$x_3 = 0.54064706037$$

$$x_4 = 0.76987823109$$

$$x_5 = 0.54921379517$$

$$x_6 = 0.76749180732$$

$$x_7 = 0.55318921233$$

$$x_8 = 0.76622981384$$

$$x_9 = 0.55527722728$$

Hasta donde se puede observar, la ecuación arroja siempre valores distintos. No parece estabilizarse en punto alguno; sin embargo, también se puede observar que a partir de la tercera iteración los valores que arroja la ecuación dan la impresión de que oscilan entre dos puntos. Si se sigue iterando, los dos puntos entre los que oscilan las soluciones que va arrojando la ecuación se presentan de manera más clara. Estos dos puntos son: 0.5580141 y 0.7645665. (ibid., 2009, p. 29). Debido a que la ecuación bajo esas condiciones iniciales arroja, a partir de cierta iteración, soluciones que oscilan entre dos valores, se dice que bajo esas condiciones la ecuación tiene un *período* igual a 2. Para el caso, por ejemplo, en que  $R$  tome valores entre 3.4 y 3.5, para cualquier valor inicial  $x_0$ , la ecuación alcanzará un período igual a 4. Es decir, llegará un momento en que sus soluciones oscilen entre 4 valores, que son, aproximadamente, 0.872, 0.389, 0.829, y 0.494 (ibid., p. 31).

Si consideramos valores para  $R$ , entre 3.54 y 3.55, la ecuación aumentará su período el doble; es decir, saltará de período 4 a período 8. Para valores de  $R$ , entre 3.564 y 3.565, la ecuación nuevamente aumenta su período el doble; pasa de período 8 a 16. Entre 3.5687 y 3.5688, pasa de período 16 a 32. Así, parece haber un patrón de crecimiento del período de la ecuación. El período parece incrementarse el doble continuamente (ibid., p. 31). Sin embargo, a partir de cierto valor de  $R$ , la ecuación deja de tener algún patrón identificable entre las sucesivas soluciones que va arrojando. La ecuación ya no se estabiliza en ningún punto. Pero no sólo eso, sino que se puede observar con claridad la extrema sensibilidad a las condiciones iniciales característica de los sistemas caóticos. Cuando  $R$  toma aproximadamente el valor de 3.569946, la ecuación no se estabiliza en ningún punto.

Aunque la extrema sensibilidad en las condiciones iniciales quizás se observa en su máxima expresión cuando  $R$  es igual a 4. Cuando  $R$  es igual a 4, y pongamos por caso que  $x_0 = 0.2$ , si se grafica las soluciones que va arrojando la ecuación y se compara la gráfica con la que se genera partiendo no ya desde  $x_0 = 0.2$  sino desde  $x_0 = 0.2000000001$ , las gráficas difieren de manera considerable, muy a pesar de que entre ambos valores iniciales la diferencia sea extremadamente pequeña.

A continuación, la gráfica en cuestión. Donde la curva sólida con círculos negros representa las soluciones que va arrojando la ecuación al iniciar con  $x_0 = 0.2$ ; mientras que la curva punteada y con círculos blancos representa las soluciones al iniciar las iteraciones con  $x_0 = 0.2000000001$ .

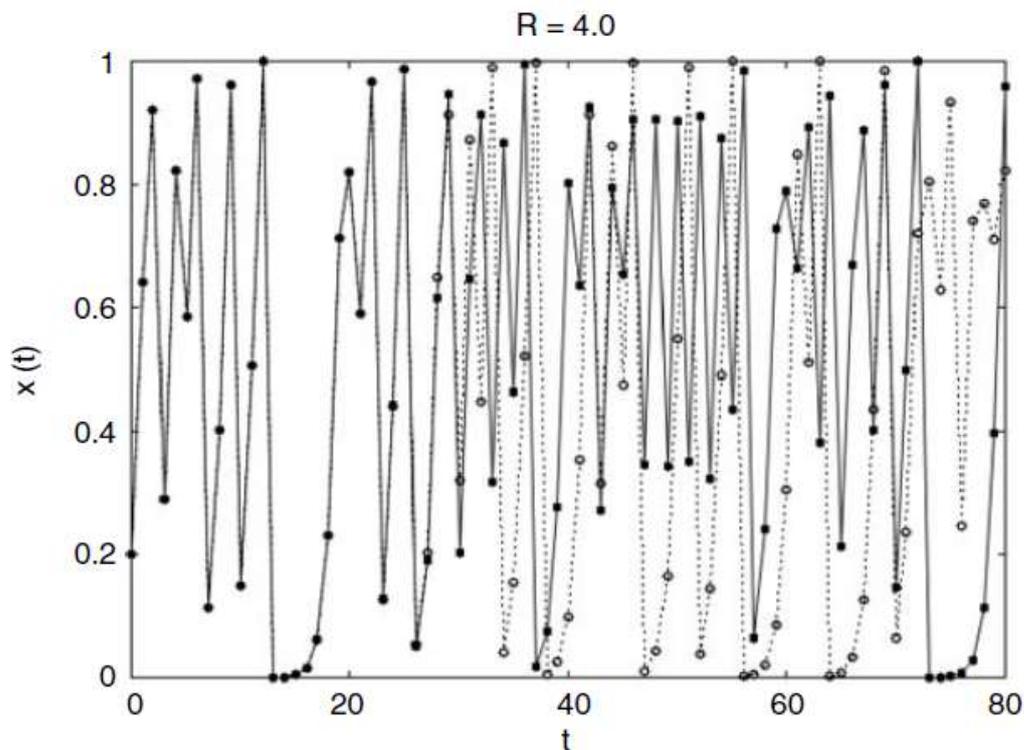


Figura tomada de (Mitchell, 2009, p. 32)

A partir de esta simple ecuación entonces, se pueden observar, con algo de claridad, las características distintivas del fenómeno del caos. Características que podemos resumir de la siguiente manera:

En un sistema con una dinámica determinista puede emerger un comportamiento aparentemente aleatorio sin que intervenga alguna fuente de aleatoriedad externa. Es decir, una dinámica determinista puede conllevar a un comportamiento *macroscópico* aleatorio, sin que intervenga un factor de aleatoriedad externo (ibid., p. 38). Se dice que el sistema es determinista en el sentido de que cualquiera de sus estados viene estrictamente condicionado por su estado anterior. Cualquier estado  $x_i$  viene determinado por las condiciones del sistema en el estado  $x_{i-1}$ . Si se pudiese conocer con total precisión esas condiciones, se podría

predecir, con total precisión, la evolución del sistema. El gran problema radica, justamente, en el cálculo de esas condiciones con total precisión; pues se sabe que una diferencia ínfima en los valores de esas condiciones puede dar lugar a trayectorias evolutivas completamente distintas. Y, sobre todo, que existe una indeterminación cuántica entre ciertas magnitudes, como la posición y el momento, que puede llegar a tener repercusiones a la escala macroscópica de los sistemas complejos que acá nos competen.

Una segunda característica distintiva del fenómeno del caos, estrechamente relacionada a la emergencia de un comportamiento aleatorio a partir de una dinámica determinista, es que incluso un sistema con un comportamiento simple y de carácter determinista puede resultar completamente *impredecible*, debido a su extrema sensibilidad a las condiciones iniciales. (ibid., p. 38).

Una tercera característica es que:

“Aunque el comportamiento detallado de un sistema caótico no pueda ser predicho, existe cierto “orden en el caos”, observado en propiedades comunes presentes en un extenso conjunto de estos sistemas. Como, por ejemplo, el camino hacia el caos a través de una duplicación de período, y, la [famosa] constante de Feigenbaum. Así, aunque establecer “predicciones resulte imposible” a un nivel detallado, a una escala mayor existen ciertos aspectos que sí pueden ser en efecto predichos.”<sup>38</sup> (Ibid., p. 38)

La constante de Feigenbaum podemos definirla como la razón entre el ancho de cada segmento de bifurcación y el ancho del segmento siguiente. Mitchell Feigenbaum notó que la ruta hacia el caos por medio de una duplicación del período se iba haciendo cada vez más

---

<sup>38</sup> *Although the detailed behavior of a chaotic system cannot be predicted, there is some “order in chaos” seen in universal properties common to large sets of chaotic systems, such as the period-doubling route to chaos and Feigenbaum’s constant. Thus even though “prediction becomes impossible” at the detailed level, there are some higher-level aspects of chaotic systems that are indeed predictable.* (Mitchell, 2009, p. 38)

rápida. Esto se pudo observar por ejemplo cuando analizábamos el aumento doble de período para valores de  $R$  cada vez más cercanos entre sí. Entre 3.564 y 3.565, la ecuación pasa de período 8 a 16. Entre 3.5687 y 3.5688, pasa de período 16 a 32. A medida que los períodos se duplican, los valores de  $R$  se van haciendo cada vez más y más cercanos.

Usando estos valores Feigenbaum midió la tasa a la que las bifurcaciones se van haciendo cada vez más y más cercanas; es decir, la tasa a la que *converge* los valores de  $R$ . Y descubrió que la tasa es (aproximadamente) 4.6692016. (ibid., p. 36).

Con lo dicho hasta ahora queda expresada la idea de que, a partir de dinámicas simples y deterministas, pueden surgir aspectos aleatorios, incluso sin que una fuente de aleatoriedad intervenga. Si bien, por ejemplo, en el caso de la ecuación o mapa logísticos, las matemáticas involucradas son sencillas, y los comportamientos que con ella se podrían hasta cierto punto modelar son comportamientos relativamente sencillos, es cierto también que hay sistemas caóticos donde confluyen tanto complicación matemática (para poderlos modelar) como complejidad comportamental. Son sistemas dinámicos alejados del equilibrio.

Sobre este tipo de sistemas, los estudios llevados a cabo por el reconocido premio Nobel de química Ilya Prigogine, nos ofrecen algunas luces. Sobre todo, en su estrecha relación con el fenómeno de la complejidad en los sistemas biológicos de nuestro interés (por ejemplo, *Acetabularia acetabulum* y biomoléculas -más concretamente: proteínas). Aunque también dejan de manifiesto algunas limitaciones para abordar desde ese marco general la complejidad en sus distintos aspectos.

En el marco de sus estudios acerca de sistemas en estados muy alejados del equilibrio termodinámico, como suelen ser los sistemas complejos, Prigogine concentró su atención en ciertos fenómenos que pueden ocurrir en tal régimen, los cuales conducen *espontáneamente*

a un cierto tipo de organización macroscópica al que denominó “estructura disipativa”. Las estructuras disipativas introducen en el sistema un tipo de *auto-organización* que rompe con la simetría del espacio y del tiempo (Nicolis y Prigogine 1989). En algunos casos, el sistema adquiere una organización espacial supramolecular, que puede describirse mediante magnitudes macroscópicas del orden de los centímetros, y que permite definir una noción intrínseca de distancia en un sistema inicialmente homogéneo y carente de toda escala espacial. Prigogine se refiere a este fenómeno como “ruptura de la simetría espacial”. En otros casos, ciertas magnitudes del sistema, como las concentraciones de algunas sustancias, comienzan a oscilar con un período bien definido que depende de los parámetros del sistema. En esta situación se produce una “ruptura de la simetría temporal”, y el sistema se convierte en un reloj químico.

Puesto que en biología la noción de evolución se suele asociar al aumento de la complejidad (Goodwin, 1998; Alberch, 1998; Margalef, 1998; McKinney; Hull, 1998) y a la formación de estructuras y funciones cada vez más complejas, la noción de estructura disipativa parece suministrar una herramienta fundamental para la comprensión de tal aumento de complejidad. En particular, las estructuras disipativas rompen con la equiprobabilidad entre las secuencias de elementos moleculares que forman las biomoléculas, como en el caso concreto de las proteínas, que constituyen un ejemplo central en nuestra argumentación. Lejos del equilibrio, las concentraciones de los componentes pueden adquirir regularidades espaciales o temporales que incrementen la probabilidad de ocupación para alguno de los componentes, de modo tal que ciertas secuencias resulten probabilísticamente privilegiadas. Es así como, lejos del equilibrio, pueden surgir las propiedades que caracterizan, por poner un ejemplo de biomoléculas, a las proteínas: no

repetitividad, existencia de correlaciones de largo alcance y ruptura de la simetría espacial (Prigogine y Stengers 1988).

En la década de los '90, Brian Goodwin aplicó la idea de ruptura de simetría en sistemas altamente inestables al ámbito de la biología evolutiva en su conocida obra *How the Leopard Changed Its Spots* (1994). Según el autor, todos los organismos comienzan como entidades altamente simétricas, como una única célula esféricamente simétrica. A medida que el organismo crece, este estado altamente simétrico se convierte en inestable debido a sus propias tensiones y tendencias internas, o debido a influencias externas provenientes del entorno. Es aquí donde ingresa la ruptura espontánea de simetría: el organismo pasa a un estado que pertenece a un conjunto de estados posibles estables, pero menos simétricos, como ocurre con la formación de verticilos en *Acetabularia* y los distintos tipos de disposición foliar en las plantas; casos de los que hablaremos más adelante. De este modo, la dinámica de estabilidad y ruptura espontánea de simetría restringe las formas generales posibles que un organismo puede adoptar durante su crecimiento. Cuál de los posibles estados el organismo adopta en cada etapa de su desarrollo puede estar controlado internamente -por su DNA, por ejemplo-, o por el entorno -a través de la temperatura o de componentes químicos, como veremos en el ejemplo de *Acetabularia acetabulum*.

Según Goodwin, estos nuevos conceptos de la ciencia de la complejidad tienen implicaciones radicales para la teoría de la evolución. De acuerdo con el enfoque darwinista tradicional, la evolución tiene la libertad de “explorar” una inmensa variedad de posibilidades, restringidas únicamente por las leyes de la física y de la química. Este enfoque, sin embargo, deja algunos problemas sin resolver, de los cuales dos de los más importantes son la emergencia de las mismas formas generales en diferentes linajes, y el hecho de que no vemos a la evolución explorando todas las posibilidades sino, por el contrario, sólo un muy

limitado subconjunto de ellas, como ocurre en el caso de las proteínas. De acuerdo con las ideas que surgen de las ciencias de la complejidad, el dominio de lo biológicamente posible se encuentra fuertemente constreñido por la inestabilidad dinámica, donde la ruptura espontánea de simetría cumple un papel central. Sobre esta base, Goodwin define el *aumento* de la complejidad en la evolución biológica como el número de rupturas de simetría que experimenta una especie al pasar de estados de mayor simetría a estados de menor simetría.

## II.2. Evolución

### II.2.1. Los orígenes de la idea de evolución

Hacia finales del siglo XVIII, países europeos como Francia, Gran Bretaña y Alemania vivían una época de expansión y difusión del conocimiento, con un marcado optimismo respecto del poder explicativo de la visión científica del mundo. A esta época perteneció Jean Baptiste de Lamarck, naturalista francés a quien se debe la primera teoría de la evolución plenamente articulada. Pues, si bien el problema de la aparición de los organismos por medios naturales y de la transformación de estos fue un problema planteado y discutido desde la época de los antiguos griegos -como lo ponen de manifiesto las propuestas de Empédocles y de Anaximandro-, no es sino hasta Lamarck que se tiene una propuesta explicativa del mecanismo subyacente a dichas transformaciones.

Desde la época de los griegos se creía en la existencia de un proceso en la naturaleza a través del cual las especies sufren cierto tipo de transformaciones; sobre todo, transformaciones que las hacen *superiores* en algún sentido respecto de las anteriores, por ejemplo, más complejas. Como afirma Jorge Wagensberg (1998), no se puede negar que entre una bacteria y William Shakespeare algo ha sucedido. No es éste el problema. Lo que sí ha

resultado problemático, ya desde la época de los griegos, ha sido la explicación del *modus operandi* de tal proceso evolutivo.

Por un lado, Lamarck sostiene la idea de que en los organismos existe cierta tendencia, o impulso natural, hacia el aumento de organización y de complejidad. Los organismos buscan instintivamente ascender en la Gran Cadena del Ser; esto es, pasar a niveles superiores en la organización jerárquica de los organismos, según la complejidad de sus anatomías, gracias a transformaciones que hacen de su morfología una más compleja. Es importante recordar que, para Lamarck, la naturaleza sólo ha producido directamente, por generación espontánea, organismos sencillos como protozoarios y bacterias, y a partir de estos ha producido sucesivamente los demás. (Ruiz y Ayala, 2002). Por otro lado, Lamarck considera que los organismos buscan satisfacer sus necesidades biológicas en ciertas circunstancias ambientales particulares; por lo tanto, si éstas cambian, los organismos tendrán que cambiar, con el fin de seguir satisfaciendo sus necesidades y su impulso natural de ascender en la Gran Cadena del Ser.

Esto significa que, si las circunstancias ambientales cambian, los organismos muy posiblemente tendrán que emplear estrategias alternativas para satisfacer sus necesidades; y con esto es muy probable que presenten un cambio en sus acciones. Esas nuevas acciones requerirán cierto esfuerzo de algún o algunos órganos, minimizando quizás el uso de otros. En tanto que esas nuevas acciones persistan durante un período considerable de tiempo, lo que ocurrirá es que algunos órganos se desarrollarán y muy posiblemente otros se atrofiarán. Estas adquisiciones y pérdidas se conservarán por el mecanismo de la herencia, siempre y cuando el nuevo carácter morfológico de la especie sea común a ambos sexos (Ruiz y Ayala, 2002).

Sin duda, la perspectiva científica de Lamarck contiene una fuerte dosis de esencialismo y teleología, al suponer que los organismos poseen una tendencia natural a ascender en la Gran Cadena del Ser. No obstante, Lamarck se ocupa de la pregunta ¿qué es lo que se esconde en el fondo de las transformaciones que sufren los organismos en ese movimiento ascendente? Su respuesta consiste en afirmar que aquello que subyace a todo el proceso de transformación de los organismos es el fenómeno de la *adaptación*: la acomodación de los organismos a las presiones ejercidas por las circunstancias ambientales.

### II.2.2. Selección natural, adaptación y gradualismo

Con la publicación en 1859 de *El Origen de las Especies*, se propone un mecanismo alternativo al lamarckiano para dar respuesta al problema de la adaptación. Otra respuesta al mismo problema es la que brindaba la Teoría del Diseño Inteligente, formulada por William Paley en su famosa obra *Teología Natural*. Tanto los trabajos de Paley como los de Lamarck habían sido estudiados por Charles Darwin, constituyéndose ambos en referencia directa para el naturalista inglés. De Lamarck, Darwin conservó, como hipótesis secundarias, por un lado, aquella que explica los cambios morfológicos de las especies por la acción del uso y el desuso de ciertos órganos, y por otro aquella que apela a la acción directa que el medio ejerce en la confección de estos cambios (Cadevall, 2005. p. 179). Asimismo, de Paley conservó la caracterización general de los organismos que se adaptan al medio, a saber: “*complejos, con partes coadaptadas y con propósito*” (Ginobili, 2014, p. 381).

Sin embargo, de acuerdo con algunos autores, uno de los aspectos revolucionarios del darwinismo es justamente la manera diferente en la que concibió la noción de adaptación. Gustavo Caponi, por ejemplo, afirma que la manera en la que Darwin concibió la adaptación

no se hallaba, ni en teólogos naturales, ni en naturalistas anteriores a él; por ello, considera al adaptacionismo predarwiniano como un mito (Caponi, 2011).

La respuesta que ofrece Darwin al problema de la adaptación se basa sobre la adopción de al menos dos hipótesis. La primera de ellas es la idea de corte malthusiano según la cual no todos los individuos de una población tendrán la posibilidad de reproducirse. Recordemos que Malthus afirmaba que los crecimientos de poblaciones y de recursos siguen progresiones distintas: mientras las poblaciones crecen en progresión geométrica, los recursos crecen en progresión aritmética. La segunda hipótesis darwiniana es la que supone la existencia de ciertos rasgos distinguibles que varían entre los miembros de una población y que son heredables.

La respuesta que ofrece Darwin al problema de la adaptación, mediante la postulación del mecanismo evolutivo de la selección natural, consiste entonces en afirmar que la adaptación es el resultado de la selección de rasgos distinguibles y heredables vinculados a una función específica, rasgos que favorecen a algunos individuos en la competencia por la supervivencia. En esta competencia, quienes sobrevivan serán justamente los que logren dejar mayor descendencia, retroalimentando así positivamente el bucle, ya que dicho mecanismo es iterativo. Ahora bien, al hablar de selección, es claro que se está presuponiendo un abanico de alternativas. Como ya fue señalado, deben existir variaciones de dichos rasgos entre los miembros de una población. Por lo tanto, de manera concisa, puede decirse que la selección natural opera sobre las variaciones distinguibles de rasgos heredables entre los miembros de una población (Godfrey-Smith, 2009).

Ahora bien, si para que el mecanismo de la selección natural opere es necesaria una variabilidad entre los rasgos distinguibles y heredables ligados a ciertas funciones, la

pregunta inmediata que se impone es de dónde surge esa variabilidad. La respuesta que ofreció Darwin es que existen ciertos factores causales que determinan dichos rasgos, de suerte que alteraciones graduales en esos factores serían los responsables de esta variabilidad. De este modo, Darwin presentaba una visión alternativa al saltacionismo, según el cual la aparición de nuevas estructuras era explicada como consecuencia de grandes cambios ocurridos por eventos súbitos (Cadevall, 2005, p. 179). En otras palabras, Darwin postuló un mecanismo basado en una dinámica gradual: la selección natural opera sobre pequeñas variaciones, de manera que la aparición de nuevas estructuras se explica como resultado de cambios graduales.

Para Darwin, los pequeños cambios en los rasgos distinguibles y heredables, ligados a ciertas funciones que dan lugar al proceso de evolución biológica, eran debidos a pequeñas alteraciones en los factores causales que determinan dichos rasgos. Sin embargo, Darwin no contaba con teoría alguna para explicar tales alteraciones. Pocos años después de la publicación de *El Origen de las Especies*, en 1866, tales factores causales fueron identificados por Gregor Mendel, quien los denominó ‘factores hereditarios’; más tarde recibieron el nombre de *genes*.

### II.2.3. *Teoría sintética de la evolución*

Hacia comienzos del siglo XX, gran parte de las ideas de Darwin habían caído en descrédito, y las teorías que se presentaban como hegemónicas en la época en que Darwin presentaba su propuesta recuperaban el terreno cedido. Fueron varios los años en los que el elemento central del darwinismo, esto es, la selección natural, recibió un rechazo casi

generalizado. Sólo permanecían en pie algunas de sus ideas, como la de la importancia de los factores geográficos al momento de explicar la diversidad biológica (Folguera, 2009, p. 42).

Con los trabajos de Mendel y Morgan en genética y el planteamiento de modelos o esquemas matemáticos para sistemas biológicos por parte de Fisher, Haldane y Wright, la teoría darwiniana recibió un nuevo aire. Los nuevos modelos permitieron la formulación de ciertas relaciones interteóricas que lograron ampliar el panorama diseñado por Darwin, particularmente en lo referido a la conceptualización de los rasgos distinguibles y heredables, sobre los que opera la selección natural. A su vez, estas relaciones permitieron establecer puentes entre otras disciplinas biológicas, como la paleontología, la genética y la sistemática (Folguera 2009, p. 43). De este modo, métodos, estrategias investigativas y posturas filosóficas lograron hasta cierto punto confluir en esta síntesis.

Se dice entonces que uno de los aspectos representativos de la teoría sintética de la evolución, al menos en lo que para autores como Gould (2002) fue su primera etapa, fue su pluralismo. Luego, a partir de la década de los '50, se inicia una etapa durante la cual la selección natural se convierte nuevamente en hegemónico como mecanismo evolutivo. No obstante, aun en este segundo período reconocido por Gould, en el contexto de la síntesis se reconoce la importancia de *eventos aleatorios* para la comprensión del proceso evolutivo. Si bien el mecanismo de la selección natural se concibió como el más importante, en la teoría sintética además se admiten como co-actantes del proceso otros mecanismos: la migración, el aislamiento y la deriva.

Además de la preponderancia del mecanismo de la selección natural en la explicación del problema de las adaptaciones, la teoría sintética adopta del darwinismo la concepción de que los mecanismos evolutivos operan de forma gradual. Las adaptaciones siguen siendo el

resultado de la selección de variaciones graduales en la expresión de los factores causales – genes– en ciertos rasgos. El mecanismo de la selección natural determina una tendencia hacia mayores adaptaciones.

#### II.2.4. *¿Cambio o direccionalidad? Progreso panglosiano y el reto de Hull*

Si bien es cierto que desde los primeros organismos unicelulares hasta William Shakespeare ha sucedido algo, la pregunta que surge inmediatamente es qué es lo que ha sucedido. Sin duda, ha existido un cambio, pero la cuestión reside en decidir qué tipo de cambio. Según Darwin, el proceso evolutivo se dirige hacia una mayor adaptación. Posteriormente, algunas respuestas prefieren señalar que en el proceso evolutivo se han visto favorecidas ciertas tendencias, tales como el aumento en los niveles de organización, un aumento en la complejidad, entre otras. Sin duda, la idea de un aumento de la complejidad durante el proceso evolutivo es fuertemente intuitiva. Sin embargo, cuando la idea se pretende precisar para su tratamiento filosófico, surgen de inmediato algunas cuestiones nada fáciles de resolver; en particular, no resulta sencillo hallar una magnitud significativa en el ámbito biológico con la cual organizar jerárquicamente a los organismos en términos morfológicos o en términos funcionales, y que represente una manifestación de la complejidad.

En 1759, bajo el seudónimo “el Señor Doctor Ralph” (“Monsieur le docteur Ralph”), Voltaire publica su novela ‘*Cándido*’, donde intenta, de una manera fina y disimulada, ridiculizar en cierto sentido la idea leibniziana, sostenida en la obra por el doctor Pangloss, de que vivimos en el mejor de los mundos posibles. Para el doctor Pangloss, tutor de Cándido, protagonista de la obra, del muestrario de acontecimientos indeseados o infortunados para la

especie humana en el siglo XVIII, se puede concluir, no obstante, que todos ellos ocurrieron para bien de la especie ya que “todo sucede para bien” (“*tout est au mieux*”).

En su famoso artículo de 1979, ‘*The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme*’, Stephen Jay Gould y Richard Lewontin denunciaron el “panglossianismo” extendido en el ámbito de la biología evolutiva, e incluso en ciertos sectores de la paleontología. Según estos autores, el paradigma “panglossiano” se extendió en la comunidad de los biólogos evolutivos e hizo que muchos autores encontraran pautas filogenéticas progresivas; o, en palabras más directas, cierta direccionalidad progresiva en el proceso evolutivo.

En su artículo ‘*Progreso Panglossiano*’ donde analiza la idea de progreso evolutivo, David Hull (1998) se hace eco de las denuncias de Gould y Lewontin y, en consecuencia, considera indispensable tratar en primera instancia el tema de la direccionalidad en la evolución. En particular, Hull se pregunta acerca del tradicional supuesto de que existe una transición de las formas inferiores hacia formas superiores que brinda al proceso evolutivo una direccionalidad progresiva.

A juicio de Hull, si la evolución no muestra siquiera rasgos de direccionalidad, no habría mayores razones para aceptar el progreso en la filogenia. Por este motivo es que el autor se pregunta, en primer lugar, si la filogenia sigue alguna dirección. Hull responde afirmativamente a esta pregunta, señalando que, de hecho, no existe sólo una, sino muchas direcciones. El asunto está en que estas direcciones son discernibles precisamente por las propias reconstrucciones que hemos hecho de la filogenia.

Hull también se opone al uso de conceptos informacionales en biología. Es una idea casi totalmente aceptada de que los genes, además de replicarse, también transcriben y traducen el tipo de proteínas que una célula puede producir. Hull señala que los conceptos de

transcripción o traducción se encuentran claramente elucidados en el ámbito de la matemática, ya que la teoría de la información es una teoría matemática precisa y bien articulada. Por el contrario, cuando se habla, por ejemplo, del contenido de información del código genético, o de la traducción de aminoácidos en proteínas, no existe un análisis formal que brinde contenido preciso a tales conceptos.

Hull se encarga de mostrar que muchas de las reconstrucciones de la filogénesis están sesgadas, por lo que cualquier pronunciamiento acerca de tendencias o direcciones en la evolución debe ser tomado con cierto escepticismo. El sesgo más fuerte que Hull identifica es el sesgo de la complejidad y la información. En el ámbito académico se encuentra ampliamente difundida la idea según la cual la vida se ha convertido en un fenómeno cada vez más complejo a lo largo de la evolución biológica. La afirmación “un elefante es más complejo que un extremófilo” parece expresar una perogrullada. Pero Hull señala que la noción de complejidad involucrada es completamente intuitiva: no existe una dilucidación precisa del concepto de complejidad. Lo mismo sucede con el uso del concepto de información: la noción de *contenido de información genética* se utiliza ampliamente en biología, a pesar de que no existe una definición formal del concepto de información en este ámbito. Para Hull, mientras no se realice un análisis formal y detallado del concepto de complejidad, será dudosa cualquier afirmación acerca de alguna dirección evolutiva progresiva en términos de un aumento de la complejidad, puesto que en teoría es siempre posible favorecer cierto aspecto de la filogenia en términos de definiciones alternativas del concepto de complejidad. En esto consiste *el reto de Hull*: en exigir una definición formal de la noción intuitiva de complejidad para evaluar la posible direccionalidad del proceso evolutivo.

La propuesta de la complejidad y emergencia modelo-paramétricas, de la que se hablará un poco más adelante, si bien expresamente no pretende ser propiamente un intento de respuesta a este reto, al que hemos denominado '*el reto de Hull*', lo cierto es que en principio sí parecería lograr alcanzar cierta relevancia frente a este; constituyendo así una muy interesante propuesta de investigación. A pesar de que los fines dispuestos en esta tesis no incluyen particularmente un intento de respuesta a este reto, existe un interés futuro en encararlo. Esto hace en cierta medida que haya una motivación de interpretar esta propuesta a partir de ejemplos en el ámbito de la biología. Incluso algunos pretendidamente en el ámbito propiamente de la biología evolutiva.

### II.3. *Estadística y probabilidad*

#### II.3.1 *Aleatoriedad y complejidad algorítmica*

La teoría matemática de la complejidad algorítmica es una que tiene sus orígenes en la década del 60, cuando tres autores, de manera independiente, se dieron a la tarea de elucidar la noción de longitud de un programa de cómputo; pues veían en ella una idea importante. Estos tres autores fueron, el reconocido matemático ruso Andréi N. Kolmogorov, el estadounidense Ray Solomonoff, quien para la época trabajaba en Inteligencia Artificial, y el joven matemático Gregory Chaitin (15 años tenía para la fecha) (Chaitin, 2002, p. 57).

Solomonoff, amigo cercano de Marvin Minsky, uno de los pensadores más representativos en el área de la Inteligencia Artificial, pensó precisamente en la importancia de la descripción de la longitud de un programa en el contexto de esta área. Estaba pensando en una reinterpretación de la famosa navaja de Occam (ibid., 2002, p. 57); la mejor teoría vendría a hacer ahí la que tenga una longitud más corta en términos computacionales. El intento de

elucidar la noción de longitud de un programa cobra en él importancia porque le permitiría reinterpretar el famoso principio filosófico en términos del programa más corto.

Kolmogorov y Chaitin, por otra parte, gracias a que contaban con formación en matemáticas, concibieron el asunto como uno de carácter matemático. Particularmente, como uno de tipo probabilístico. La consideración de la importancia de la longitud de un programa les permitió formalizar la noción de ‘*aleatoriedad*’ en dicha área; noción que será central en esta sección, y, como se ha podido notar, de gran importancia para la tesis en general. Una secuencia de unos y ceros es aleatoria cuando el programa más corto que la puede generar posee la misma longitud que la secuencia original (Fodor 1983; Chaitin 2002). Es decir, posee tantas instrucciones como datos contenga la secuencia. Para cada dato, se necesitaría una instrucción que lo generase.

Chaitin conceptualizó en principio la noción de longitud de un programa fijando su atención en las máquinas de Turing. La longitud de un programa se establecía mediante el número de estados que tuviese la máquina. Luego optó por usar un lenguaje de programación particular (LISP) y, con relación a él, medir la longitud del programa por sus líneas de código (contar el número de caracteres).

En la teoría clásica de la probabilidad, teoría de la que Kolmogorov es referente directo, la noción informal de aleatoriedad se circunscribía a un plano ontológico, en tanto que era de secuencias generadas por procesos *estocásticos* o *azarosos* que se hablaba de aleatorias. Por ejemplo, una secuencia generada por el lanzamiento de una moneda  $n$  veces, produciría  $2n$  secuencias, todas iguales de aleatorias. Si  $n = 10$ , secuencias tales como 0000000000, 1111111111 o 1010101010 eran concebidas igual de aleatorias que 1100111000, 1111101001 o 0000010011 (Chaitin, 2002, p. 58). Con la introducción de lo que podría

entenderse como un plano epistemológico, gracias a la teoría de la complejidad algorítmica, secuencias como las primeras dejan de serlo en este plano, pues podrían generarse por programas de longitud bastante cortas: “*Imprima 1, 10 veces*” o “*Imprima 0, 10 veces*”.

Se dice entonces que Kolmogorov y Chaitin formalizan la noción de aleatoriedad para caracterizar secuencias sin estructura o patrón reconocible en aquellas generadas por procesos estocásticos o azarosos. Por lo que la teoría de la complejidad algorítmica en sus comienzos lo que permite es asentar una definición de *aleatoriedad*. Con relación a ella vieron estos dos autores la importancia de la idea de la longitud de un programa.

Esta definición de aleatoriedad está estrechamente relacionada con uno de los aspectos principales que trata la teoría de la información formulada por Shannon, a saber, el de la eficiencia o efectividad (*efficient*) en el proceso de transmisión de la información (Shannon, 1948; Chaitin, 2007, p. 8). En la exposición hecha por Shannon se establece que para una fuente de información que no presente un máximo de entropía es posible comprimir algunos de sus mensajes. Es decir, para una fuente que presente redundancia o que no sea completamente aleatoria, es posible siempre generar esos mismos mensajes con un menor requerimiento de información. De ahí que la teoría de la complejidad algorítmica reciba también el nombre de teoría de la información algorítmica, pues intenta establecer el mínimo de información que se requiere para generar una secuencia de bits.

Justamente este mínimo representará la complejidad de la secuencia computada. De manera que una secuencia de bits será más compleja que otra en tanto la longitud del programa más corto que la computa sea mayor que la del programa de menor longitud que computa a la otra. Esta definición de complejidad está sujeta al lenguaje de programación que se emplee. En el caso de Chaitin el lenguaje utilizado fue LISP.

Sin embargo, muy a pesar de que en efecto se pueda calcular la complejidad de una secuencia de bits, la conceptualización hecha por Chaitin, según sus propias palabras (Chaitin, 2002, p. 64), no buscaba esto con fines concretos; de hecho, considera que en términos estrictos la complejidad no es calculable. Lo que se busca con esta medida es más bien indicar los límites del conocimiento. Propiamente, es una reinterpretación del teorema de Gödel y del desarrollo de Turing.

### II.3.2. *Tres tipos de azar*

En un artículo publicado hace ya más de un siglo, bajo el título de '*The Doctrine of Necessity Examined*' (1892), Charles Sanders Peirce, en contra de una visión determinista en la naturaleza, y resaltando por su parte la importancia del factor azar en la diversificación y especificación que en ella se genera de manera continua, propone una serie de ideas que lo presentarían como un exponente referencial de lo que podría llamarse *Teoría de la Teleología Natural*, que es otra de las posibles vías, además por ejemplo de la teoría de evolución por selección natural, de análisis para encarar el problema de la complejificación en la naturaleza (Rescher, 1998, p. 5). En este artículo, Peirce aboga por una complejidad creciente. Así, en un apartado donde propone un intercambio de ideas con una persona inclinada hacia el determinismo, sostiene lo siguiente:

“Muy bien, mi estimado oponente, hemos alcanzado ahora un asunto problemático. Usted piensa que todas las especificaciones arbitrarias del universo se introdujeron en una sola dosis, al principio, si hubo un principio, y que la variedad y complejidad de la naturaleza siempre ha sido tan grande como ahora. Yo, por otra parte, creo que la diversificación y la especificación ha tenido lugar continuamente. Si condescendiera en preguntarme por qué pienso eso, debería dar mis razones de la siguiente manera:

Cuestione cualquier ciencia que se ocupe del curso del tiempo. Considere la vida de un animal o planta individual, o de una mente. Eche un vistazo a la historia de los estados, de las instituciones, del lenguaje, de las ideas. Examine las sucesiones de formas mostradas por la paleontología; la historia del globo tal como se expone en geología; lo que el astrónomo es capaz de comprender acerca de los cambios de los sistemas estelares. En todos estos casos, el hecho principal es el surgimiento y crecimiento de la complejidad. [...] De estos hechos amplios y omnipresentes podemos inferir con justicia, mediante la lógica más irreprochable, que probablemente exista en la naturaleza algún agente mediante el cual se pueda incrementar la complejidad y diversidad de las cosas; y que, en consecuencia, la regla de la necesidad mecánica encuentra, de alguna manera, interferencia.”<sup>39</sup> (Pierce, 1892, p. 333)

Azar y espontaneidad serían entonces para Pierce los factores que intervendrían en la generación de más complejidad en la naturaleza. El punto crítico quizás de esta postura es que le atribuye a la naturaleza una tendencia inherente hacia una mayor complejidad en el curso del tiempo. No se proponen mecanismos específicos a partir de los cuales, con la intervención del azar, se pueda generar un aumento de complejidad; sino que este ocurre más bien de manera espontánea. Atribuir esta característica teleológica a la naturaleza, en relación

---

<sup>39</sup> *Very well, my obliging opponent, we have now reached an issue. You think all the arbitrary specifications of the universe were introduced in one dose, in the beginning, if there was a beginning, and that the variety and complication of nature has always been just as much as it is now. But I, for my part, think that the diversification, the specification, has been continually taking place. Should you condescend to ask me why I so think, I should give my reasons as follows:*

*Question any science which deals with the course of time. Consider the life of an individual animal or plant, or of a mind. Glance at the history of states, of institutions, of language, of ideas. Examine the successions of forms shown by paleontology; the history of the globe as set forth in geology, of what the astronomer is able to make out concerning the changes of stellar systems. Everywhere the main fact is growth and increasing complexity. [...] From these broad and ubiquitous facts we may fairly infer, by the most unexceptionable logic, that there is probably in nature some agency by which the complexity and diversity of things can be increased; and that consequently the rule of mechanical necessity meets in some way with interference* (Pierce, 1892, p. 333).

con el problema de intentar dar cuenta de cómo surge y aumenta la complejidad en la misma, no nos posiciona en un mejor lugar. Es un enfoque descriptivo mas no explicativo.

Siguiendo la misma línea trazada por Pierce de considerar el azar un factor principal en el intento de dar cuenta de una posible complejidad creciente en la naturaleza podría mencionarse la *Teoría del Azar-Más-Estabilidad* como un enfoque ligeramente alternativo, o quizás mejor, complementario al de Pierce. Este enfoque es ya uno de corte plenamente naturalista. En una posible explicación que se ofrezca desde él, no se apela ni a un tipo de inteligencia supra-natural, como sucede con la Teoría del diseño inteligente, ni a una especie de teleología intra-natural, como en el caso de Pierce. La explicación se daría más bien considerando la problemática como una cuestión de probabilidad. *La mera fluctuación fortuita de las cosas ocasionalmente abre nuevos dominios para que la complejidad aumente.* Pero, además, y es la característica más importante que nos ofrece este enfoque, aquí se adiciona la idea la *estabilidad temporal*. Desde este enfoque se considera que una de las características de la complejidad es que le representa hasta cierto punto una suerte de estabilidad a las cosas en las que se hace manifiesta; hasta que un evento aleatorio interviene, generalmente a grandes intervalos de tiempo. La complejidad traería aparejado la estabilidad; estabilidad, claro está, definida bajo ciertos límites. Y es a partir de esta relación entre ambas que, de manera ocasional, se genera un aumento de complejidad en la naturaleza. Desde este enfoque podría decirse que la complejidad sería ella misma la que posibilita su aumento.

No se está diciendo que la complejidad se auto-potencie. Lo que se quiere decir más bien es que la complejidad tendría, entre otras características, la de hacer posible su aumento. Una de las características de la complejidad, por todo lo que ella implica y por todo lo que trae

consigo aparejado, es la de posibilitar su aumento; posibilidad en la que la idea de estabilidad temporal juega un papel muy importante.

Los sistemas complejos son sistemas que presentan comportamientos caóticos<sup>40</sup>, que, debido a su extrema sensibilidad a las condiciones iniciales, suelen experimentar variaciones en sus dinámicas debido a la intervención del azar; aunque, no necesariamente, claro está, a causa únicamente de esto. Pero es innegable el rol que como agente de cambio en las dinámicas de los sistemas complejos suele presentar. Sin embargo, no parece estar muy claro qué se entiende propiamente por azar, en tanto “la palabra «azar» es notablemente polisémica.” (Bunge, 2007, p. 143). Bunge nos plantea que se debe “distinguir al menos tres clases diferentes de azar: el accidente o encuentro casual, el desorden y el evento espontáneo o no causado.” (ibid., 2007, p. 143).

Desde hace poco más de 22 siglos el filósofo estoico Crisipo de Solos nos planteaba que un encuentro casual o accidente o coincidencia era el cruce de líneas causales inicialmente independientes. Como por ejemplo cuando se haya un tesoro mientras se cava para plantar un árbol, o, quizás más tradicionalmente, cuando nos tropezamos con algún conocido en algún lugar. También son ejemplos típicos de esta clase de azar la colisión involuntaria de dos vehículos, *la exaptación* y *la serendipia* (ibid., p. 143).

El papel de esta clase de azar en la evolución biológica cobra, según algunos autores destacados como François Jacob (1976), Stephen Jay Gould (2002), entre otros, un papel ciertamente no menor. Algunas bioespecies, “en particular la nuestra, podrían no haber emergido si no hubiese sido por una combinación accidental de mutaciones genéticas y condiciones ambientales favorables” (Bunge, 2007, p. 144). La evolución biológica parece

---

<sup>40</sup> Cabe resaltar que la caoticidad no es permanente. Los sistemas complejos presentan comportamientos caóticos bajo ciertos parámetros y bajo ciertos dominios temporales.

ser, antes que un proceso guiado o coordinado por algún agente o algunos agentes, más bien un proceso exploratorio de *concoirs de circonstances*. Circunstancias coincidentes que representan oportunidades únicas e irrepetibles para el surgimiento y crecimiento de la complejidad. Como sucede por ejemplo cuando un rasgo termina siendo apropiado para realizar ciertas funciones de las que tenía originalmente y que elevan la aptitud biológica. Que es a lo que algunos han dado en llamar exaptación. (ibid., 2007, p. 144)

Asimismo, algunos descubrimientos científicos han sido accidentes con suerte, como el caso de los rayos X, el de la radiactividad, el de la penicilina, etc. (ibid., 2007, p. 144)

No cabe entonces duda de que el accidente, la contingencia o el encuentro casual ha posibilitado y marcado no sólo nuestro desarrollo biológico, cultural y científico, sino también el desarrollo general de la evolución biológica.

El azar, como desorden o aleatoriedad, es aquel que se encuentra por ejemplo en el ruido blanco, la propagación de impulsos nerviosos a lo largo de un axón, la dispersión de genes y semillas, los resultados de los juegos de azar (ibid., p. 145). En suma, en todas aquellas situaciones en las que lo que se genera no parece ser comprimible en un esquema fácilmente identificable; lo que se genera de estos procesos presenta un alto contenido de complejidad algorítmica. Este es el tipo de azar con el que lidia, de manera particular, por ejemplo, la mecánica estadística, la estadística notarial, la sociología de la conducta marginal (ibid., p. 145), etc.

Este tipo de azar se considera que fue plenamente identificado a comienzos de la Revolución Científica; revolución que, como se suele estimar, tiene en Galileo quizás el referente fundacional más representativo. Galileo fue justamente uno de los primeros en identificar, aunque parezca un oxímoron, una suerte de leyes del azar. Descartes, Pascal, Fermat y el caballero de Méré fueron igualmente pioneros en esta dirección (ibid., p. 145).

Sin embargo, quizás quien se lleve la parte del león en el tratamiento de este tipo de azar sea Ludwig Boltzmann. La razón es que Boltzmann compatibilizó, nada más y nada menos que la mecánica, pináculo del determinismo y de la pretensión de predictibilidad, con un tratamiento estadístico, haciendo por ejemplo de dicha pretensión de predictibilidad un asunto ahora de carácter probabilístico.

Lo que Boltzmann nos planteó básicamente con su tratamiento estadístico de gases ideales fue que:

“Aunque las partículas microscópicas que conforman un gas siguen trayectorias newtonianas perfectamente predecibles, el sistema completo es tan grande y tan complicado que no resulta práctico resolver los sistemas de ecuaciones que los describen. Lo que sí se puede hacer es considerar las componentes microscópicas del gas como si fueran agentes aleatorios y hacer predicciones de las propiedades macroscópicas del gas como la temperatura, la presión, etc. (lo cual, por cierto, se logró llevar a cabo con enorme éxito). Tenemos así un enfoque determinista en principio, pero estadístico en la práctica.” (Fontanelli et al, 2020, p. 12).

Es decir, “que el azar en un nivel puede ser el resultado de la causalidad en un nivel inferior (Bunge, 1951). Esto no convierte el azar en algo ilusorio: únicamente lo hace relativo al nivel de organización” (Bunge, 2007, p. 146).

Una tercera clase de azar es aquella vislumbrada por Epicuro poco menos del siglo tres antes de Cristo, y que Lucrecio dio en llamar *clinamen*. Para Epicuro, este tipo de azar consistía en la desviación espontánea de los átomos en su trayectoria rectilínea (ibid., 2007, p. 146). Azar se está entiendo en este sentido como movimiento errático de carácter espontáneo. Este es el tipo de azar presente, por ejemplo, de manera conspicua, en la

mecánica cuántica. Piénsese “en las desintegraciones radiante y radioactiva espontánea” (ibid., p. 147) o en las fluctuaciones de vacío atribuidas al electrón. (ibid., p. 147).

### II.3.3. Orden estadístico: teorema del límite central

Queda entonces de manifiesto que el azar, sea de este último tipo, o sea de los otros dos mencionados, es un agente activo de nuestra naturaleza, y que, como tal, debe ser seriamente estudiado, ampliando así nuestras nociones de determinación con leyes de tipo probabilistas. El azar, antes que un asunto de ignorancia u otro de carácter subjetivo, es más bien un aspecto no trivial de nuestra naturaleza, que se manifiesta principalmente mediante regularidades estadísticas. Irregularidad, desorden, y confluencia de líneas causales independientes en los componentes individuales de una colección, abordados mediante leyes probabilistas, dan lugar a una suerte de orden en el agregado o colección. Un ejemplo muy significativo es el famoso teorema del límite central. Que nos dice que:

Sean  $X_1, X_2, \dots, X_n$  variables aleatorias continuas independientes, cuyo número crece a infinito, y de varianza finita, “entonces la suma normalizada de todas ellas (es decir, su promedio) converge a una variable aleatoria con distribución normal (y será exactamente la distribución normal cuando la muestra sea infinita)” (Fontanelli et al, 2020, p. 17)

Lo que quiere decir que esta suma:

$$X_{norm} \equiv \frac{\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma^2}}$$

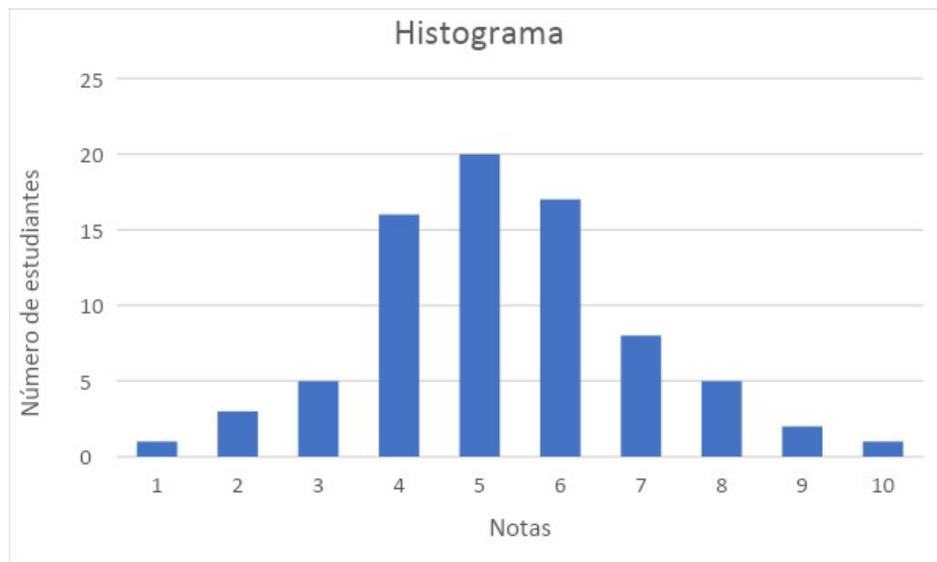
converge a una variable con función de distribución normal.

La distribución de una variable aleatoria podemos entenderla como una forma de descripción comprimida del comportamiento de los valores que dicha variable puede tomar. Así, por ejemplo, si la variable en consideración es la altura de los miembros de una población dada, la distribución de esa variable son todos los posibles valores que puede tomar dentro de dicha población con sus respectivas probabilidades de ocurrencia. En este caso particular los valores que puede tomar dicha variable están acotados tanto inferiormente, como superiormente.

En una distribución normal las probabilidades de ocurrencia de un evento se concentran en valores cercanos a la media de la población; o mejor, de la muestra. Entendiendo por muestra un subconjunto de la población. Ejemplos típicos de distribución normal además de la altura, son las notas en un examen, la presión sanguínea, errores de medida, entre otros.

Si se hace una tabla de frecuencia con alguna de estas variables, ocurren cosas bastante interesantes. Por ejemplo, que la media coincide con la moda; además, hay simetría con respecto a la media, y la desviación nunca es muy grande.

Con fines meramente ilustrativos considérese la situación en un salón de clases con relación a las notas obtenidas por los estudiantes. Lo que regularmente ocurre es que muy pocos estudiantes reprueban con la nota mínima, así como muy pocos aprueban con la nota máxima. Las notas se suelen acumular más bien alrededor de la media. Siendo la media incluso la moda.



Una situación como esta puede ser descrita con la ayuda de una función ideada por quien fuese uno de los matemáticos más destacados de todos los tiempos: el gran Carl Friedrich Gauss. Gauss se ideó una función cuya curva arropase o cubriera, junto al eje  $x$ , todos estos posibles valores que la variable aleatoria puede tomar en una situación como la descrita.

La función es la siguiente:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

A esta función se le conoce como *función de densidad de probabilidad*; donde  $\mu$  es la media y  $\sigma^2$  es la variancia ( $\sigma$  es la desviación). La curva de esta función, que tiene forma de una especie de campana, la hace merecedora del nombre alternativo de “campana de Gauss”.

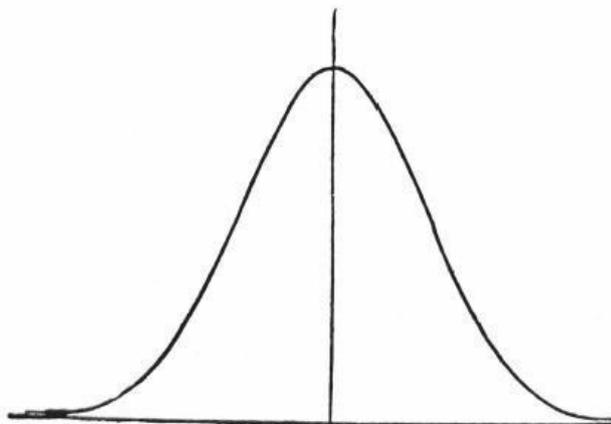


FIG. 80.—Curva normal de la probabilidad.

Figura tomada de (Kasner y Newman, 1988, p. 256)

A este tipo de distribuciones se le conoce como distribuciones “bien portadas”; sin embargo, hay distribuciones en las que la probabilidad de que la variable tome valores extremos no es despreciable. A este tipo de distribuciones se las conoce como distribuciones de cola pesada. Ejemplos de ellas son: la distribución de Cauchy, la distribución de Pareto, la distribución de Lévy, la de Burr, entre otras.

Entonces, el teorema del límite central lo que nos dice es que independientemente de la distribución que se esté trabajando, si se toman submuestras suficientemente grandes<sup>41</sup>, la distribución que siguen las medias de cada una de estas muestras será la distribución normal. Es decir, no importa qué tipo de distribución siga el parámetro o la variable considerada dentro de la población de estudio, que, si se toman submuestras suficientemente grandes y se estudia el comportamiento de las medias de dichas submuestras, estas seguirán siempre una

---

<sup>41</sup> Se entiende regularmente por ‘suficientemente grande’ submuestras con  $n \geq 30$ . Es decir, submuestras con mínimo 30 eventos.

distribución normal. Constituyendo así un muy bonito ejemplo de orden estadístico a partir del azar. Este resultado, junto con otros que nos brindan la teoría de la probabilidad y la estadística, cobra una gran importancia en el terreno de los modelos y simulaciones computacionales, en el que incursionaremos en el capítulo IV, particularmente para el método de simulación de Monte Carlo, bastante discutido en la literatura filosófica acerca de dichas herramientas en la investigación científica.

#### II. 4. *Complejidad modelo paramétrica ( $C^*$ )*

En el marco de la dinámica no lineal y la teoría del caos podemos encontrar una propuesta de abordaje de la complejidad, y particularmente de la emergencia, acorde a los objetivos pretendidos en esta tesis; propuesta además que hace uso de la teoría de la probabilidad y la estadística, y que se relaciona positivamente con una idea concerniente al proceso evolutivo de que las rupturas espontáneas de simetría han sido un factor distinguible en el *aumento* de complejidad entre las especies biológicas. Quizás el aspecto más representativo de esta propuesta sea su dependencia de la modelación matemática, vía, principalmente, ecuaciones diferenciales parciales.

“El nivel macroscópico de descripción es usualmente el más familiar en el que se manifiesta la complejidad, ya sea a través de la generación de soluciones (que se modelan *vía* ecuaciones matemáticas en el campo de la ciencia de la complejidad), que muestran propiedades estructurales y dinámicas inesperadas o de la dinámica propia del sistema, como es el caso del caos clásico. [...] Los procesos de bifurcación son uno de los elementos básicos de la complejidad, ya que proporciona el mecanismo fundamental para la aparición de nuevas posibilidades a partir de regímenes correspondientes a comportamientos simples del sistema.” (Fuentes, 2018, p. 89)

Pensemos en el siguiente ejemplo. En el ciclo vital del organismo unicelular *Acetabularia acetabulum*, un alga verde miembro del orden *Dasycladades*, que es un grupo de algas muy antiguo reducido ahora a una veintena de especies, se da algo muy llamativo, y es la formación, en la que invierte una cantidad considerable de recursos, de delicadas estructuras verticiladas sin que parezcan desempeñar alguna función. (Goodwin, 1998, p. 140)

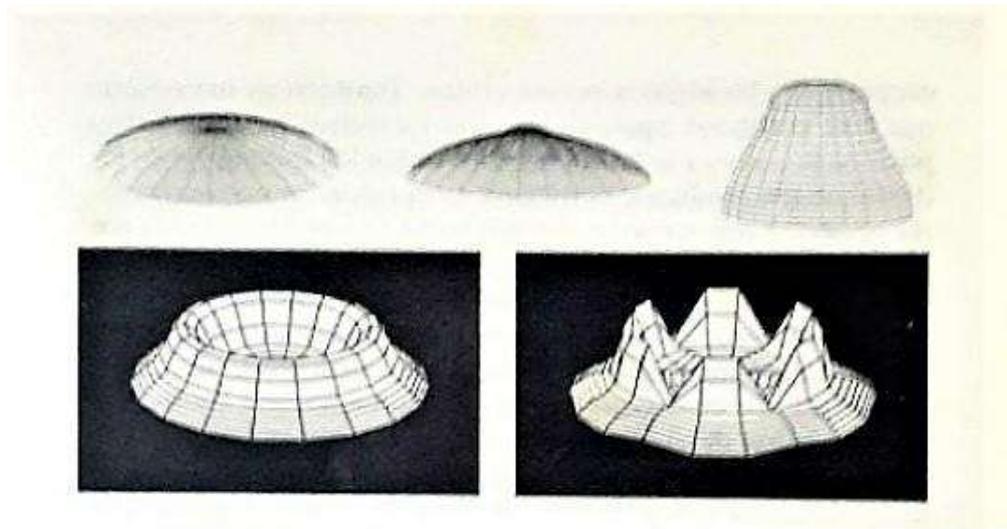
“¿por qué se producen estas estructuras? Ante este apasionante rompecabezas, los biólogos recurren a una explicación histórica. Aunque en *Acetabularia* no tienen ninguna función conocida, probablemente los verticilos fueron alguna vez funcionales en los ancestros de estas algas. [...] [D]e manera que los verticilos de *Acetabularia* son el resultado de un patrón ontogenético ancestral persistente heredado de un antecesor común. [...] [L]os verticilos se siguen formando por una especie de inercia morfogenética, y el precio que se paga por ello no es tan alto como para que la selección natural haya eliminado los verticilos o la misma *Acetabularia*. Pero esta explicación no sirve. [...] [I]nvoacar un ancestro común no explica cómo surgieron los verticilos en un principio. La propuesta de que la inercia morfogenética resiste la selección natural tampoco es una explicación, ya que simplemente redescrive el fenómeno en otros términos. [...] He aquí un interesante problema sin resolver.” (ibid., 1998, pp. 141, 142)

Sobre este problema las simulaciones computacionales, vía modelamiento matemático, han dado grandes luces. Manipulando los valores de ciertos parámetros se pueden observar, haciendo correr la simulación, cómo estas delicadas formas resultaban ser representaciones de soluciones estables (atractor) de un sistema dinámico representado mediante un sistema de ecuaciones diferenciales parciales.

“Una forma de investigar esta clase de problema es construir un modelo del proceso ontogenético que genera la forma. Esto requiere una teoría de campos morfogenéticos. [...] El núcleo del modelo es la descripción del citoplasma como un medio excitable en el que puede haber una ruptura espontánea de simetría con generación de patrones espaciales en las variables primarias, que son la concentración de calcio libre en el citoplasma y la tensión mecánica (grado de estiramiento o compresión), estando ligadas ambas variables por las propiedades del citoesqueleto.

[...] Desde el punto de vista biológico el modelo es muy simple, aunque matemáticamente es un *complejo* sistema no lineal de ecuaciones diferenciales parciales acopladas. Hay 26 parámetros, aunque la dinámica es sensible a cambios en sólo 6 de ellos. Los valores de los parámetros se escogieron según dos criterios: (1) están dentro del intervalo de bifurcación, de manera que puedan surgir patrones espontáneos; (2) las longitudes de onda de los patrones son menores que el dominio de crecimiento y regeneración, de manera que se pueden desarrollar formas interesantes. Después se dejó que el modelo evolucionara a su aire produciendo cualquier forma que pudiera generar. Aquí había un espacio paramétrico muy grande por explorar en términos de valores que satisfagan nuestros criterios, y no teníamos ni idea de cuánto tiempo nos llevaría encontrar algo interesante, ni siquiera si surgiría algo remotamente parecido a formas algales.

Resultó más fácil de lo que esperábamos encontrar conjuntos de parámetros que daban lugar a un tallo a partir de una esfera (un cigoto) o un hemisferio (un ápice regenerativo). Pero lo más sorprendente fue ver que estos «tallos» generaban verticilos en su proceso de crecimiento, pero que nunca habían sido explicados.” (ibid., 1998, pp. 143, 144)



Modelo informático del desarrollo de los verticilos de *Acetabularia*  
 Figura tomada de (Goodwin, 1998, p. 145)

“[La propuesta de la *complejidad modelo paramétrica* ( $C^*$ ) consiste en asignarle, vía una función], a cada valor del parámetro de control de un determinado modelo utilizado para explicar el sistema en estudio, [...] el valor de la complejidad de Kolmogorov.

La complejidad de Kolmogorov se calcula utilizando solo la señal en los datos y descartando los ruidos. La discriminación entre señal y ruido se realiza de manera tal que el set de datos que se extrae como ‘señal’ sean datos típicos del sistema, es decir, sean estadísticamente significativos para ser tomados como tales, mientras que el ‘ruido’ no cumple estas características estadísticas.” (Fuentes, 2018, p. 106)

Tratando de interpretar esta propuesta a la luz del ejemplo ofrecido del alga unicelular *Acetabularia acetabulum*, se dirá que en primer lugar habría que identificar un parámetro cuya dinámica del sistema en consideración sea sumamente sensitiva a su variación. Que en este caso sería, principalmente, la concentración de calcio libre en el citoplasma; parámetro,

valga la pena enfatizar, frente a la dinámica de formación de estructuras verticiladas en la parte superior del tallo.

Variando los valores de dicho parámetro, en su estrecha relación principalmente con la tensión mecánica del citoplasma (grado de estiramiento o compresión), el sistema presenta alteraciones en su estabilidad. Por ejemplo:

“Sabemos que [estas] algas pueden crecer perfectamente sin producir verticilos, cosa que sucede cuando la concentración de calcio en el agua se reduce a  $2mM$ , una quinta parte de su valor normal en el agua de mar (Goodwin et al., 1983). Si, cuando ya ha alcanzado algunos centímetros, se vuelven a colocar estas células sin verticilos, consistentes sólo en un rizoide y un tallo, en agua de mar con una concentración de calcio de  $10mM$ , forman parasoles y pueden completar un ciclo vital normal.”

(Goodwin, 1998, p. 140)

La idea sería entonces, posicionándonos en cada uno de estos valores, aplicar la complejidad de Kolmogorov al sistema dinámico en consideración, centrándonos en algún aspecto estadísticamente significativo, y el resultado de dicha aplicación sería un estimativo de la complejidad del sistema en dicho punto. Por ejemplo, si nos posicionamos en un valor de la concentración de calcio libre en el citoplasma de  $2mM$ , la complejidad de Kolmogorov arrojaría un valor menor que la que arrojaría si nos posicionáramos en un valor de la concentración de calcio libre de  $10mM$ . Esto, porque el programa que describiría el sistema en el segundo caso emplearía un número mayor de líneas de código que el que se emplearía para describir el sistema en el primer caso. La descripción del sistema en el segundo caso involucra una estructura física -el parasol- que en el primer caso no figura. Por lo que su descripción sería un poco más larga que la que se pueda ofrecer en el primer caso.

#### II.4.1. *Propiedad emergente modelo paramétrica (E\*)*

En estrecha relación con la propuesta de la *complejidad modelo paramétrica (C\*)*, se presenta a continuación una alternativa de abordaje de propiedades emergentes; a saber, el abordaje modelo paramétrico de propiedad emergente, o quizás en mejores palabras: *propiedad emergente modelo paramétrica (E\*)*.

Decimos entonces que:

“Una propiedad será emergente, ( $E^*$ ), para el valor del parámetro de control si la complejidad modelo paramétrica ( $C^*$ ) en ese valor del parámetro de control presenta una discontinuidad de tipo escalonado, es decir si se cumple que a un lado del valor del parámetro de control la complejidad es notoriamente menor que al otro lado del parámetro de control.” (Fuentes, 2018, p. 107)

Esto quiere decir que, por ejemplo, en el caso de *Acetabularia*, la formación de verticilos es una propiedad emergente modelo paramétrica en  $10mM$ .

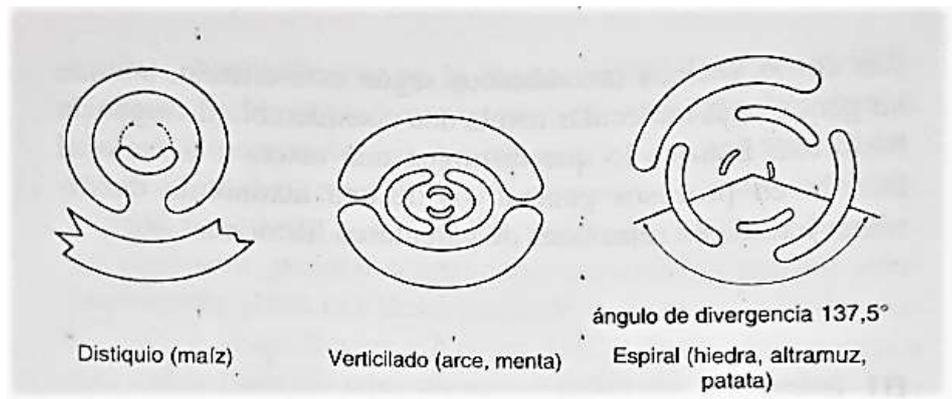
Esta propuesta de abordaje, tanto de la complejidad, como de la emergencia, en términos modelo paramétricos, encuentra, por ejemplo, en el ámbito de la evolución biológica, ciertamente no pocos casos relevantes.

La explicación de los distintos tipos de filotaxis es uno de ellos. Hay tres grandes tipos de disposición foliar: verticilado, helicoidal y dístico (Goodwin, 1998, p. 149). En el modelamiento del meristema como sistema dinámico todo parece indicar que estos tres tipos de patrones foliares se corresponderían a tres atractores (morfogenéticos) en el sistema de ecuaciones con las que se intenta llevar a cabo dicho modelamiento (ibid., 1998, p. 149).

“Un modelo reciente de la filotaxis, obra de Douady y Couder [...], da otro paso significativo en esta dirección. Lo que han conseguido demostrar es la naturaleza más que probable de la bifurcación con ruptura de simetría en el campo morfogenético del

meristema que resulta en los patrones filotáticos principales, y al mismo tiempo demuestran que las series menores surgen como discontinuidades de las series principales. Emplearon un modelo físico simple de generación de patrones consistente en hacer gotear un ferrofluido en el centro de un disco cubierto de una película de aceite en la que las gotas quedaban flotando. Un campo magnético polarizaba las gotas convirtiéndolas en pequeños dipolos magnéticos que se repelían entre sí. El campo morfogenético del meristema era así simulado por un campo magnético. A medida que las gotas caían en el centro del disco eran repelidas por las gotas polarizadas ya presentes, y quedaban expuestas a un campo magnético estacionario que las empujaba hacia el borde del disco. El resultado es que surgen diferentes patrones dependiendo de las condiciones del experimento, pero todos corresponden a patrones filotáticos observados.” (ibid., p. 150)

En este caso el parámetro de control propuesto es un número sin dimensiones que Douady y Couder (1992) definieron como  $G = V_0 \frac{T}{R_0}$ , “donde  $V$  es la rapidez con que las gotas se alejan del centro del disco en respuesta a un campo magnético fijo,  $T$  es el período entre gota y gota, y es  $R_0$  el radio de la región que corresponde al centro del meristema y alrededor de la cual se generan los primordios foliares” (ibid., p. 153).



Patrones filotácticos vistos desde arriba  
 Figura tomada de (Goodwin, 1998, p. 150)

La funcionalidad de las proteínas es otro claro ejemplo en el ámbito de la biología de propiedad emergente modelo paramétrica ( $E^*$ ). Sus aminoácidos constituyentes no logran desempeñar estas funciones. Es a partir de sus interacciones, y de una configuración estructural bien específica -que se da de estas interacciones-, que surge dicha funcionalidad. La complejidad de Kolmogorov, aplicada sobre el conjunto de valores paramétricos<sup>42</sup>, que representan en la simulación computacional la configuración estructural específica de una proteína, arrojaría un mayor valor que si se aplica sobre un conjunto de valores paramétricos distintos. Ya que en el primer caso la descripción del sistema incluiría dicha función. En ese conjunto de valores paramétricos específico la complejidad modelo paramétrica ( $C^*$ ) es mayor que en otro conjunto de valores paramétricos. Su complejidad efectiva, que vendría siendo la complejidad de Kolmogorov discriminando el ruido, sería mayor. Pues habría regularidades que comprimir a partir de su función, que harían del esquema en el que se

<sup>42</sup> Sobre este conjunto de valores hablaremos más adelante, en el marco de las redes neuronales artificiales de tipo *Deep learning*. La simulación que acá se está considerando está posibilitada por un modelo de redes neuronales artificiales.

encapsulen dichas regularidades uno más largo que los esquemas en los que no se compriman estas regularidades.

Con lo anterior se está intentando resaltar la importancia y aplicabilidad de la propuesta de la complejidad y emergencia modelo paramétricas en el ámbito de la biología, por ser este quizás el terreno donde la noción de complejidad se encuentra más incorporada. Como habíamos mencionado anteriormente: en biología la noción de evolución se suele asociar al *aumento* de la complejidad y a la formación de estructuras y funciones cada vez más complejas.

La propuesta entonces de la complejidad y la emergencia modelo paramétricas, nos permite, no sólo un acercamiento teórico -de carácter métrico, principalmente- a estas nociones, sino que además se presenta como una explicación complementaria, en el ámbito de la biología evolutiva, a ciertos rasgos fenotípicos difíciles de abordar únicamente bajo el accionar de mecanismos evolutivos. En no pocas ocasiones, antes que ser el resultado propiamente del accionar de los distintos mecanismos evolutivos, ciertos rasgos fenotípicos en varias especies pueden verse más bien como resultado de soluciones estables a sus dinámicas internas. En estas explicaciones, vía modelamiento matemático, la identificación de parámetros, como se ha insistido, se hace indispensable.

Sin embargo, para el caso por ejemplo del problema del plegamiento de las proteínas, la identificación de este conjunto de valores paramétricos no parece posible mediante métodos analíticos. Y es ahí donde se hace muy valiosa y significativa la posibilidad brindada por las redes neuronales artificiales de aprendizaje profundo; como veremos a continuación.

En el siguiente capítulo, en estrecha relación con esta propuesta, presentaremos la alternativa de abordaje de la complejidad mediante redes neuronales artificiales de tipo *Deep learning*; siendo el caso del éxito abrumador de *AlphaFold* nuestro ejemplo fundamental.

## CAPÍTULO III

### REDES NEURONALES ARTIFICIALES Y EL PROBLEMA DEL PLEGAMIENTO DE LAS PROTEÍNAS

*¿Se está produciendo un giro hacia la predicción en ciencia?*

#### Abstract

En este capítulo analizaremos el poder predictivo de las simulaciones computacionales de la complejidad mediante redes neuronales artificiales de aprendizaje profundo. Presentaremos el estudio de caso de *AlphaFold*. Nos permitirá enfatizar el compromiso manipulacionista la distintiva centralidad de la predicción científica en este tipo de enfoque.

En este capítulo presentaremos una propuesta de abordaje de la complejidad igualmente en términos modelo paramétricos. Sin embargo, el modelamiento aquí pensado vendría principalmente dado por redes neuronales artificiales de aprendizaje profundo. Se habla de propuesta modelo paramétrica en el sentido de que acá igualmente se hace fundamental identificar ciertas variables, que, bajo ciertos valores, permiten establecer una correlación con el fenómeno a abordar. Adicional a esto, en la propuesta aquí pensada se pone especial énfasis en su carácter «manipulacionista», vía modelo y simulación computacional, principalmente de los valores de dichos parámetros, hasta encontrar, por ensayo y error, una correlación entre estos y el fenómeno o la propiedad -emergente- a abordar; búsqueda por ensayo y error optimizada por los algoritmos de ‘*backpropagation*’ y ‘*stochastic gradient descent*’.

El ejemplo fundamental en este capítulo, e incluso de la tesis misma que acá se intenta defender, ha de ser el *AlphaFold*, y su éxito abrumador en la *predicción* de las estructuras de un amplio número de proteínas a partir de sus secuencias de aminoácidos. Predicción nada fácil de conseguir, y que ha representado un gran y significativo problema en el ámbito de la biología. El ejemplo de *AlphaFold* no sólo nos sirve para intentar evidenciar la gran relevancia que cobran las redes neuronales artificiales para el abordaje de la complejidad, sino que además nos sirve para evidenciar una suerte de giro epistémico hacia la predicción que se viene dando en ciencias en años recientes, como bien lo hacen notar, por ejemplo, Mazviita Chirimuuta (2018) y Daniel A. Weiskopf (2022).

Comenzaremos este capítulo introduciendo el marco general en el que se insertan las redes neuronales artificiales; que son las protagonistas de este capítulo. Más concretamente, iniciaremos el capítulo planteando los tipos de aprendizaje considerados paradigmáticos dentro del *Machine learning*.

### III.1. *Paradigmas de aprendizaje dentro del campo del 'Machine learning'*

Si bien podría considerarse que nuestra realidad es estructurada; que hay orden dentro de ella, y que, por tanto, hay cierto margen de predictibilidad, también es evidente que es caótica, compleja y que las señales que recibimos de ella son señales que viajan a través de canales ruidosos. En el procesamiento de señales debemos enfrentarnos a una cantidad considerable de ruido. Aun así, logramos efectivamente hacer predicciones; logramos identificar patrones, y, con base en estos, desplegamos ciertos tipos de comportamientos.

Cuando se habla de aprendizaje se está hablando, primeramente, de procesamiento de información; procesamiento en el que se intenta, regularmente, identificar patrones a partir de un cúmulo de señales sumamente finas. Señales que viajan a través de canales ruidosos.

La identificación de estos patrones se realiza expresamente con fines predictivos. Lo que se pretende es, a partir de las regularidades o patrones percibidos del entorno, poner en marcha ciertos comportamientos con el fin de sortear las presiones selectivas que este imparta. Lo que quiere decir que, el despliegue de los esquemas en los que se encapsulen las regularidades percibidas tiene un carácter adaptativo. Si, con los esquemas desplegados, no logran sortearse las presiones selectivas que el entorno imparta, de no ser estos modificados, el resultado puede ser que la selección natural termine eliminado al sistema en cuestión.

Aprendizaje es, en términos generales, identificación de patrones; comprensión de regularidades. En los intentos por dar cuenta de cómo, por ejemplo, nuestro cerebro comprime regularidades, han surgido tres grandes paradigmas de análisis en el campo de la inteligencia artificial. Estos tres paradigmas son: el '*paradigma del aprendizaje supervisado*', el '*paradigma del aprendizaje no supervisado*' y el '*paradigma del aprendizaje reforzado*'. En cada uno de estos lo que está a la base es la identificación de patrones. Patrones que de una u otra forma permitan realizar predicciones acerca del medio.

### III.1.1. *Aprendizaje supervisado*

Por aprendizaje supervisado se entiende básicamente el tipo de aprendizaje en el que se logra establecer un patrón entre unas variables de entrada y unas variables de salida. Matemáticamente hablando lo que se logra bajo este paradigma es establecer una función que procese los valores de entrada, en los valores de salida dados. Bajo este tipo de aprendizaje se logran establecer correlaciones entre un conjunto de datos de entrada y un conjunto de datos de salida. Siendo ambos conjuntos unos, generalmente, muy grandes. Se requiere, regularmente, de conjuntos muy grandes de entrada y de salida para un establecimiento confiable de una correlación entre los elementos de ambos conjuntos. El

establecimiento de esta correlación, que requiere la identificación de ciertos patrones, opera regularmente sobre un cúmulo de señales sumamente finas comparado con el nivel de resolución que presentan los datos. La característica de supervisado se debe al hecho de que los datos de salida son dados. Esto de alguna manera representa una guía en el proceso de aprendizaje. Al suministrarle al sistema el conjunto de salida, se le esté guiando hacia él, representando esto un tipo de supervisión en su aprendizaje.

### III.1.2 *Aprendizaje no supervisado*

El aprendizaje no supervisado, por su parte, es uno en el que se no se suministra dato de salida alguno. No hay ninguna pauta a seguir sobre la cual intentar establecer algún tipo de correlación. Bajo este paradigma no hay ningún tipo de guía o supervisión sobre el tipo de respuesta que se deba generar. Lo que se realiza, regularmente, bajo este paradigma, es una búsqueda de patrones de similitud entre los datos de entrada. A partir de estos patrones de similitud el sistema genera lo que en el *machine learning* se conoce como *espacio latente*. Que es un espacio constituido de variables latentes. Estas son variables implícitas; son variables no observables pero que se pueden inferir a partir de variables observables.

Piénsese por ejemplo en la siguiente imagen como valor de una variable de entrada



Las siguientes imágenes representan valores de variables latentes

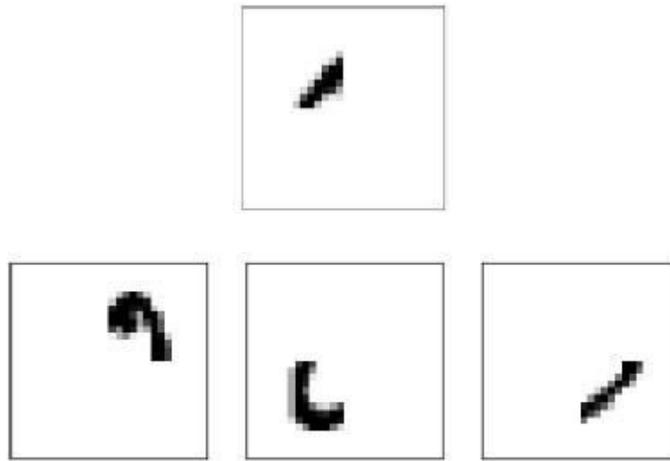


Figura tomada de (Nielsen, 2015, p, 14)

La variable observable en este ejemplo es de tipo numérica. Esa variable observable de tipo numérica puede expresarse por medio de variables no categorizables de manera sencilla; por medio de variables latentes. Estas son variables, por decirlo de alguna manera, ocultas, en el sentido que son variables implícitas. Son variables que aportan pequeños trozos de información para la conformación de una variable con un carácter observable y definida de manera explícita.

Quizás una manera un poco más clara de plantear esto de espacio y variables latentes sea diciendo que el espacio latente constituye el espacio matemático de representación de un concepto, y las variables latentes son aquellas variables involucradas en las características básicas de tales conceptos. Piénsese por ejemplo en el concepto de silla. El espacio latente sería el espacio constituido por todos esos puntos que permiten representar distintas formas y tipos de sillas. Las variables latentes serían aquellas variables involucradas en las

características básicas que permiten generar esas distintas formas y tipos de sillas.  
Características básicas identificadas por el sistema.

### III.1.3 *Aprendizaje reforzado*

Por último, el aprendizaje reforzado, como su nombre lo indica, es un tipo de aprendizaje donde, por decirlo de alguna manera, se premia cierto tipo de respuestas. Asimismo, hay en él algún tipo de penalización de no ofrecer respuestas acordes a la situación problema a la que el sistema está sujeto. Esto quiere decir que el mecanismo básico tras este tipo de aprendizaje es uno de penalización y premiación con relación a una presión selectiva dada. Y lo que se despliega en este tipo de aprendizaje son estrategias de ataque a la situación problema planteada. Es decir, aquí la respuesta como tal es una serie de comportamientos que constituyen un plan de acción, o un plan de ruta para la resolución de cierto tipo de problema.

El ejemplo inmediato que se podría dar es el tipo de aprendizaje que se desarrolla en el contexto de juegos como el ajedrez, el Go y todo tipo de juegos que impliquen el despliegue de estrategias. Acá, claramente, no hay una única respuesta. No hay una única estrategia conducente a una victoria. Además, el tipo de estrategia desplegada en un momento específico es una entre muchísimas alternativas en competencia. El abanico de posibilidades estratégicas es extremadamente grande. Y las posibles combinaciones entre los elementos involucrados en la situación problema es prácticamente inconmensurable. No es posible establecer una estrategia explorando las posibles combinaciones de elementos involucrados. Se exploran sólo algunas posibles alternativas de planes de acción según los resultados de premiación o penalización que se vayan obteniendo. Esto hace que el aprendizaje sea aquí principalmente una cuestión de ensayo y error.

No hay manera deductivamente de explorar estrategias conducentes a una victoria, en tanto que eso implicaría contemplar las posibles consecuencias de un número extremadamente grande de combinaciones de piezas. Lo que opera más bien es el reconocimiento de patrones de similitud (analogías) en las partidas. Patrones de similitud generados a partir de la integración de pequeños trozos de información. No suelen ser elementos (piezas) particulares donde se centra la atención, sino en configuraciones estructurales con cierto nivel de generalidad. Las estrategias se van desarrollando a partir de la identificación de similitudes de ciertos patrones estructurales entre partidas. Patrones asociados, o bien con un peso positivo, o bien con uno negativo. Es decir, patrones vinculados a una derrota tendrían una valoración negativa y patrones vinculados a una victoria tendrían una valoración positiva (*reinforcement*).

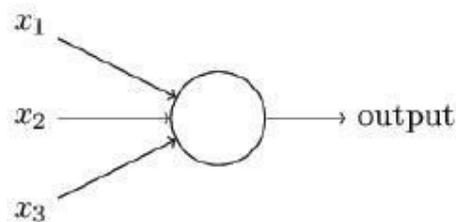
### III.2 *Redes neuronales artificiales de aprendizaje profundo*

Dentro del *Machine Learning*, que es una de las disciplinas de la Inteligencia Artificial, existe una familia de algoritmos que se conoce bajo el nombre de '*Redes neuronales artificiales*'. Esta familia de algoritmos está constituida de una enorme cantidad de tipos. Sin embargo, nuestra atención se centrará aquí en la familia de algoritmos que se conoce como redes neuronales artificiales de tipo *Deep learning*. Se les considera de aprendizaje profundo, básicamente, por el nivel de procesamiento que ocurre en sus capas ocultas. En estos algoritmos el aprendizaje se da de manera jerarquizada. Es por la estructuración del procesamiento de información, en varios niveles, que estos algoritmos reciben este nombre. Los outputs en estos algoritmos son producto de un procesamiento por niveles de la información percibida en los inputs. El aprendizaje se da por un procesamiento a distintos niveles, simulando así la manera en la que nuestro cerebro procesa la información que

percibe. El procesamiento de información en nuestro neocórtex se da de manera jerárquica. Y, además, con una alta dosis de retroalimentación; característica también notoria en esta familia de algoritmos.

En las redes neuronales artificiales la unidad básica de procesamiento, igual que en nuestro cerebro, presenta un funcionamiento interno relativamente sencillo. Una neurona, en el terreno de la inteligencia artificial, es básicamente un modelo de regresión lineal. Es una suma ponderada de sus estímulos externos o valores de entrada. Estímulos externos que recibe por medio de unas conexiones de entrada; pero, además, presenta una conexión de salida. Por medio de esta conexión de salida la neurona expresa el resultado de su procesamiento interno de la información que recibe. Este resultado constituirá, a menos que sea una neurona de la capa de salida, un estímulo externo para la neurona o para las neuronas a la que esté o a las que esté conectada. A este modelo básico de neurona se le conoce con el nombre de *perceptrón*.

“Entonces, ¿cómo funcionan los perceptrones? Un perceptrón toma varias entradas binarias,  $x_1, x_2, \dots$ , y produce una única salida binaria:



En la imagen mostrada el perceptrón recibe tres inputs  $x_1, x_2, x_3$ . En general, podría recibir más o menos variables de entrada. Rosenblatt propuso una simple regla para

computar el valor de la variable de salida. Introdujo unas variables que tomaban valores del conjunto de los reales que expresaban la importancia de cada input en el valor de la variable de salida, que se conocen como pesos  $w_1, w_2, \dots$ . El valor de 0 o 1 de la neurona de salida estaba determinado por si la suma ponderada  $\sum_j w_j x_j$  era mayor o menor que cierto valor umbral. De igual forma que en el caso de los pesos, el umbral es un número real, que vendría siendo un parámetro de la neurona. Para poner esto en términos algebraicos más precisos:

$$\text{output} = \begin{cases} 0 & \text{if } \sum_j w_j x_j \leq \text{threshold} \\ 1 & \text{if } \sum_j w_j x_j > \text{threshold} \end{cases} \quad (1)$$

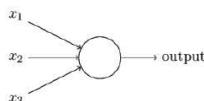
¡Esto es todo lo que está involucrado en el funcionamiento de un perceptrón!

Este es el modelo matemático básico. Una manera de concebir el perceptrón es como un dispositivo que toma decisiones ponderando la evidencia de la que dispone.”<sup>43</sup>

(Nielsen, 2015, p. 3)

---

<sup>43</sup>So how do perceptrons work? A perceptron takes several binary inputs,  $x_1, x_2, \dots$ , and produces a single binary output:



In the example shown the perceptron has three inputs,  $x_1, x_2, x_3$ . In general it could have more or fewer inputs. Rosenblatt proposed a simple rule to compute the output. He introduced weights,  $w_1, w_2, \dots$ , real numbers expressing the importance of the respective inputs to the output. The neuron's output, 0 or 1, is determined by whether the weighted sum  $\sum_j w_j x_j$  is less than or greater than some threshold value. Just like the weights, the threshold is a real number which is a parameter of the neuron. To put it in more precise algebraic terms:

$$\text{output} = \begin{cases} 0 & \text{if } \sum_j w_j x_j \leq \text{threshold} \\ 1 & \text{if } \sum_j w_j x_j > \text{threshold} \end{cases} \quad (1)$$

That's all there is to how a perceptron works!

That's the basic mathematical model. A way you can think about the perceptron is that it's a device that makes decisions by weighing up evidence. (Nielsen, 2015, p. 3)

Si bien el procesamiento interno que realiza el perceptrón es bastante sencillo, lo cierto es que con este modelo se puede procesar la información necesaria para llevar a cabo problemas de clasificación, de selección, de decisión, entre otros. Con una neurona se puede modelar el funcionamiento de una compuerta lógica AND, OR y NAND. No obstante, si se quiere modelar el funcionamiento de una compuerta lógica XOR, una sola neurona no basta. Esto quiere decir que para modelar situaciones un poco más complicadas se requiere de una concatenación de neuronas. Para el caso de una compuerta lógica XOR, con dos neuronas la situación se puede modelar.

El problema de la concatenación de neuronas, bajo el modelo básico del perceptrón, es que el modelo termina colapsando en una regresión lineal. La concatenación de funciones lineales termina siendo otra función lineal. Por lo que, modelar situaciones no lineales, requiere agregar cierta deformidad del modelo de regresión lineal, característico del perceptrón. A aquellas funciones que confieren no linealidad dentro del modelo interno básico del perceptrón se les llama *'funciones de activación'*; siendo la función *'sigmoide'* una de las más utilizadas. Otras funciones de activación comunes son *'la tangente hiperbólica'*, *'la función escalonada'* y *'la unidad rectificadora lineal'*. Estas funciones de activación permiten la concatenación de varias neuronas, evitando que dicha concatenación termine colapsando en un modelo de regresión lineal.

La función sigmoide, que se define de la siguiente manera:

$$\sigma(x) = \frac{1}{1 + e^x}$$

al ser pasada por la suma ponderada básica del modelo del perceptrón, da lugar al modelo de neurona que se conoce como '*modelo sigmoide*'. Este modelo es muy útil, principalmente, por la ventaja que ofrece de representar probabilidades. La gráfica de la función nos permite observar que, para valores grandes, la distorsión que produce conlleva a que estos vayan al valor uno (1), mientras que para valores muy pequeños la distorsión hace que vayan al valor cero (0).

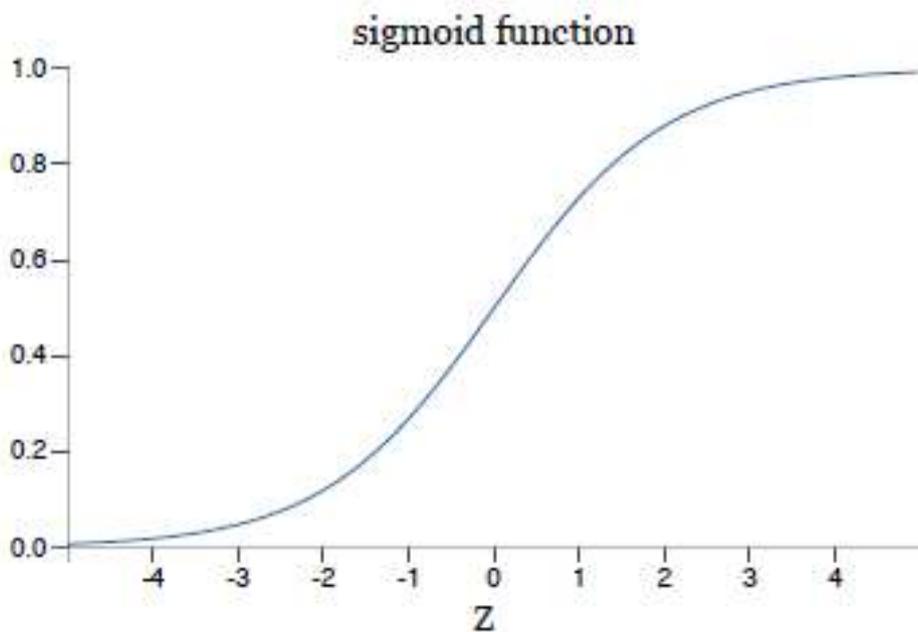


Figura tomada de (ibid., 2015, p. 9)

Acá,  $Z$ , vendría representando la suma ponderada de los parámetros de la neurona; es decir, los pesos, y el factor de sesgo o *bias*.

$$Z = \sum_j w_j x_j + b,$$

donde  $b$  es el factor de sesgo.

En el caso de la función sigmoide, el output que generaría la neurona vendría dado por:

$$\sigma(Z) \equiv \frac{1}{1 + e^Z}$$

Otra manera de expresarlo sería:

$$\sigma(Z) = \frac{1}{1 + \exp(-\sum_j w_j x_j + b)}$$

Con las funciones de activación, que se suelen representar con la letra ‘ $a$ ’, se puede procesar información de manera no lineal. La interconexión no lineal de elementos es, como hemos venido presentando, una característica importante de la complejidad. Es a partir de las interrelaciones no lineales de elementos que emergen comportamientos y respuestas de sistemas estimados complejos. Las redes neuronales entonces presentan características distintivas de la complejidad; en el sentido, principalmente, de que a partir de sus estructuraciones podría considerarse que dan lugar a propiedades emergentes. Por ejemplo, una neurona por sí sola no puede aprender a diferenciar de una imagen el rostro de una persona, mientras que una red de ellas sí. Es a partir de la interconexión entre ellas que pueden desplegar esquemas y respuestas inteligentes.

Las neuronas se estructuran en capas, habiendo tres tipos de ellas: la capa de entrada, las capas ocultas y la capa de salida. La profundidad de las capas ocultas da nombre al tipo de redes que acá nos interesa. Son redes en las que se procesan muchísimos pequeños trozos de información en bloque dando lugar a un aprendizaje jerárquico.

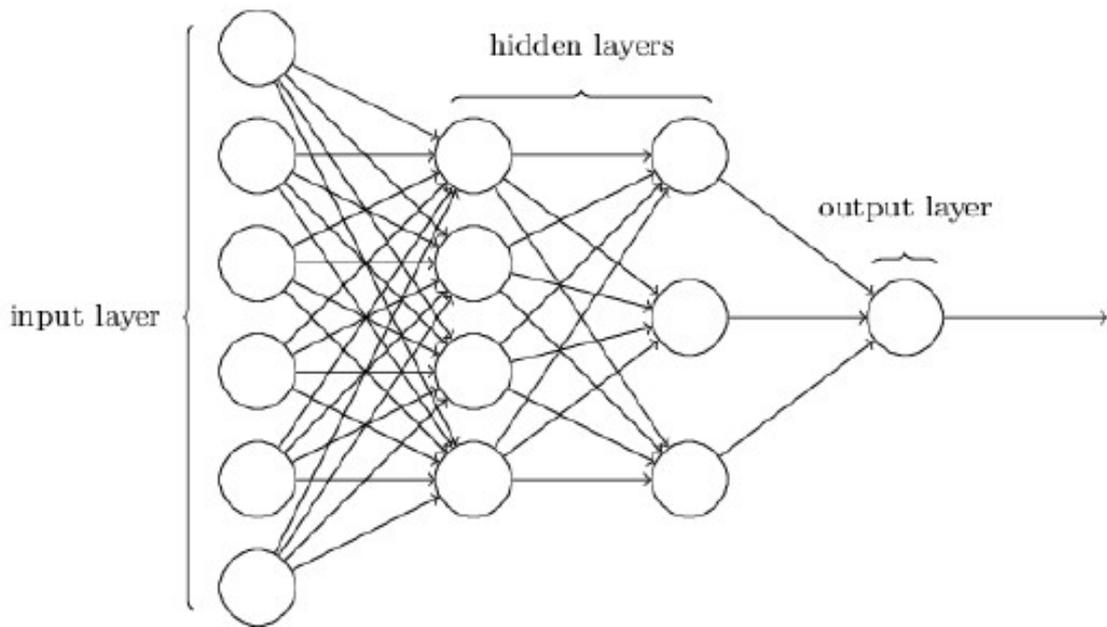
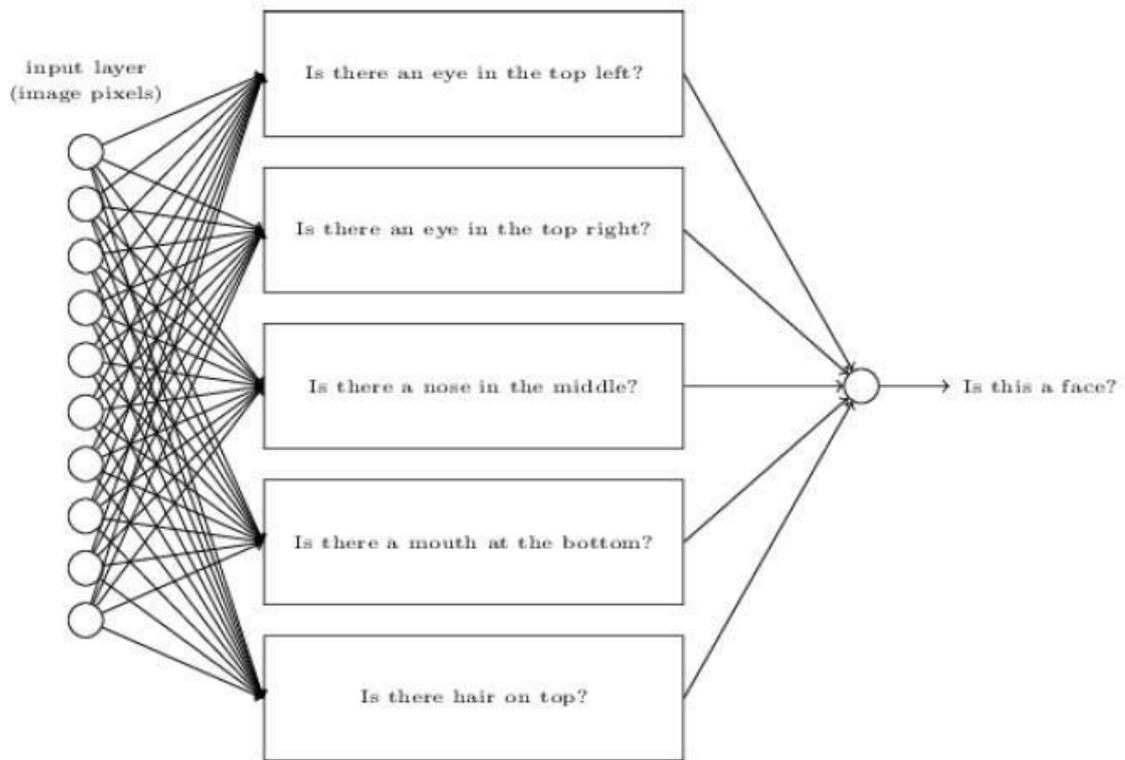


Figura tomada de (ibid., 2015, p. 11)

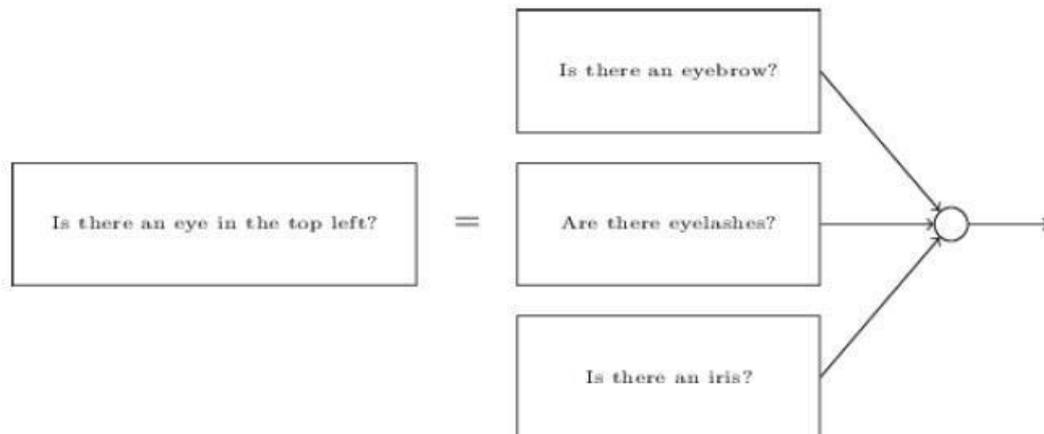
Veamos un ejemplo

“Supongamos que queremos determinar si una imagen muestra un rostro humano o no. Podríamos atacar este problema de la misma manera que atacamos el reconocimiento de escritura a mano: utilizando los píxeles de la imagen como entrada a una red neuronal, con la salida de la red siendo una sola neurona que indique: “Sí, es una cara” o “No, no es una cara”. Supongamos que hacemos esto, pero que no estamos usando un algoritmo de aprendizaje. En su lugar, intentaremos diseñar una red a mano, eligiendo pesos y sesgos adecuados. ¿Cómo podríamos hacerlo? Olvidando por completo las redes neuronales por el momento, una heurística que podríamos utilizar sería descomponer el problema en subproblemas: ¿la imagen tiene un ojo en la parte superior izquierda? ¿Tiene un ojo en la parte superior derecha? ¿Tiene una nariz en el medio? ¿Tiene una boca en el medio inferior? ¿Hay pelo en la parte superior? Etcétera.

Si las respuestas a varias de estas preguntas son “sí”, o incluso simplemente “probablemente sí”, entonces concluiríamos que es probable que la imagen sea una cara. Por el contrario, si las respuestas a la mayoría de las preguntas son “no”, entonces la imagen probablemente no sea una cara. Por supuesto, esto es solo una heurística aproximada y adolece de muchas deficiencias. Tal vez la persona sea calva, por lo que no tiene pelo. Tal vez solo podamos ver una parte de la cara, o la cara está en cierto ángulo que hace que algunas de las características faciales estén oscurecidas. Aun así, la heurística sugiere que, si podemos resolver los subproblemas usando redes neuronales, entonces quizás podamos construir una red neuronal para la detección de rostros, combinando las redes para los subproblemas. Aquí hay una posible arquitectura, con rectángulos que denotan las subredes. Tenga en cuenta que esto no pretende ser un enfoque realista para resolver el problema de detección de rostros; más bien, es para ayudarnos a desarrollar una intuición sobre cómo funcionan las redes. Aquí está la arquitectura:



También es plausible que las subredes se puedan descomponer. Supongamos que estamos considerando la pregunta: “¿hay un ojo en la parte superior izquierda?” Esto se puede descomponer en preguntas como: “¿hay una ceja?”; “¿hay pestañas?”; “¿hay un iris?”; etcétera. Por supuesto, estas preguntas también deberían incluir información posicional: “¿está la ceja en la parte superior izquierda y por encima del iris?”. Pero hagámoslo simple. La red para responder a la pregunta “¿Hay un ojo en la parte superior izquierda?” ahora se puede descomponer:



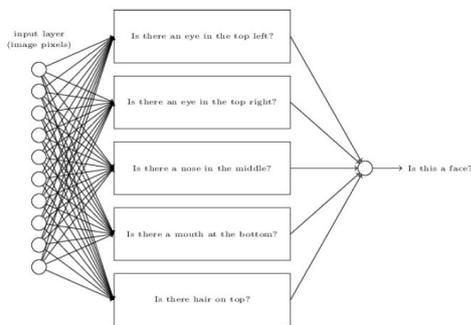
Esas preguntas también pueden desglosarse, cada vez más, a través de múltiples capas. En última instancia, trabajaremos con subredes que responden preguntas tan simples que pueden responderse fácilmente a nivel de píxeles individuales. Esas preguntas podrían, por ejemplo, referirse a la presencia o ausencia de formas muy simples en puntos particulares de la imagen. Estas preguntas pueden responderse mediante neuronas individuales conectadas a los píxeles sin procesar de la imagen.

El resultado final es una red que desglosa una pregunta muy complicada - ¿esta imagen muestra una cara o no? - en preguntas muy simples que se pueden responder a nivel de píxeles individuales. Lo hace a través de una serie de muchas capas, con capas iniciales que responden preguntas muy simples y específicas sobre la imagen de entrada, y capas posteriores que construyen una jerarquía de conceptos cada vez más complejos y abstractos.”<sup>44</sup> (ibid., 2015, pp. 35, 36 y 37)

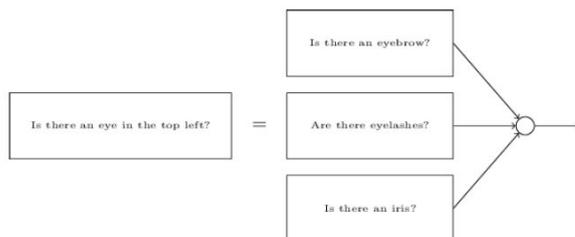
---

<sup>44</sup> *Suppose we want to determine whether an image shows a human face or not: We could attack this problem the same way we attacked handwriting recognition – by using the pixels in the image as input to a neural network, with the output from the network a single neuron indicating either “Yes, it’s a face” or “No, it’s not a face”. Let’s suppose we do this, but that we’re not using a learning algorithm. Instead, we’re going to try to design a network by hand, choosing appropriate weights and biases. How might we go about it? Forgetting neural networks entirely for the moment, a heuristic we could use is to decompose the problem into sub-*

problems: does the image have an eye in the top left? Does it have an eye in the top right? Does it have a nose in the middle? Does it have a mouth in the bottom middle? Is there hair on top? And so on. If the answers to several of these questions are “yes”, or even just “probably yes”, then we’d conclude that the image is likely to be a face. Conversely, if the answers to most of the questions are “no”, then the image probably isn’t a face. Of course, this is just a rough heuristic, and it suffers from many deficiencies. Maybe the person is bald, so they have no hair. Maybe we can only see part of the face, or the face is at an angle, so some of the facial features are obscured. Still, the heuristic suggests that if we can solve the sub-problems using neural networks, then perhaps we can build a neural network for face-detection, by combining the networks for the sub-problems. Here’s a possible architecture, with rectangles denoting the sub-networks. Note that this isn’t intended as a realistic approach to solving the face-detection problem; rather, it’s to help us build intuition about how networks function. Here’s the architecture:



It’s also plausible that the sub-networks can be decomposed. Suppose we’re considering the question: “Is there an eye in the top left?” This can be decomposed into questions such as: “Is there an eyebrow?”; “Are there eyelashes?”; “Is there an iris?”; and so on. Of course, these questions should really include positional information, as well—“Is the eyebrow in the top left, and above the iris?”, that kind of thing – but let’s keep it simple. The network to answer the question “Is there an eye in the top left?” can now be decomposed:



Those questions too can be broken down, further and further through multiple layers. Ultimately, we’ll be working with sub-networks that answer questions so simple they can easily be answered at the level of single pixels. Those questions might, for example, be about the presence or absence of very simple shapes at particular points in the image. Such questions can be answered by single neurons connected to the raw pixels in the image.

The end result is a network which breaks down a very complicated question – does this image show a face or not – into very simple questions answerable at the level of single pixels. It does this through a series of many layers, with early layers answering very simple and specific questions about the input image, and later layers building up a hierarchy of ever more complex and abstract concepts. (ibid., pp. 35, 36 y 37).

El ejemplo anterior sirve básicamente para ilustrar algunos aspectos importantes de este tipo de arquitecturas algorítmicas. Principalmente, poner de manifiesto el procesamiento de tipo jerárquico que se da en este tipo de redes y la profundidad en dicha estructura, razón por lo que reciben el nombre de *Deep learning*. Igualmente, de la gran cantidad de pequeños trozos de información por considerar para lograr el resultado deseado. Son muchos los parámetros que habría que ajustarse hasta lograr el resultado deseado. Pero, lo pretendido es que sea la propia red la que pueda ir autoajustándose, hasta lograr establecer un patrón que le permita, jerárquicamente, ir perfilando -pretendidamente con un bajo costo computacional- la abstracción conceptual -en este caso- de lo que constituye un rostro. De modo que, ir explorando de manera parcialmente exhaustiva las consecuencias de dichos ajustes resulta inmanejable. Se requiere entonces de un algoritmo que la red pueda usar para minimizar el costo computacional de ir explorando exhaustivamente las consecuencias de los ajustes que vaya realizando de sus parámetros.

### III.2.1 '*Backpropagation*' y '*gradient descent*' (*Retropropagación del error y el descenso del gradiente*)

Los algoritmos de '*Backpropagation*' y '*Gradient descent*' quizás sean lo más utilizados en el proceso de entrenamiento de las redes neuronales artificiales de aprendizaje profundo. De hecho, se estima que con la postulación del algoritmo de *backpropagation*, el campo de la inteligencia artificial cobra un nuevo aire; pues, este abrió la posibilidad de tratar nuevos problemas, muy difíciles de abordar a partir de las técnicas de entrenamiento previas. No había, previo a la formulación de este algoritmo, una manera de que la red se entrenase de forma eficiente a partir de autoajustar sus parámetros hasta conseguir resultados favorables, ya que no había una manera de optimizar esta exploración de resultados. Eran ajustes dados

a la fuerza bruta, por decirlo de alguna manera. Eran horas de exploración, muchas veces sin observar disminución en la función de coste del modelo. Es decir, eran varias horas de cómputo sin observar muchas veces una disminución en el error del modelo. Con la aparición del algoritmo de *backpropagation* se lograba mostrar experimentalmente cómo una red podía autoajustar sus parámetros para aprender así qué tipo de información era la que debía procesar.

“El algoritmo de retropropagación del error se introdujo originalmente en la década de 1970, pero su importancia no se apreció por completo hasta un famoso artículo de 1986 de David Rumelhart, Geoffrey Hinton y Ronald Williams. Ese artículo describe varias redes neuronales en las que la propagación inversa funciona mucho más rápido que los enfoques anteriores de aprendizaje, lo que hace posible el uso de redes neuronales para resolver problemas que anteriormente eran insolubles. Hoy en día, el algoritmo de retropropagación es el caballo de batalla del aprendizaje en redes neuronales [...]

En el corazón de la retropropagación hay una expresión para la derivada parcial  $\frac{\partial C}{\partial w}$  de la función de coste  $C$  con respecto a cualquier peso  $w$  (o sesgo  $b$ ) en la red. La expresión nos dice qué tan rápido cambia el coste cuando cambiamos los pesos y el sesgo. Y aunque la expresión es algo compleja, también tiene una belleza, y cada elemento tiene una interpretación natural e intuitiva. Por tanto, la propagación hacia atrás no es solo un algoritmo rápido de aprendizaje. En realidad, nos brinda información detallada

sobre cómo, al cambiar los pesos y el sesgo, cambia el comportamiento general de la red.”<sup>45</sup> (ibid., 2015, p. 39)

En el cálculo de las derivadas parciales de los parámetros con respecto al coste del modelo está el punto principal de este algoritmo. Es decir, lo que este algoritmo nos pide, en principio, es el cálculo de la derivada parcial del coste del modelo con respecto a los pesos de la neurona, y, la derivada parcial del coste del modelo con respecto al factor de sesgo o *bias*. En cuanto a la función de coste hay que decir que la oferta es variada, como pasa con la función de activación. Sin embargo, la función ‘*error cuadrático medio*’ es una de las más utilizadas. Así, si  $a$  es nuestra función de activación, entonces nuestra función de error, usando el error cuadrático medio, vendría dada por la siguiente expresión:

$$C(a_j^L) = \frac{1}{2} \sum_j (y_j - a_j^L)^2$$

Donde, el superíndice  $L$ , está referenciando la última capa de la red, y,  $y_j$  a las salidas (outputs) de las neuronas de la última capa. El cálculo de las derivadas parciales se realiza a partir de la última capa, pues, como su nombre lo indica, el objetivo de dicho algoritmo es retropropagar el error hacia las capas iniciales.

---

<sup>45</sup> *The backpropagation algorithm was originally introduced in the 1970s, but its importance wasn't fully appreciated until a famous 1986 paper by David Rumelhart, Geoffrey Hinton, and Ronald Williams. That paper describes several neural networks where backpropagation works far faster than earlier approaches to learning, making it possible to use neural nets to solve problems which had previously been insoluble. Today, the backpropagation algorithm is the workhorse of learning in neural network...*

*At the heart of backpropagation is an expression for the partial derivative  $\partial C/\partial w$  of the cost function  $C$  with respect to any weight  $w$  (or bias  $b$ ) in the network. The expression tells us how quickly the cost changes when we change the weights and biases. And while the expression is somewhat complex, it also has a beauty to it, with each element having a natural, intuitive interpretation. And so backpropagation isn't just a fast algorithm for learning. It actually gives us detailed insights into how changing the weights and biases changes the overall behaviour of the network. (ibid., 2015, p. 39)*

Para obtener la derivada parcial del coste con respecto a los parámetros mencionados, recuérdese que esta es pasada por la función de activación, y esta última a su vez es pasada por la suma ponderada a la que referenciamos con la letra  $Z$ . Lo que quiere decir que hay una serie de composición de funciones involucradas en el cálculo de dichas derivadas. Tenemos entonces que la función  $C$  recibe a la función  $a$ , que recibe a su vez a la función  $Z$ . Esto es,

$$C(a(Z^L)) = y^L$$

donde  $y^L$  vendría siendo el error del modelo, fijados únicamente en la última capa.

Por tanto, si se quiere calcular el valor de la derivada de  $C$ , con respecto a los parámetros considerados, se tiene que:

$$\frac{\partial C}{\partial w^L} = \frac{\partial C}{\partial a^L} \frac{\partial a^L}{\partial Z^L} \frac{\partial Z^L}{\partial w^L}$$

y

$$\frac{\partial C}{\partial b^L} = \frac{\partial C}{\partial a^L} \frac{\partial a^L}{\partial Z^L} \frac{\partial Z^L}{\partial b^L}$$

Acá, es importante anotar que  $Z^L$  se puede reescribir de la siguiente forma:

$$Z^L = W^L a^{L-1} + b^L,$$

donde  $W^L$  es la matriz de los pesos de las neuronas de la última capa, y  $a^{L-1}$  son los outputs de las neuronas de la capa anterior.

De las derivadas anteriores, se puede anotar que:

$$\frac{\partial C}{\partial a^L} \frac{\partial a^L}{\partial z^L} = \frac{\partial C}{\partial z^L}$$

representa el valor imputado a la neurona en el resultado; en el error del modelo. Este valor nos cuenta la responsabilidad de la neurona en el error del modelo, lo que quiere decir que, si el valor de la derivada es grande, la neurona en consideración tiene una gran responsabilidad en el error del modelo. Por el contrario, si el valor de la derivada es pequeño, quiere decir que su responsabilidad en dicho error es baja. De ahí que a dicho valor se le llame '*error imputado*' a la neurona; valor que se suele referenciar con la letra  $\delta$ . Esto permite reescribir nuestras derivadas parciales de la siguiente manera:

$$\frac{\partial C}{\partial w^L} = \delta^L \frac{\partial z^L}{\partial w^L}$$

$$\frac{\partial C}{\partial b^L} = \delta^L \frac{\partial z^L}{\partial b^L}$$

pero, además,

$$\frac{\partial z^L}{\partial w^L} = a_i^{L-1},$$

y

$$\frac{\partial z^L}{\partial b^L} = 1,$$

por lo que,

$$\frac{\partial C}{\partial w^L} = \delta^L a_i^{L-1}$$

$$\frac{\partial C}{\partial b^L} = \delta^L$$

Con esto tendríamos un indicador del error del modelo centrados en la última capa. Restaría realizar los cálculos en las  $L - 1$  capas restantes. Veamos cómo sería para la capa  $L - 1$ .

$$\frac{\partial C}{\partial w^{L-1}} = \frac{\partial C}{\partial a^L} \frac{\partial a^L}{\partial z^L} \frac{\partial z^L}{\partial a^{L-1}} \frac{\partial a^{L-1}}{\partial z^{L-1}} \frac{\partial z^{L-1}}{\partial w^{L-1}}$$

$$\frac{\partial C}{\partial b^{L-1}} = \frac{\partial C}{\partial a^L} \frac{\partial a^L}{\partial z^L} \frac{\partial z^L}{\partial a^{L-1}} \frac{\partial a^{L-1}}{\partial z^{L-1}} \frac{\partial z^{L-1}}{\partial b^{L-1}}$$

Si bien puede parecer que el cálculo se hace cada vez más largo y tedioso, lo cierto que es que gran parte de esas derivadas ya han sido calculadas en la capa previa.

$$\frac{\partial C}{\partial a^L} \frac{\partial a^L}{\partial z^L} = \delta^L$$

este producto ya ha sido calculado. Además,

$$\frac{\partial z^{L-1}}{\partial w^{L-1}} = a_j^{L-2},$$

$$\frac{\partial z^{L-1}}{\partial b^{L-1}} = 1.$$

La expresión:

$$\frac{\partial a^{L-1}}{\partial z^{L-1}}$$

es la derivada de la función de activación, y

$$\frac{\partial z^L}{\partial a^{L-1}}$$

nos dice cómo varía la suma ponderada de una capa conforme se ajustan los valores de salida de la capa previa. Esta derivada vendría siendo la matriz de parámetros que conectan dos capas contiguas. En este caso, la capa  $L$  y la capa  $L - 1$ . A esta matriz de parámetros se le suele referenciar con la letra  $W$ . En este caso concreto,  $W^L$ .

Ahora, la expresión

$$\frac{\partial C}{\partial a^L} \frac{\partial a^L}{\partial z^L} \frac{\partial z^L}{\partial a^{L-1}} \frac{\partial a^{L-1}}{\partial z^{L-1}}$$

vendría siendo

$$\frac{\partial C}{\partial z^{L-1}}$$

que, a su vez, vendría siendo

$$\delta^{L-1}$$

Así,

$$\frac{\partial C}{\partial a^L} \frac{\partial a^L}{\partial Z^L} \frac{\partial Z^L}{\partial a^{L-1}} \frac{\partial a^{L-1}}{\partial z^{L-1}} = \frac{\partial C}{\partial z^{L-1}} = \delta^{L-1}.$$

Este procedimiento ya es extensible al resto de capas. A partir de estas derivadas se puede ir retropropagando el error a las capas anteriores; encontrando las responsabilidades de cada neurona en dicho error. A partir de estas cuatro derivadas el algoritmo de *backpropagation* nos permite obtener el error *imputado* a cada neurona en el error del modelo; y, además, nos permite obtener el gradiente de la función de coste, resultado importante para poder utilizar el algoritmo de *stochastic gradient descent*, que es un algoritmo de optimización del aprendizaje en este tipo de redes. Las cuatro derivadas son:

$$1. \quad \frac{\partial C}{\partial a^L} \frac{\partial a^L}{\partial Z^L} = \delta^L$$

$$2. \quad \delta^{l-1} = W^l \delta^l \frac{\partial a^{l-1}}{\partial z^{l-1}}$$

$$3. \quad \frac{\partial C}{\partial w^{l-1}} = \delta^{l-1} a^{l-2}$$

$$4. \quad \frac{\partial C}{\partial b^{l-1}} = \delta^{l-1}$$

El algoritmo de *stochastic gradient descent* lo que nos ofrece básicamente es una estrategia de cómo ajustar los parámetros de un modelo para disminuir la función de coste de este. Con *backpropagation* la red tiene una manera de determinar cuáles son las neuronas mayormente implicadas en el resultado de error del modelo; con *stochastic gradient descent* la red tiene una manera de autoajustar los parámetros de esas neuronas mayormente implicadas en el error del modelo para así disminuirlo.

La lógica del algoritmo del descenso del gradiente, como su nombre lo indica, es descender iterativamente en la dirección del gradiente en los puntos (neuronas) fuertemente implicados en el error del modelo; gradiente que viene dado por la derivada parcial del coste del modelo en esos puntos (neuronas). Si el gradiente nos indica hacia dónde el error se incrementa, la idea es ir en sentido contrario a dicho gradiente.

La función de coste se puede representar en un plano tridimensional donde las coordenadas sean los dos tipos de parámetros involucrados en nuestras neuronas y el error del modelo en dichas neuronas. Así, en el eje  $x$  podemos ubicar a los factores de peso de las neuronas involucradas en el modelo, en el eje  $z$  los factores de sesgo, y en el eje  $y$  los valores de la función de coste. Lo que se haría entonces es ir descendiendo iterativamente sobre la superficie que representa la función de coste en sentido inverso del gradiente de la función en los puntos críticos, con el fin de ir encontrando los puntos mínimos de la función de coste. Es decir, los valores de los parámetros que hacen del error uno muy bajo. Con esto se tiene una estrategia de exploración eficiente de los resultados de ir ajustando los parámetros de las neuronas hasta encontrar el resultado deseado.

“Resumiendo, la forma en que funciona el algoritmo del descenso de gradiente consiste en calcular repetidamente el gradiente  $\nabla C$ , y luego moverse en la dirección opuesta, “cayendo” por la ladera del valle.”<sup>46</sup> (Nielsen, 2015, p. 27)

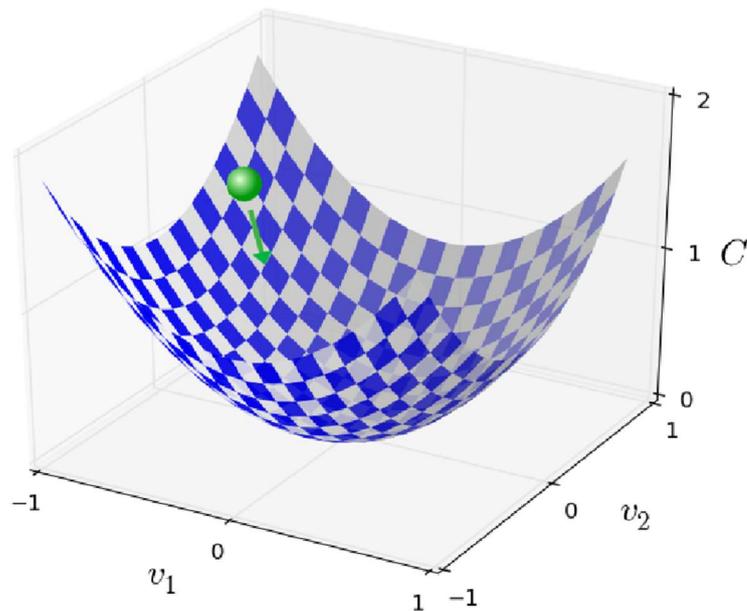


Figura tomada de (ibid., 2015, p. 27)

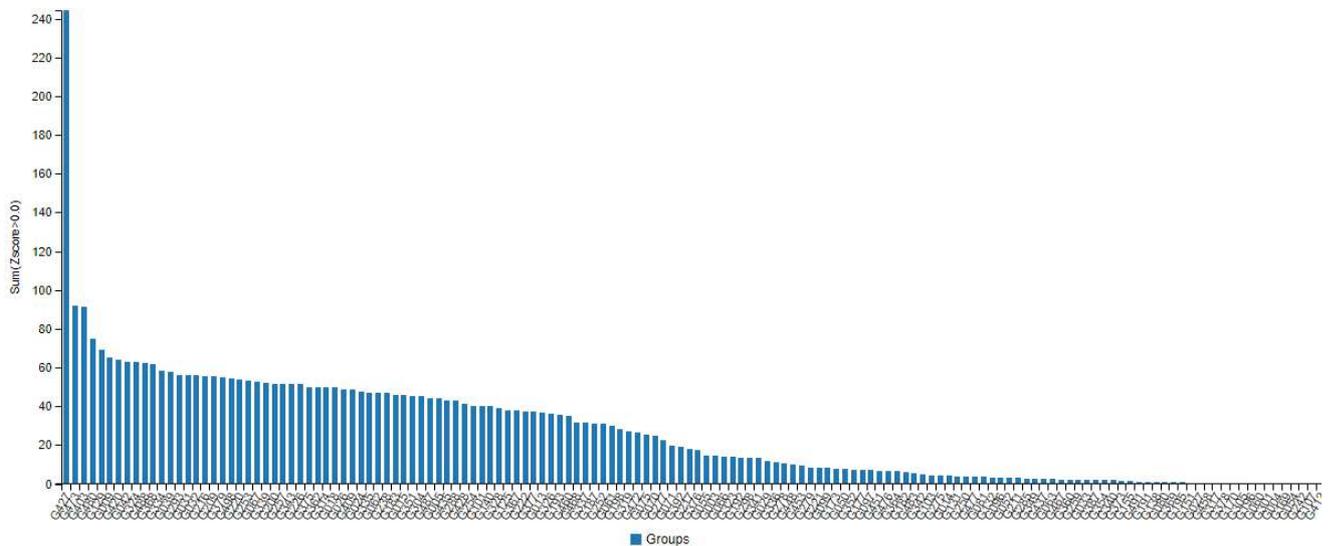
### III.3 El problema del plegamiento de las proteínas y su asombrosa resolución por parte de AlphaFold.

El problema del plegamiento de las proteínas ha sido uno que ha interesado sobremanera por décadas a los científicos en el área de la biología, la bioquímica, la biomedicina, etc. Tanto así, que desde el año 1994 ha venido realizándose un concurso (*Critical Assessment of Techniques for Protein Structure Prediction -CASP-*), donde se premia al grupo o centro

---

<sup>46</sup> Summing up, the way the gradient descent algorithm works is to repeatedly compute the gradient  $\nabla C$ , and then to move in the opposite direction, "falling down" the slope of the valley. (Nielsen, 2015, p. 27)

investigativo -con una buena suma de dinero- que logre, mediante algún procedimiento, *predecir* la estructura o configuración precisa de un amplio número de proteínas a partir de su secuencia de aminoácidos. Hasta el momento, el concurso se ha llevado a cabo en catorce ocasiones. En dos de ellas, -las dos últimas más concretamente (CAPS13: 2018 y CASP14: 2020<sup>47</sup>)-, ha participado *DeepMind*, que es un laboratorio de inteligencia artificial; siendo quizás el más prestigioso del mundo. Laboratorio adquirido por el gigante californiano (Google) en el 2014. En ambas ocasiones resultó ganador. Y ambas con un buen margen de ventaja. Siendo el obtenido en su segunda participación -con *AlphaFold2*- uno extremadamente amplio; sin duda, asombroso.



Gráfica tomada del: *Protein Structure Prediction Center*  
<https://predictioncenter.org/index.cgi>

<sup>47</sup> El concurso se realiza cada dos años. En el 2020 se llevó a cabo en dos momentos; uno expresamente concerniente a las proteínas involucradas en SAR-2-coV (Covid-19).

#	GR Code	GR name	Domains Count	SUM Zscore (>-2.0)	Rank SUM Zscore (>-2.0)	AVG Zscore (>-2.0)	Rank AVG Zscore (>-2.0)	SUM Zscore (>0.0)	Rank SUM Zscore (>0.0)	AVG Zscore (>0.0)	Rank AVG Zscore (>0.0)
1	427	AlphaFold2	92	244.0217	1	2.6524	1	244.0217	1	2.6524	1
2	473	BAKER	92	90.8241	2	0.9872	2	92.1241	2	1.0013	2
3	403	BAKER-experimental	92	88.9672	3	0.9670	3	91.4731	3	0.9943	3
4	480	FEIG-R2	92	72.5351	4	0.7884	4	74.5627	4	0.8105	4
5	129	Zhang	92	67.9065	5	0.7381	5	68.8922	5	0.7488	5
6	009	fFold_human	92	61.2858	7	0.6861	8	65.2157	6	0.7089	7
7	420	MULTICOM	92	63.2689	6	0.6877	7	64.0531	7	0.6962	8
8	042	QUARK	92	60.0226	10	0.6524	11	62.9711	8	0.6845	9
9	324	Zhang-Server	92	60.8875	8	0.6618	9	62.9122	9	0.6838	10
10	488	fFold-IDT_human	92	57.6435	11	0.6266	12	62.0795	10	0.6748	11
11	368	fFold-CaT_human	92	60.5423	9	0.6581	10	61.8464	11	0.6722	12
12	334	FEIG-R3	92	48.4424	20	0.5265	23	58.5809	12	0.6367	13
13	039	ropiusQQA	92	55.7086	12	0.6055	13	57.8135	13	0.6284	14
14	293	MUFOLD_H	92	47.7806	21	0.5194	24	55.9608	14	0.6083	15
15	031	Zhang-CEIthreader	92	49.5742	18	0.5389	21	55.9467	15	0.6081	16
16	032	MESHI	92	53.0953	14	0.5771	15	55.9047	16	0.6077	17
17	216	EMAP_CHAE	92	53.1597	13	0.5778	14	55.4235	17	0.6024	18
18	209	BAKER-ROSETTASERVER	92	46.1861	25	0.5020	28	55.2993	18	0.6011	19
19	379	Wallner	92	51.7365	15	0.5624	16	55.1852	19	0.5998	20
20	498	VoroMQA-select	92	51.4288	17	0.5590	18	54.5710	20	0.5932	21
21	220	McGuffin	92	49.5443	19	0.5385	22	53.7384	21	0.5841	22
22	253	Bhattacharya	92	51.5239	16	0.5600	17	52.9898	22	0.5760	23
23	067	ProQ2	92	47.7403	22	0.5189	25	52.3893	23	0.5694	25
24	339	ProQ3D	92	46.1913	24	0.5021	27	52.3390	24	0.5689	26
25	200	Bilbiu2020	90	44.6502	27	0.5406	20	51.7027	25	0.5745	24
26	257	P3De	92	42.0842	34	0.4574	36	51.4179	26	0.5589	27
27	343	VoroCNN-select	92	44.9630	26	0.4887	29	51.2603	27	0.5572	28
28	226	Zhang-TBM	92	46.2470	23	0.5027	26	51.2506	28	0.5571	29

Gráfica tomada del: *Protein Structure Prediction Center*  
<https://predictioncenter.org/index.cgi>

De las gráficas anteriores se puede observar el amplio margen de ventaja que toma *AlphaFold2* con relación a los demás. Por ejemplo, con relación al segundo puesto hay una diferencia de aproximadamente 150 puntos: 244,0217 para *AlphaFold2* y 92,1241 para BAKER. Todo esto sin ser *DeepMind* un centro especializado en bioquímica, biomedicina o biología.

Ahora, ¿por qué es tan relevante ese logro?

“Las proteínas son moléculas, grandes y complejas, esenciales para el sostenimiento de la vida. Casi todas las funciones que realiza nuestro cuerpo (contraer músculos, detectar la luz o convertir alimentos en energía) pueden ser rastreadas hasta una o más proteínas y cómo ellas se mueven o cambian. Las recetas para esas proteínas –llamadas genes– están codificadas en nuestro ADN. Lo que cualquier proteína dada puede hacer

depende de su estructura 3D única. Por ejemplo, los anticuerpos que forman nuestro sistema inmunológico tienen forma de “Y” y son similares a unos ganchos. Al aferrarse a virus y bacterias, estas proteínas anticuerpos pueden detectar y etiquetar microorganismos causantes de enfermedades para su exterminio [...]

Pero descubrir la forma 3D de una proteína exclusivamente a partir de su secuencia genética es una tarea compleja que los científicos han encontrado desafiante durante décadas. El desafío es que el ADN solo contiene información sobre la secuencia de los bloques de construcción de una proteína llamados residuos de aminoácidos, que forman largas cadenas. Predecir cómo esas cadenas se plegarán en la intrincada estructura 3D de una proteína es lo que se conoce como el “problema del plegamiento de proteínas”.

Cuanto más grande es la proteína, tanto más complicado y difícil resulta de modelar, pues hay más interacciones entre los aminoácidos para tener en cuenta. Como se señaló en la paradoja de Levinthal, tomaría más tiempo que la edad del universo enumerar todas las configuraciones posibles de una proteína típica antes de alcanzar la estructura 3D correcta.”<sup>48</sup> (Evans et al., 2018)

---

<sup>48</sup>*Proteins are large, complex molecules essential in sustaining life. Nearly every function our body performs—contracting muscles, sensing light, or turning food into energy—can be traced back to one or more proteins and how they move and change. The recipes for those proteins—called genes—are encoded in our DNA. What any given protein can do depends on its unique 3D structure. For example, antibody proteins that make up our immune systems are ‘Y-shaped’ and are akin to unique hooks. By latching on to viruses and bacteria, antibody proteins are able to detect and tag disease-causing microorganisms for extermination [...]*

*But figuring out the 3D shape of a protein purely from its genetic sequence is a complex task that scientists have found challenging for decades. The challenge is that DNA only contains information about the sequence of a protein’s building blocks called amino acid residues, which form long chains. Predicting how those chains will fold into the intricate 3D structure of a protein is what’s known as the “protein folding problem”.*

*The bigger the protein, the more complicated and difficult it is to model because there are more interactions between amino acids to take into account. As noted in [Levinthal’s paradox](#), it would take longer than the age of the universe to enumerate all the possible configurations of a typical protein before reaching the right 3D structure. (Evans et al., 2018)*

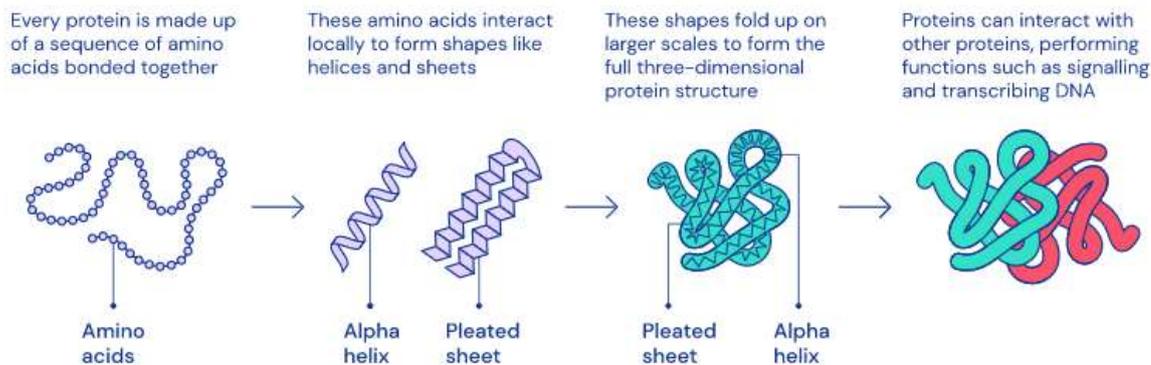


Figure 1: Complex 3D shapes emerge from a string of amino acids.

Figura tomada de la página de DeepMind:  
<https://deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery>

La red neural artificial de aprendizaje profundo utilizada para obtener un modelo tridimensional de la proteína, a partir de la secuencia de aminoácidos que la constituye, se conoce como *AlphaFold*. *AlphaFold* se origina básicamente como resultado del éxito abrumador obtenido por *DeepMind* con las versiones *AlphaGo* y *AlphaZero*, desarrolladas para jugar al Go y al ajedrez respectivamente; juegos en los que se evidencia que una exploración de todas las posibles partidas para determinar estrategias conducentes a una victoria resulta un reto de nunca acabar. Téngase en cuenta, como acercamiento preliminar, que el Go es un juego con  $19 \times 19$  posiciones o “casillas”, y que, además, tiene movimientos más flexibles que los del ajedrez. El abanico de posibilidades de combinaciones de movimientos es prácticamente inconmensurable.

Las posibilidades exploratorias de las distintas configuraciones tridimensionales que puede adquirir una proteína, a partir de la secuencia de aminoácidos que la constituye, representa un escenario de búsqueda, en esencia, similar al de la búsqueda de estrategias

favorables para la obtención de una victoria, bien sea en el ajedrez o en el Go. Claramente, la búsqueda no puede ser exhaustiva. La búsqueda se realiza mediante un árbol de búsqueda de Markov; esto es, un árbol de búsqueda aleatorio. La búsqueda se realiza, en esencia, de manera *aleatoria*. Pero en esta búsqueda interviene el algoritmo del descenso del gradiente (*gradient descent*). Algoritmo que, como se ha dicho, lo que permite básicamente es ir minimizando de manera iterativa (recursiva) la función de error, o el coste del modelo, siguiendo rutas aleatorias. Lo que se hace con este algoritmo es básicamente moverse aleatoriamente en el espacio de búsqueda (superficie de la función de coste) intentando encontrar mínimos locales. Esto, haciendo variaciones en los valores de los parámetros considerados en dicho modelo hasta ir encontrando mínimos locales.

En *AlphaFold*, los parámetros considerados son las distancias que separan a los aminoácidos que constituyen a la proteína y los ángulos de torsión de los enlaces químicos de dichos constituyentes (aminoácidos).

“Nuestro equipo se centró específicamente en el difícil problema de modelar formas específicas de proteínas desde cero, es decir, sin utilizar proteínas previamente resueltas como plantillas. Alcanzamos un alto grado de precisión al predecir las propiedades físicas de una estructura proteica, y luego usamos dos métodos distintos para construir predicciones de estructuras proteicas completas.

Ambos métodos se basan en redes neuronales profundas que están entrenadas para predecir las propiedades de la proteína a partir de su secuencia genética. Las propiedades que predicen nuestras redes son: (a) las distancias entre pares de aminoácidos y (b) los ángulos entre los enlaces químicos que conectan esos aminoácidos. El primer desarrollo es un avance en las técnicas de uso común que estiman si los pares de aminoácidos están cerca uno del otro.

Entrenamos una red neuronal por separado para predecir una distribución de distancias entre cada par de residuos en una proteína. Estas probabilidades se combinaron en una puntuación que estima cuán precisa es una estructura de proteína propuesta. También entrenamos una red neuronal separada que usa todas las distancias en conjunto para estimar qué tan cerca está la estructura propuesta de la respuesta correcta.

Usando estas funciones de puntuación, estuvimos en capacidad de buscar en el paisaje de proteínas para encontrar estructuras que coincidieran con nuestras predicciones. Nuestro primer método se basó en técnicas comúnmente utilizadas en biología estructural, y se reemplazó repetidamente piezas de una estructura proteica con nuevos fragmentos. Entrenamos una red neuronal generativa para inventar nuevos fragmentos, que se utilizaron para mejorar continuamente la puntuación de la estructura proteica propuesta.

El segundo método optimizó los puntajes a través del descenso del gradiente —una técnica matemática comúnmente utilizada en el aprendizaje automático para realizar pequeñas y crecientes mejoras—, lo que resultó en estructuras altamente precisas. Esta técnica más bien se aplicó a cadenas proteicas completas, en lugar de a piezas que deben plegarse por separado antes de ensamblarse, reduciendo la complejidad del proceso de predicción.”<sup>49</sup> (Evans et al., 2018)

---

<sup>49</sup> *Our team focused specifically on the hard problem of modelling target shapes from scratch, without using previously solved proteins as templates. We achieved a high degree of accuracy when predicting the physical properties of a protein structure, and then used two distinct methods to construct predictions of full protein structures.*

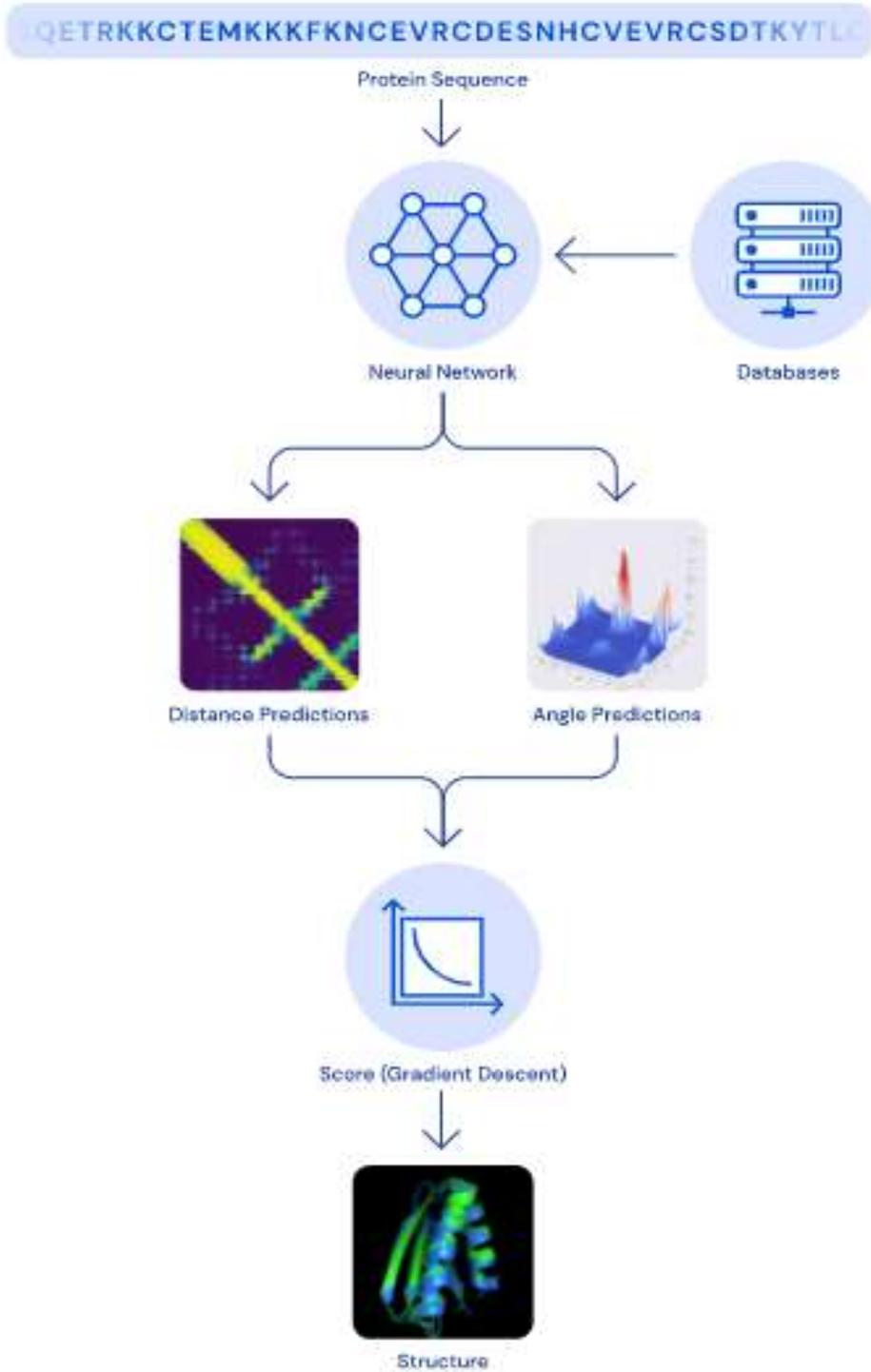
*Both of these methods relied on deep neural networks that are trained to predict properties of the protein from its genetic sequence. The properties our networks predict are: (a) the distances between pairs of amino acids and (b) the angles between chemical bonds that connect those amino acids. The first development is an advance on commonly used techniques that estimate whether pairs of amino acids are near each other.*

---

*We trained a neural network to predict a separate distribution of distances between every pair of residues in a protein. These probabilities were then combined into a score that estimates how accurate a proposed protein structure is. We also trained a separate neural network that uses all distances in aggregate to estimate how close the proposed structure is to the right answer.*

*Using these scoring functions, we were able to search the protein landscape to find structures that matched our predictions. Our first method built on techniques commonly used in structural biology, and repeatedly replaced pieces of a protein structure with new protein fragments. We trained a generative neural network to invent new fragments, which were used to continually improve the score of the proposed protein structure.*

*The second method optimized scores through [gradient descent](#)—a mathematical technique commonly used in machine learning for making small, incremental improvements—which resulted in highly accurate structures. This technique was applied to entire protein chains rather than to pieces that must be folded separately before being assembled, reducing the complexity of the prediction process. (Evans et al., 2018)*



Gráfica tomada de la página de *DeepMind*  
<https://deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery>

Esta gráfica ilustra, en términos generales, el procedimiento llevado a cabo por *AlphaFold* en el CASP-13. En primer lugar, los investigadores de *DeepMind* se valen del extenso dataset del que se dispone, desde hace ya un par de décadas, sobre la secuencia genómica de varias especies. Es decir, se valen básicamente del extenso número de proteínas de las que se conoce su secuencia de aminoácidos. Se estima que son aproximadamente 100 millones de proteínas<sup>50</sup>, de las cuales 20000 son de *Homo sapiens*. Sobre ese extenso número de secuencias se valen de una técnica que se conoce como ‘*Multiple sequence alignment*’ (MSA) (Senior et al. 2019), con la que se busca principalmente alinear secuencias de aminoácidos constituyentes de proteínas emparentadas o ligadas evolutivamente. Es decir, se busca alinear secuencias homólogas.

La idea con este alineamiento es identificar patrones entre dichas secuencias. O quizás, más concretamente, correlaciones. Correlaciones que de alguna manera ayuden a estimar la distancia a la que se encuentran dos aminoácidos entre sí. Por ejemplo, si al alinear varias secuencias de aminoácidos, constituyentes correspondientemente de varias proteínas ligadas evolutivamente de distintas especies (secuencias homólogas), se logra identificar, bajo algún patrón de alineamiento, una fuerte correlación entre dos aminoácidos, esto representaría un fuerte indicio de que esos dos aminoácidos se hallan a una distancia cercana el uno del otro.

Como situación meramente ilustrativa, sin pretensión de total exactitud, considérese la secuencia de aminoácidos que constituye la hemoglobina en *Homo sapiens*:

---

<sup>50</sup> El catálogo de estas proteínas se encuentra en: <https://www.uniprot.org/help/uniref>

MVLSPADKTNVKAAWGKVGAAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNA  
VAHVDDMPNALSALSADLHAHKLVRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTISK  
YR<sup>51</sup>

A continuación, tómese por ejemplo la secuencia de aminoácidos que constituye la hemoglobina en ratones caseros (*Mus musculus*):

MVLSGEDKSNIKAAWGKIGGHGAEYGAEALERMFASFPTTKTYFPHFDVSHGSAQVKGHGKKVADALASA  
AGHLDDLPGALSALSADLHAHKLVRVDPVNFKLLSHCLLVTLASHHPADFTPAVHASLDKFLASVSTVLTISK  
YR<sup>52</sup>

A simple vista, puede notarse que son secuencias bastante similares. Si se alinean, por ejemplo, en decenas, quizás puedan observarse aún más fácilmente algunas correlaciones entre ellos.

H.s: MVLSPADKTNVKAAWGKVGAAHAGEYGAEALERMF<sup>1</sup>LSFPTTKTYFPHFDLS<sup>2</sup>...  
M.m: MVLSGEDKSNIKAAWGKIGGHGAEYGAEALERMF<sup>1</sup>ASFPTTKTYFPHFDV<sup>2</sup>SHGSAQVKGHGKKVADALASA  
AGHLDDLPGALSALSADLHAHKLVRVDPVNFKLLSHCLLVTLASHHPADFTPAVHASLDKFLASVSTVLTISK  
YR<sup>52</sup>

En caso tal entonces que lo que se esté intentado identificar sean correlaciones entre aminoácidos en la hemoglobina en *Homo sapiens*, podríamos valernos de la información que nos brindan las secuencias homólogas tomadas en consideración, que, en este caso, sobre

---

<sup>51</sup> Secuencia tomada de:  
[https://www.ebi.ac.uk/seqdb/confluence/display/JDSAT/Multiple+Sequence+Alignment+Tool+Input+Examples#MultipleSequenceAlignmentToolInputExamples-FASTAFASTA\\_p](https://www.ebi.ac.uk/seqdb/confluence/display/JDSAT/Multiple+Sequence+Alignment+Tool+Input+Examples#MultipleSequenceAlignmentToolInputExamples-FASTAFASTA_p)

<sup>52</sup> Secuencia tomada de:  
[https://www.ebi.ac.uk/seqdb/confluence/display/JDSAT/Multiple+Sequence+Alignment+Tool+Input+Examples#MultipleSequenceAlignmentToolInputExamples-FASTAFASTA\\_p](https://www.ebi.ac.uk/seqdb/confluence/display/JDSAT/Multiple+Sequence+Alignment+Tool+Input+Examples#MultipleSequenceAlignmentToolInputExamples-FASTAFASTA_p)

simplificando un poco el asunto, sólo se ha tomado una secuencia, la de *Mus musculus*. Así, por ejemplo, podría considerarse en este caso que quizás los aminoácidos ‘prolina’ (P) y ‘valina’ (V), resaltados con color amarillo, guardan una estrecha relación de cercanía al plegarse; igualmente, la ‘adenina’ (A) y ‘valina’ (V), resaltados con color verde; la ‘treonina’ (T) y la ‘adenina’ (A), resaltados con color morado; etc.

De modo que, inicialmente, la red se entrenaría para generar una imagen que represente la distribución de distancias y ángulos de torsión entre los aminoácidos constituyentes de la proteína de la que se estaría intentando predecir su estructura tridimensional específica.

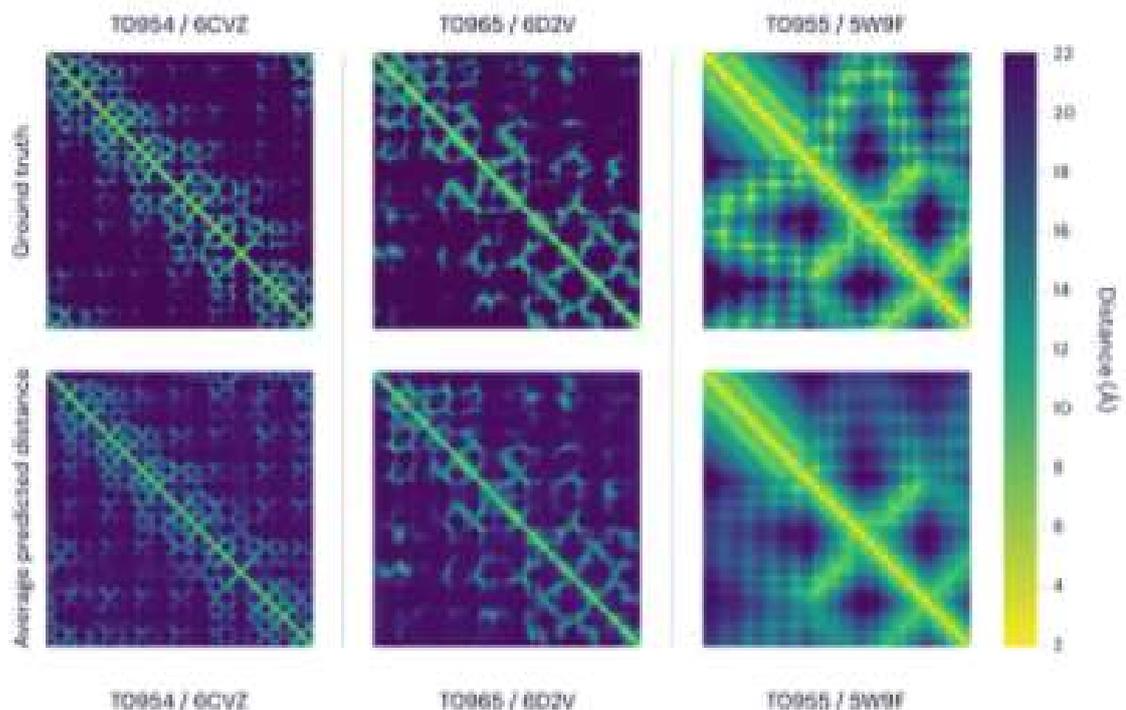


Figura tomada de la página de *DeepMind*:  
<https://deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery>

Es decir, se entrenaría para generar algo como la imagen de arriba. Este *output* representaría el *input* de otra red conectada a la anterior; que se entrenaría, mediante el algoritmo del descenso del gradiente, para generar la estructura tridimensional específica. Siendo el parámetro principal en la segunda red el ángulo de torsión entre los aminoácidos.

A continuación, una imagen un poco más detallada de lo que hace *AlphaFold*:

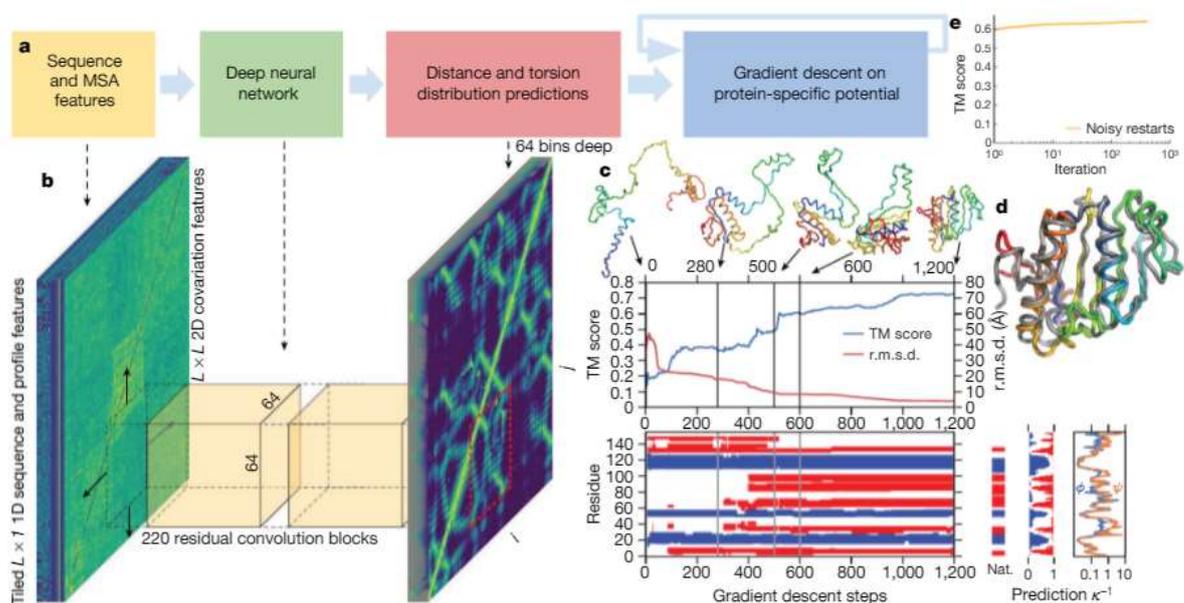


Figura tomada de (Senior et al. 2019, p. 708)

Se intenta mostrar con lo anterior que un punto clave en la resolución de este problema, que llevaba algo más de cincuenta años de planteado, consistió en la manipulación y exploración de las consecuencias de dicha manipulación, vía modelo y simulación computacional, del parámetro estimado relevante en la dinámica del sistema. Esta exploración, dada la enorme cantidad de combinaciones de valores involucrados en el

fenómeno emergente a explicar, no puede darse de manera exhaustiva; sino más bien de manera aleatoria. A partir del conocimiento de las propiedades fisicoquímicas de los aminoácidos constituyentes de una proteína particular, resulta impensable deducir la estructura o configuración específica que la hace funcional. Se requiere, sin lugar a duda, de otras vías alternativas.

Las redes neuronales artificiales son una herramienta que permiten dar cuenta de la interacción no lineal de muchísimos elementos. Esto representa una estrategia con la que se podría dar un tipo de aproximación no analítica de fenómenos emergentes.

Se habla de *aproximaciones no analíticas* en el sentido de que en ellas no se parte estrictamente de un análisis de los contenidos de conocimiento de los elementos involucrados en el fenómeno -emergente- a explicar para poder dar cuenta de este. Se parte más bien de una suerte de manipulación o intervención, por medio de simulaciones computacionales, de dichos elementos involucrados, hasta obtener una correlación entre el fenómeno a explicar y los valores de ciertos parámetros considerados relevantes en la dinámica del sistema. El factor clave en este tipo de aproximaciones está en su carácter *manipulacionista*, que se da mediante las simulaciones computacionales, posibilitadas por el altísimo poder de cálculo y procesamiento de información que ofrecen los ordenadores de los que se disponen hoy día; principalmente, de aquellos que disponen de GPUs (*Graphics Processing Unit*). En estas simulaciones computacionales las ecuaciones diferenciales parciales resultan ser herramientas básicas. Pero estas ecuaciones no suelen resolverse<sup>53</sup> de manera analítica; se

---

<sup>53</sup> Se estima que la noción de “solución” en el terreno de las matemáticas adquiere nuevas connotaciones a partir de estos métodos. La discusión que surge básicamente es si efectivamente se “resuelven” estos sistemas de ecuaciones. Sobre esto hablaremos un poco más en el siguiente capítulo.

“resuelven”, frecuentemente, por métodos de aproximación numérica, en los que se requiere mucho poder de cálculo computacional.

El fenómeno de la complejidad podría verse como invitándonos a replantear el tipo de herramientas con las que intentamos dar orden dentro de nuestra visión manifiesta de la naturaleza. La complejidad irrumpe con nuestra manera tradicional de organizar lo que podríamos llamar fenómenos de nuestra visión manifiesta de la naturaleza. *La no linealidad*, el *procesamiento en paralelo de información*, la *emergencia*, la *autoorganización*, la *retroalimentación*, la *adaptación*, todas ellas características sobresalientes de la complejidad, han hecho que se examine, quizás desde una nueva perspectiva, las herramientas con las que intentar abordar la naturaleza. Esta examinación nos ha puesto de manifiesto las maneras en las que hemos intentado organizar y estructurar nuestro conocimiento. La complejidad, entre otros fenómenos, nos ha hecho reflexionar, sobre todo, acerca de nuestras consideraciones previas de cómo es que conocemos. El estudio de los sistemas complejos nos ha puesto de manifiesto que nuestro conocimiento es quizás la mayor expresión de complejidad, al ser resultado de la actividad del que sea quizás el ejemplo *par excellence* de sistema complejo. A saber, nuestro cerebro.

Es importante resaltar que la limitación de un acercamiento expresamente analítico al fenómeno de la emergencia, y en general al fenómeno de la complejidad, va más allá de la dificultad de encontrar esquemas generales aplicables a estos fenómenos. En principio, mediante el sistema de ecuaciones diferenciales parciales para representar la función de onda de Schrödinger, se podría determinar la configuración molecular que pretendamos; haciendo alusión al problema del plegamiento de las proteínas. Sin embargo, con estas ecuaciones diferenciales parciales, a partir de más diez partículas, no es sólo ya que una solución analítica resulta prácticamente imposible, sino que, además, una solución por aproximación numérica

resulta en extremo difícil, aun con el poder de cálculo que presentan hoy día los computadores más poderosos de los que disponemos. Piénsese que sólo estamos considerando algo más de diez partículas. La situación para el *protein folding problem* es extremadamente inmanejable desde la función de onda de Schrödinger, ya que en una proteína intervienen cientos de miles de partículas.

Sin embargo, como se ha podido ir observando con los ejemplos presentados de *Acetubularia* y *AlphaFold*, donde a partir de la identificación de ciertos parámetros, y de la exploración, posibilitadas por simulaciones computacionales, de las consecuencias de la manipulación de un amplio conjunto de valores paramétricos ha resultado una estrategia conducente a resultados positivos. Siendo el caso de *AlphaFold* el más significativo. De hecho, sumamente significativo, como lo intentan atestiguar, por ejemplo, las palabras de los siguientes investigadores

Arthur D. Levinson (PhD., fundador y director ejecutivo de Calico, expresidente y director ejecutivo de Genentech):

“AlphaFold es un avance único en una generación, que predice estructuras de proteínas con una velocidad y precisión increíbles. Este salto adelante demuestra cómo los métodos computacionales están preparados para transformar la investigación en biología y son muy prometedores para acelerar el proceso de descubrimiento de fármacos.”<sup>54</sup> (Jumper et al., 2020)

---

<sup>54</sup> “AlphaFold is a once in a generation advance, predicting protein structures with incredible speed and precision. This leap forward demonstrates how computational methods are poised to transform research in biology and hold much promise for accelerating the drug discovery process.

**ARTHUR D. LEVINSON**

PHD, FOUNDER & CEO CALICO, FORMER CHAIRMAN & CEO, GENENTECH” (Jumper et al., 2020)

John Moult (Cofundador y presidente del CASP. Universidad de Maryland):

“Habíamos estado atascados en este problema durante casi 50 años. Ver a DeepMind producir una solución después de haber trabajado personalmente en este problema durante tanto tiempo, y después de tantas pausas y reanudaciones, preguntándonos si alguna vez llegaríamos a una solución, es un algo muy especial.”<sup>55</sup> (Jumper et al., 2020)

Profesor Venki Ramakrishnan (Premio Nobel de química y presidente de la Royal Society):

“Este trabajo computacional representa un avance sorprendente en el problema del plegamiento de proteínas, un gran desafío de 50 años en biología. Han transcurrido décadas antes de que personas en el campo lo hubieran predicho. Será emocionante ver las muchas formas en que cambiará fundamentalmente la investigación biológica.”<sup>56</sup> (Jumper et al., 2020)

Debido a la importancia de la identificación de ciertos parámetros en este tipo de estrategias de abordaje, sobre todo, en sistemas complejos, podríamos llamar a las aproximaciones que acá se desprenden como «*modelo-paramétricas*»; aproximaciones que dependen fuertemente de simulación computacional.

---

<sup>55</sup> “*We have been stuck on this one problem – how do proteins fold up – for nearly 50 years. To see DeepMind produce a solution for this, having worked personally on this problem for so long and after so many stops and starts, wondering if we’d ever get there, is a very special moment.*”

**PROFESSOR JOHN MOULT**

CO-FOUNDER AND CHAIR OF CASP, UNIVERSITY OF MARYLAND.” (Jumper et al., 2020).

<sup>56</sup> “*This computational work represents a stunning advance on the protein-folding problem, a 50-year-old grand challenge in biology. It has occurred decades before many people in the field would have predicted. It will be exciting to see the many ways in which it will fundamentally change biological research.*”

**PROFESSOR VENKI RAMAKRISHNAN**

NOBEL LAUREATE AND PRESIDENT OF THE ROYAL SOCIETY.” (Jumper et al., 2020)

#### III.4 ¿Se está produciendo un giro en ciencias hacia la predicción?

Con el ejemplo de *AlphaFold* se ha intentado evidenciar la gran relevancia que cobran las redes neuronales artificiales de aprendizaje profundo para un abordaje exitoso de la emergencia/complejidad, y de cómo en este abordaje lo que hemos llamado manipulacionismo se hace presente. Pero, además, con este ejemplo se intenta mostrar un caso muy dicente a favor de la idea, que vienen sosteniendo algunos, como M. Chirimuuta (2018) y Daniel Weiskopf (2022), de que recientemente se viene generando en ciencias una suerte de giro epistémico hacia la predicción; siendo en particular el de la neurociencia el caso estudiado por estos dos autores. Se traen a colación, sin embargo, por la importancia que les atribuyen a modelos computacionales en dicho giro. En el caso de Chirimuuta, abiertamente a las redes neuronales artificiales. “El giro predictivo se puede definir en términos de un compromiso con la autonomía de la predicción frente a otros principios científicos organizadores”<sup>57</sup> (Weiskopf, 2022, p. 2)

Teniendo en cuenta que en neurociencias se trata principalmente al que se estima es el sistema complejo por excelencia -a saber: nuestro cerebro-, podría pensarse que, en cierta medida, este giro viene motivado por los estudios recientes en sistemas complejos, donde la cantidad de información que se dispone para ser procesada es sumamente abrumadora.

“La neurociencia está experimentando una gran revolución de datos, donde métodos de alto rendimiento generan terabytes de registros neuronales, y algoritmos de aprendizaje automático se encuentran trabajando en la búsqueda de algún significado

---

<sup>57</sup> *The predictive turn can be defined in terms of a commitment to the autonomy of prediction vis-à-vis other organizing scientific principles.* (Weiskopf, 2022, p. 2)

y patrón entre el sinfín de disparos neuronales registrados simultáneamente.”

<sup>58</sup>(Chirimuuta, 2018, p. 1)

El estudio de sistemas complejos, además de motivar el uso extensivo de herramientas computacionales, igualmente parece motivar una valoración muy positiva hacia la predicción, aun representando esto en no pocas ocasiones la pérdida parcial de otros valores epistémicos, como la transparencia.

“El aspecto opaco del procesamiento detallado interno que llevan a cabo las redes neuronales artificiales es algo discutido por expertos dentro de la neurociencia. Por ejemplo, Omri Barak (2017:5) señala que “el aprendizaje automático nos proporciona niveles de rendimiento cada vez mayores, acompañados de un aumento paralelo de la opacidad”.

Sin embargo, sería un error decir que las redes neuronales artificiales son literalmente cajas negras, ya que los investigadores que desarrollan estos modelos conocen muchas características de su arquitectura interna y su funcionamiento. Al mismo tiempo, la forma exacta en que una red llega a las predicciones o clasificaciones suele ser bastante opaca para sus creadores, de ahí las preocupaciones. Como los neurocientíficos teóricos Gao y Ganguli (2015: 151) describen las cosas,

Cada una de estas redes [neuronales artificiales] puede resolver un problema computacional complejo. Además, conocemos la conectividad completa de la red, la dinámica de cada neurona, la regla de plasticidad utilizada para entrenar la red y, de hecho, toda la experiencia de desarrollo de la red... Sin embargo, aún nos elude una

---

<sup>58</sup> *Neuroscience is undergoing a big-data revolution, where high throughput methods generate terabytes of neural recordings, and machine learning algorithms are at work searching for meaning and pattern amongst the endless numbers of simultaneously recorded spikes and traces.* (Chirimuuta, 2018, p. 1)

comprensión significativa de cómo funcionan estas redes y cuál sería un punto de referencia adecuado para tal comprensión.”<sup>59</sup> (Chirimuuta, 2018, pp. 3-4).

Quiere decir entonces que, particularmente el uso extensivo de redes neuronales artificiales para el abordaje de sistemas complejos motiva, dentro de la discusión filosófica sobre las simulaciones computacionales, ciertas consideraciones que podrían considerarse como novedosas en cuanto a la opacidad epistémica que subyace en sus despliegues exitosos en materia predictiva. Sobre esto hablaremos más extensamente en el siguiente capítulo.

---

<sup>59</sup> *The black box flavour of artificial neural networks is something discussed by experts within neuroscience. For example Omri Barak (2017:5) points out that “machine learning provides us with ever increasing levels of performance, accompanied by a parallel rise in opaqueness”. However, it would be wrong to say that ANN’s are literally black boxes because so many features of their internal architecture and workings are known to the model builders. At the same time, the exact way that a network arrives at predictions or classifications is often quite opaque to its makers, hence the concerns. As theoretical neuroscientists Gao and Ganguli (2015:151) describe matters, Each of these [artificial neural] networks can solve a complex computational problem. Moreover, we know the full network connectivity, the dynamics of every single neuron, the plasticity rule used to train the network, and indeed the entire developmental experience of the network.... Yet a meaningful understanding of how these networks work still eludes us, as well as what a suitable benchmark for such understanding would be.* (Chirimuuta, 2018, pp. 3-4)

## CAPÍTULO IV

### ALGUNOS RETOS EPISTEMOLÓGICOS QUE REPRESENTA EL ESTUDIO DE SISTEMAS COMPLEJOS Y SU RELACIÓN CON LOS MODELOS Y SIMULACIONES COMPUTACIONALES

#### *¿Cuál es la novedad epistémica de las simulaciones computacionales?*

##### Abstract

En este capítulo defenderemos la idea, contra ciertas posiciones canónicas en la literatura filosófica, de que la modelización mediante redes neuronales artificiales de sistemas complejos presenta novedades epistemológicas específicas. Estas novedades están vinculadas con la opacidad epistémica y el papel de la manipulación y la visualización en este tipo de enfoques.

En este capítulo se intentará extender las consideraciones acerca de las dificultades epistemológicas que representa el estudio de sistemas complejos, fijando ahora mayormente la atención en la posible disrupción que se genera con el empleo extensivo de modelos y simulaciones computacionales en el abordaje de esta clase de sistemas, principalmente en materia metodológica y justificativa de la investigación científica. La *opacidad epistémica*, la *experimentación iterada* y la posibilidad de *manipulación* y *visualización* que nos significa en ocasiones el empleo de modelos y simulaciones computacionales, abren escenarios, que quizás merezcan ser considerados, pese a las importantes objeciones realizadas al respecto por Roman Frigg y Julian Reiss (2009), como novedosos en el terreno de la filosofía de la ciencia.

Comenzaremos pues este capítulo planteando lo que se suele entender por “simulación computacional”. Para ello, nos valdremos de la clasificación hecha por Eric Winsberg, quien

es uno de los representantes más significativos en la literatura filosófica sobre las simulaciones computacionales.

#### IV.1 *¿Qué se suele entender por simulación computacional?*

##### VI.1.1 *Simulación computacional en un sentido estrecho*

En un sentido estrecho o, quizás mejor, estricto, se suele entender por simulación computacional el proceso de ejecución de un programa o algoritmo en un ordenador; donde el programa o algoritmo que se ejecuta representa, en el lenguaje de programación empleado, un modelo dinámico del sistema objeto de estudio. Este modelo viene dado regularmente por un sistema de ecuaciones diferenciales parciales, cuya intratabilidad analítica conlleva a que se resuelva frecuentemente más bien por métodos de aproximación numérica.

Pero, más allá de las características del modelo subyacente para el abordaje del sistema objeto de estudio y de las estrategias que se desplieguen para obtener información relevante del sistema a partir de la implementación de dicho modelo, bajo esta interpretación de las simulaciones computacionales se hace alusión concretamente a lo que sucede cuando se pone a correr el programa; se hace alusión al proceso de ejecución de las líneas de código que constituyen el programa o algoritmo en el que se encuentra representado el modelo con el que se intenta abordar la dinámica del sistema en consideración. Esto hace que bajo esta interpretación el lenguaje de programación empleado, el tipo de compilador, el tipo de ordenador en el que se corre el programa, jueguen ciertamente un papel no menor, pues la ejecución de un programa o algoritmo es sumamente sensitiva a variaciones en estos factores.

En palabras de Eric Winsberg (2013):

“Cuando se habla de “simulación computacional” en el sentido más estricto, se está hablando de la implementación particular de un algoritmo en una computadora digital específica, escrito en un lenguaje de programación específico, utilizando un compilador específico, etc. [Donde habría que tomarlos particularmente en cuenta, pues] hay casos en los que se obtienen resultados diferentes debido a variaciones en cualquiera de estos ítems.”<sup>60</sup>

Para ser un poco más precisos se dirá que por simulación computacional en el sentido más estricto, se entiende la ejecución de un programa, con todas las especificidades recién mencionadas, que va arrojando un “panorama” numérico de la evolución de los estados del sistema, según la conceptualización de su dinámica plasmada en el modelo (matemático) subyacente (Winsberg, 2013, p. 2). Este panorama numérico, o “secuencia de valores para las variables del modelo [que va arrojando el programa en ejecución], puede guardarse como una larga colección de ‘información’ y visualizarse en una pantalla de computador mediante algunos métodos”.<sup>61</sup> Mediante esta posibilidad de visualización se estaría logrando observar una representación de la dinámica propia del sistema objeto de estudio.

#### IV.1.2 *Simulación computacional en un sentido amplio*

Cuando se habla de simulación computacional puede igualmente no sólo hacerse alusión específicamente al proceso de ejecución de un programa, sino que puede además tomarse en consideración un proceso más amplio en la elaboración del programa que se va a ejecutar. Si

---

<sup>60</sup> Finally, when speaking of “a computer simulation” in the narrowest sense, we should be speaking of a particular implementation of the algorithm on a particular digital computer, written in a particular language, using a particular compiler, etc. There are cases in which different results can be obtained as a result of variations in any of these particulars. (Winsberg, 2013, p. 2)

<sup>61</sup> This sequence of values for the model variables can be saved as a large collection of “data” and is often viewed on a computer screen using methods of visualization. (Winsberg, 2013, p. 2)

bien desde esta óptica de las simulaciones computacionales se siguen considerando -como sucede en la visión estrecha- todos los detalles involucrados propiamente en el proceso de ejecución, como el lenguaje de programación empleado, el tiempo de ejecución, las características de procesamiento del ordenador en el que es ejecutado el programa, el compilador, etc., también se centra la atención en los aspectos creativos y justificativos que subyacen a la ejecución de un programa como representación de un modelo.

Es decir, podría considerarse, como lo hace Johannes Lenhard (2016, p, 3), que la visión estrecha es parte constituyente (*part-and parcel*) de la visión en un sentido más amplio de las simulaciones computacionales. Y es dentro de esta visión un poco más amplia que, tanto Lenhard, como Winsberg, Humphreys, Morrison, entre otros autores representativos de la discusión filosófica acerca de las simulaciones computacionales, consideran es en donde surgen mayormente los aspectos relevantes y significativos de dicha discusión.

“En un sentido más general, podemos pensar en las simulaciones computacionales como un método integral para el estudio de ciertos sistemas. En este sentido más amplio del término se hace alusión a todo un proceso. Este proceso incluye elegir un modelo; encontrar una manera de implementar ese modelo en una forma que pueda ejecutarse en una computadora; calcular los datos de salida del algoritmo; y visualizar y estudiar dichos datos resultantes. El método incluye todo este proceso, utilizado para hacer inferencias sobre el sistema que se intenta modelar, así como los procedimientos utilizados para aprobar esas inferencias. [...] Cuando desde la filosofía de la ciencia se escribe acerca de las simulaciones computacionales y se pretenden plantear cuáles serían las características epistemológicas y metodológicas que estas poseen, se hace

usualmente tomando en consideración este segundo sentido en que pueden entenderse.<sup>62</sup> (Winsberg, 2013, p. 3)

#### IV.1.3 *Un punto de vista alternativo sobre las simulaciones computacionales*

Existe una tercera acepción que resulta por parte de algunos, como por ejemplo Hartmann (1996), que realizan un análisis de los términos involucrados por separado y luego tratan de compaginarlos. Es decir, realizan un análisis del término “simulación” por separado y luego ese análisis tratan de particularizarlo en el terreno de lo computacional. Por simulación se estaría entendiendo básicamente desde esta perspectiva una relación de similitud entre las dinámicas de dos sistemas (Winsberg, 2013, p. 3). Por ejemplo, la dinámica de activación en las neuronas artificiales simularía, desde esta perspectiva, la dinámica de las neuronas biológicas. Suponiendo esto, lo que se buscaría básicamente sería extrapolar algunos resultados en el estudio dinámico de uno de los sistemas hacia el otro sistema en consideración. Y la extrapolación se haría a partir de los resultados obtenidos del empleo de ordenadores en el abordaje dinámico de uno de los sistemas. Siendo el que se modela quizás uno menos complejo que el otro.

---

<sup>62</sup> *More broadly, we can think of computer simulation as a comprehensive method for studying systems. In this broader sense of the term, it refers to an entire process. This process includes choosing a model; finding a way of implementing that model in a form that can be run on a computer; calculating the output of the algorithm; and visualizing and studying the resultant data. The method includes this entire process—used to make inferences about the target system that one tries to model—as well as the procedures used to sanction those inferences. [...] When philosophers of science write about computer simulation, and make claims about what epistemological or methodological properties “computer simulations” have, they usually mean the term to be understood in this broad sense of a computer simulation study. (Winsberg, 2013, p. 3)*

## IV.2 Métodos de simulación

### IV.2.1 Método de Monte Carlo

La gran importancia que empiezan a adquirir las simulaciones computacionales en la investigación científica podríamos referenciarla durante la segunda guerra mundial, y en el período inmediatamente posterior, principalmente con los estudios llevados a cabo en meteorología -con Edward Lorenz y su *Royal McBee LGP-30* como uno de sus abanderados- y en física nuclear, siendo el matemático John von Neumann quizás uno de los que más resalte por su papel en el Proyecto Manhattan, ya que una parte significativa de las investigaciones que se llevaban a cabo en física nuclear estaban relacionadas al desarrollo de armas nucleares y el matemático húngaro en el marco de dicho proyecto hizo aportes determinantes.

Fue precisamente en el contexto de estos aportes donde podría considerarse el método de simulación de Monte Carlo tuvo inicialmente lugar en el terreno de la investigación científica. Sobre el desarrollo de este método y su discusión filosófica en el terreno de la investigación científica, Peter Galison nos ofrece un análisis detallado en un interesante artículo de 1996 titulado '*Computer Simulations and the Trading Zone*'. En él nos planteaba, entre otras cosas, que las simulaciones computacionales, centrando su atención principalmente en el método de Monte Carlo, irrumpen con las tres tradiciones de investigación que identifica -teórica, experimental e instrumental- (Lenhard, 2016, p.), instaurando así un nuevo modo de producción de conocimiento científico (Galison, 1996, p. 119)

“Fundadas en estadística, teoría de juegos, teoría de muestreo y codificación computacional, estas simulaciones constituyen lo que he llamado “zona de

intercambio”, un terreno en el que actividades completamente diferentes pueden coordinarse *localmente*, pero no globalmente.”<sup>63</sup>

Desarrollo de armas termonucleares, mejoramiento de bombas atómicas, desarrollo de gases venenosos, predicción del clima, estudio de las interacciones entre piones y nucleones, algunos estudios sobre mecánica cuántica, entre otras actividades, por más que parezcan desconectadas entre sí, sin un marco carnapiano *-framework-* claro; sin leyes comunes, y obviamente, sin entidades ontológicas comunes, pueden englobarse aun así bajo un nuevo tipo de metodología fundada en el uso de computadores (Galison, 1996, pp. 118, 119). La aleatoriedad es pilar fundamental de esta nueva metodología, y el método de Monte Carlo, propuesto en gran medida por el matemático polaco Stanisław Ulam, constituye una muy interesante manera de trabajar con ella.

“[D]ebido a que tanto el método de Monte Carlo como la naturaleza siguen procesos estocásticos, el método podría ofrecer entonces un vistazo de la realidad previamente oculta para el matemático de mentalidad analítica.”<sup>64</sup> (Galison, 1996, p. 144)

Justamente lo que el método nos permite, en términos generales, es obtener una muy buena aproximación de una descripción de un fenómeno o de un sistema, regularmente constituido de muchísimas partes -como lo son los sistemas complejos-, mediante una enorme iteración de un proceso de aleatorización; de ahí su nombre. ‘Monte Carlo’ hace referencia acá a la meca, en el principado de Mónaco, de los juegos de azar, con la ruleta como un gran ícono.

---

<sup>63</sup> *Grounded in statistics, game theory, sampling, and computer coding, these simulations constituted what I have been calling a "trading zone, an arena in which radically different activities could be locally, but not globally, coordinated.*

<sup>64</sup> *“because both Monte Carlo and nature were stochastic, the method could offer a glimpse of reality previously hidden to the analytically minded mathematician”* (Galison, 1996, p. 144)

La idea básicamente consiste en generar muchísimas muestras estadísticas, pretendidamente<sup>65</sup> sobre la población, prácticamente inconmensurable, de elementos involucrados en el fenómeno a explicar, mediante muchísimas iteraciones de un proceso de aleatorización; pues lo que se busca es que dichas muestras no presenten, en lo posible, algún tipo de sesgo. La enorme iteración de un proceso de aleatorización sobre los elementos constituyentes de una población serviría, en principio, para la generación de una enorme cantidad de muestras no sesgadas, de modo tal que los estimativos estadísticos sobre esas muestras no reflejen dichos posibles sesgos. Es el enorme poder de cálculo que nos significan los ordenadores lo que posibilita la enorme cantidad de iteraciones necesarias para lograr mejores estimativos.

Un ejemplo interesante y típico de este método de simulación consiste en la estimación del valor de  $\pi$  a partir del lanzamiento aleatorio, repetidas veces, de una colección de objetos, -por ejemplo, canicas-, sobre una superficie que contiene un cuadrado -pongamos por caso, de lado ( $l$ ) 2- y un círculo -pongamos por caso, de radio ( $r$ ) 1- inscrito en dicho cuadrado. La estimación se lleva a cabo calculando, en cada iteración del proceso, la proporción de los objetos que caen dentro del círculo entre los que caen dentro del cuadrado. A medida que el número de iteraciones se hace cada vez mayor, la media estadística de estas proporciones se hace cada vez más próxima a  $\frac{\pi}{4}$ .

Es decir, la media de la distribución de probabilidad que siguen estas proporciones de objetos que caen dentro del círculo entre los que caen dentro del cuadrado, es muy próxima a la proporción de las áreas de las figuras geométricas consideradas; que en este caso sería:

---

<sup>65</sup> No necesariamente las muestras se obtienen a partir de los elementos constituyentes del propio fenómeno del que se quiera dar cuenta.

$$\frac{\text{Área}_{\text{círculo}}}{\text{Área}_{\text{cuadrado}}} = \frac{\pi r^2}{l^2} = \frac{\pi(1)^2}{2^2} = \frac{\pi}{4}$$

Acá, claramente, el teorema del límite central, visto en el capítulo II, cobra una gran importancia, pues sobre este número sumamente grande de muestras generadas mediante la iteración de un proceso de aleatorización, se le puede pedir al programa de computadora que se esté empleando que arroje la media de la distribución de probabilidad que sigue cada una de estas muchísimas muestras, y a partir de cada una de estas muchísimas medias estimar su distribución de probabilidad, que en este caso, por dicho teorema, sabemos que seguiría una distribución normal. Toda la información que se obtenga a partir de esta distribución de probabilidad podría estimarse, por lo que se conoce como la ley de los grandes números en estadística, como una muy buena aproximación a la información que se desee obtener de la población original, en principio intratable analíticamente por la enorme cantidad de elementos involucrados.

#### IV.2.2 Simulaciones basadas en ecuaciones

Las simulaciones basadas en ecuaciones son aquellas en las que se intenta ofrecer alguna predicción, descripción o aproximación explicativa de algún sistema o fenómeno a partir un esquema matemático pretendidamente abarcador de las regularidades que operan sobre dicho sistema o fenómeno. Es decir, el esquema matemático encapsula, mediante las relaciones entre las variables y parámetros estimados relevantes, un patrón sobre el sistema o fenómeno. Desde el esquema matemático se tiene, por decirlo de alguna manera, un panorama *general* del fenómeno o sistema que se intenta abordar.

Este esquema puede concebirse como uno basado en ítems particulares o discretos (*particle-based*), donde hay  $n$  de estos ítems en interacción y es un sistema de ecuaciones

diferenciales el que permite dar cuenta de esas interacciones, o como uno basado en un campo (*field-based*), donde el sistema de ecuaciones describe la evolución en el tiempo de un medio continuo o campo. (Winsberg, 2013, p. 5)

“Un ejemplo del primer tipo es una simulación de formación de galaxias, en la que la interacción gravitatoria entre una colección finita de cuerpos se discretiza en el tiempo y el espacio. Un ejemplo del segundo tipo es la simulación de un fluido, como un sistema meteorológico -como una tormenta severa. Aquí, el sistema se trata como un medio continuo, un fluido, y un campo que representa su distribución de las variables relevantes en el espacio se discretiza y luego se actualiza en intervalos discretos de tiempo.”<sup>66</sup> (Winsberg, 2013, p. 5)

#### IV.2.3 Simulaciones basadas en agentes

Las simulaciones computacionales basadas en agentes (*agent-based*) son aquellas en las que los elementos o ítems que se modelizan se rigen por pautas o reglas separadas. Cada ítem estimado en la simulación, opera, por decirlo de alguna manera, bajo ciertas reglas específicas. No hay un sistema de ecuaciones que gobierne las interacciones de los ítems considerados, sino más bien que el comportamiento de dichos ítems viene regido por reglas locales (ibid., 2013, p 5). Las redes neuronales artificiales son un ejemplo representativo de este tipo de simulación.

---

<sup>66</sup> *An example of the former is a simulation of galaxy formation, in which the gravitational interaction between a finite collection of discrete bodies is discretized in time and space. An example of the latter is the simulation of a fluid, such as a meteorological system like a severe storm. Here the system is treated as a continuous medium—a fluid—and a field representing its distribution of the relevant variables in space is discretized in space and then updated in discrete intervals of time.* (Winsberg, 2013, p. 5)

#### IV.2.4 Simulaciones multiescalas

Las simulaciones multiescala son aquellas en las que se combinan en el esquema matemático tanto el enfoque basado en partículas (*particle-based*) como el enfoque basado en un campo (*field-based*).

“Un buen ejemplo de esto sería un modelo que simule la dinámica de la materia granular al tratar el material como un campo sometido a tensión y deformación en un nivel de descripción relativamente bajo, pero que se acerque a regiones particulares del material donde se producen efectos importantes a pequeña escala, y modelar esas regiones más pequeñas con métodos de modelado relativamente más detallados.

Dichos métodos pueden basarse en dinámica molecular, mecánica cuántica, o ambas, cada una de las cuales es una descripción más detallada de la materia que la que se ofrece al tratar el material como un campo. Los métodos de simulación multiescala pueden dividirse en métodos de multiescala en serie y multiescala en paralelo. El método más tradicional es el modelado multiescala en serie. La idea aquí es elegir una región, simularla en el nivel inferior de descripción, resumir los resultados en un conjunto de parámetros admisibles por el modelo a un nivel de descripción superior y pasarlos a la parte del algoritmo que calcula en dicho nivel superior.

Los métodos multiescala seriales no son efectivos cuando las diferentes escalas están fuertemente acopladas. Cuando las diferentes escalas interactúan fuertemente para producir el comportamiento observado, lo que se requiere es un enfoque que simule cada región simultáneamente. Esto se llama modelado multiescala paralelo. El modelado multiescala paralelo es la base de un método de simulación casi omnipresente: el llamado modelado de "subcuadrícula" (*sub-grid*). El modelado de subcuadrícula se refiere a la representación de procesos físicos importantes a pequeña

escala que ocurren en escalas de longitud que no pueden resolverse adecuadamente en el tamaño de cuadrícula de una simulación particular. (Recuerde que muchas simulaciones discretizan ecuaciones continuas, por lo que tienen un "tamaño de cuadrícula" finito relativamente arbitrario). En el estudio de la turbulencia en fluidos, por ejemplo, una estrategia práctica común para el cálculo es tener en cuenta los vórtices de pequeña escala (o remolinos) que caen dentro de las celdas de la cuadrícula. Esto se hace agregando al movimiento a gran escala una viscosidad de remolino que caracteriza el transporte y la disipación de energía en el flujo a menor escala, o cualquier característica similar que ocurra a una escala demasiado pequeña para ser capturada por la red.”<sup>67</sup> (Winsberg, 2013, pp. 5 y 6)

#### IV.3 Algunos posibles aspectos epistemológicos novedosos en las simulaciones computacionales

Es una idea extendida desde hace poco más de veinte años que las simulaciones computacionales abren *nuevos* escenarios de discusión en el terreno de la filosofía de la

---

<sup>67</sup> *A good example of this would be a model that simulates the dynamics of bulk matter by treating the material as a field undergoing stress and strain at a relatively coarse level of description, but which zooms into particular regions of the material where important small scale effects are taking place, and models those smaller regions with relatively more fine-grained modeling methods. Such methods might rely on molecular dynamics, or quantum mechanics, or both—each of which is a more fine-grained description of matter than is offered by treating the material as a field. Multiscale simulation methods can be*

*further broken down into serial multiscale and parallel multiscale methods. The more traditional method is serial multi-scale modeling. The idea here is to choose a region, simulate it at the lower level of description, summarize the results into a set of parameters digestible by the higher level model, and pass them up to into the part of the algorithm calculating at the higher level.*

*Serial multiscale methods are not effective when the different scales are strongly coupled together. When the different scales interact strongly to produce the observed behavior, what is required is an approach that simulates each region simultaneously. This is called parallel multiscale modeling. Parallel multiscale modeling is the foundation of a nearly ubiquitous simulation method: so called “sub-grid” modeling. Sub-grid modeling refers to the representation of important small-scale physical processes that occur at length-scales that cannot be adequately resolved on the grid size of a particular simulation. (Remember that many simulations discretize continuous equations, so they have a relatively arbitrary finite “grid size.”) In the study of turbulence in fluids, for example, a common practical strategy for calculation is to account for the missing small-scale vortices (or eddies) that fall inside the grid cells. This is done by adding to the large-scale motion an eddy viscosity that characterizes the transport and dissipation of energy in the smaller-scale flow—or any such feature that occurs at too small a scale to be captured by the grid. (Winsberg, 2013, pp. 5 y 6)*

ciencia; acarrear nuevos aspectos epistemológicos que requieren discusión. Peter Galison (1996, pp. 119-120), por ejemplo, nos planteaba que los físicos y los ingenieros elevaron prontamente el método de Monte Carlo por encima del modesto estatus de un mero esquema de cálculo numérico. En esta elevación, sostiene Galison, motivada por su gran relevancia frente a lo que a su consideración era el problema más complejo en ese momento emprendido en la historia de la ciencia -a saber, el diseño de la primera bomba de hidrógeno-, se instauró una suerte de realidad alternativa; que, si bien marginal, preferible en varias ocasiones para llevar a cabo cierto tipo de “experimentación”.

Eric Winsberg, por su parte, en su artículo de 2001 (p. 443) se propuso discutir algunos de los aspectos epistemológicos distinguibles en el despliegue de simulaciones computacionales, y termina sosteniendo que dichos aspectos son novedosos dentro de la filosofía de la ciencia. Las distintas maneras en las que se intentan justificar las simulaciones computacionales no suelen aparecer contempladas en la literatura filosófica sobre justificación (p. 447).

Paul Humphreys (2004, p. 54), sin duda uno de los pensadores más destacado dentro de este debate, sostiene por ejemplo que el modo “como entendemos y evaluamos estos nuevos métodos [es decir, métodos por simulación computacional] es esencialmente diferente del que se emplea en la comprensión y evaluación de las teorías tradicionales.”<sup>68</sup>

En un importante artículo sobre este debate de las posibles novedades epistemológicas que representa el despliegue extensivo de simulaciones computacionales, Roman Frigg y Julian Reiss (2009) se encargaron de analizar precisamente las citas recién presentadas, junto con otras, a fin de poder ellos también responder a la pregunta de si representan aspectos

---

<sup>68</sup> “*how we understand and appraise these new methods [i.e., computer simulation methods] is essentially different from the understanding and evaluation of traditional theories.* (Humphreys, 2004, p. 54)

novedosos en el terreno de la filosofía de la ciencia. El artículo cobra gran importancia, entre otras cosas, debido a la argumentada cautela que nos invita a sostener frente a estas consideraciones extendidas dentro de la filosofía de la ciencia sobre las simulaciones computacionales. Básicamente, lo que plantean es que, en efecto, el despliegue de simulaciones computacionales acarrea aspectos novedosos. Sin duda que las simulaciones computacionales introducen nuevos y significativos aspectos que merecen discusión. Pero esos nuevos aspectos tienen que ver más, por ejemplo, con la psicología o con las propias matemáticas que con la filosofía de la ciencia (Frigg y Reiss, 2009, pp. 595 y 607). En todo caso, las discusiones que suscitan en el terreno de la filosofía pueden abordarse desde las que se dan bajo la temática de los modelos en ciencia.

Analizando los argumentos que han dado, por ejemplo, Galison, Winsberg, Humphreys, entre otros, a favor de la idea de la novedad epistemológica a partir de las simulaciones computacionales, Roman Frigg y Julian Reiss identificaron cuatro aspectos filosóficos fundamentales donde podrían considerarse dichas posibles novedades. Estos cuatro aspectos son:

*“Metafísico:* Las simulaciones crean una especie de mundo paralelo en el que los experimentos pueden llevarse a cabo en condiciones más favorables que en el "mundo real".

*Epistémico:* Las simulaciones exigen una nueva epistemología.

*Semántico:* Las simulaciones exigen un nuevo análisis de cómo los modelos/teorías se relacionan con los fenómenos concretos.

*Metodológico*: Simular es una actividad Sui Generis que se encuentra ‘entre’ la teorización y la experimentación.”<sup>69</sup> (Frigg y Reiss, 2009, p. 595)

Dentro del plano semántico tratan el aspecto de la *visualización*, que es uno de gran importancia para fines dispuestos en esta investigación doctoral. Contrario a lo que sostiene Humphreys, Frigg y Reiss plantean que la posible novedad que bajo este aspecto pueda significar las simulaciones computacionales quizás deba evaluarse más dentro del terreno de la psicología que desde la filosofía de la ciencia.

“*Visualización*. En las páginas 112-114 Humphreys parece sugerir que el hecho de que los resultados de las simulaciones computacionales se representan visualmente tiene un peso especial. Cuando se hace uso de las simulaciones, no solo se obtiene una gran cantidad de números, se obtienen además gráficos que representan formas geométricas, imágenes que se asemejan a las que experimentamos en otros contextos, etc. Seguramente esto es correcto, pero el problema nuevamente parece ser que las representaciones visuales no son particulares de las simulaciones. Utilizamos fotografías, y en los últimos años también vídeo secuencias para representar científicamente ciertos sistemas, incluso cuando no hay simulación involucrada (por ejemplo, imágenes por microscopio para procesos en neurobiología). Incluso en disciplinas ‘tradicionales’ las representaciones visuales son comunes: en nuestras clases de matemáticas nos pidieron que dibujáramos funciones, y en geometría diferencial a menudo usamos dibujos para visualizar los hallazgos. Lo mismo es cierto

---

<sup>69</sup> *Metaphysical*: Simulations create some kind of parallel world in which experiments can be conducted under more favourable conditions than in the ‘real world’.

*Epistemic*: Simulations demand a new epistemology.

*Semantic*: Simulations demand a new analysis of how models/theories relate to concrete phenomena.

*Methodological*: Simulating is a Sui Generis activity that lies ‘in between’ theorising and experimentation. (Frigg y Reiss, 2009, p. 595)

cuando usamos cualquier tipo de modelo 'icónico', como aquellos usados en química, o el modelo de doble hélice del ADN, etc. Además, nuevamente, en el mejor de los casos no está claro si este punto tiene alguna importancia epistémica. Parece ser más bien una cuestión de psicología antes que de epistemología el hecho de que preferimos lo visual a otras formas de representación.”<sup>70</sup> (Frigg y Reiss, 2009, p. 607)

Efectivamente, como lo plantean Frigg y Reiss, las representaciones visuales, más allá del nivel de resolución que posean, de la cantidad de información que en ellas se logre comprimir, de que puedan ser dinámicas, etc., no son propias y exclusivas de las simulaciones computacionales. La visualización en general es un aspecto epistemológico para considerar más allá del caso particular de las simulaciones computacionales. Se está de acuerdo además en que la preferencia por cierta disposición de la información -visual-, y quizás mayor capacidad de procesamiento bajo esa disposición que los humanos parecemos tener, es algo que atañe más a la psicología que a la filosofía de la ciencia.

Sin embargo, las imágenes que suelen ser arrojadas por métodos de simulación computacional, se considera aquí, están tan estrechamente relacionadas con un aspecto epistemológico que se ha intentado resaltar a lo largo de esta tesis, que parecería que deberían evaluarse junto con ese otro aspecto. Visualización y manipulación en simulación computacional suelen ir estrechamente de la mano. Y juntas abren un escenario -nuevo- no

---

<sup>70</sup> *Visualisation.* On pages 112–114 Humphreys seems to suggest that the fact that the outputs of simulations are represented visually bears special weight. When using simulations, we don't just get a pile of numbers, we get graphs representing geometrical shapes, images resembling those we experience in other contexts and so on. This is surely correct, but the problem again seems to be that visual representations are not particular to simulations. We use photographs, and in recent years also video sequences to represent scientifically certain target systems even where no simulation is involved (e.g., microscope footage for processes in neurobiology). And even in 'traditional' disciplines visual representations are common: in our mathematics classes we were asked to draw functions, and in differential geometry we often use drawings to visualise the findings. The same is true when we use any sort of 'iconic' model such as ball and stick models in chemistry, the double helix model of DNA and so on. Moreover, it is, again, at best unclear whether this point has any epistemic import. It seems to be a matter of psychology rather than epistemology that we prefer visual to other forms of representation. (Frigg y Reiss, 2009, p. 607)

contemplado por Frigg y Reiss en su famoso artículo ‘*The philosophy of simulation: hot new issues or same old state*’ (2009).

A lo largo de esta tesis, se ha intentado plantear que, en el abordaje de sistemas complejos, constituidos de cantidades abrumadoras de elementos, métodos de simulación computacionales han representado una poderosa herramienta, principalmente porque han permitido una exploración -no sin cierta opacidad- manipulacionista de *posibles* consecuencias en sus dinámicas a partir de ciertas alteraciones en algún o algunos de los parámetros estimados relevantes para intentar dar cuenta de dicha dinámica.

El asunto es que dicha exploración no puede ser exhaustiva, como bien se intentó evidenciar con el ejemplo del problema del plegamiento de proteínas. Incluso para los potentes ordenadores de los que se disponen hoy día, explorar exhaustivamente las posibles consecuencias en la dinámica de los sistemas complejos de alterar los valores de cierto parámetro o de ciertos parámetros estimados relevantes en el intento de dar cuenta de dicha dinámica bajo cierto modelo, resultaría una tarea de nunca acabar. Si bien se establecen ciertos criterios para dicha exploración<sup>71</sup>, aun dentro del dominio de valores paramétricos resultante la exploración exhaustiva resultaría algo improcedente. De modo que, una vez establecidos ciertos criterios para la exploración en cuestión, que trazan, por decirlo de alguna manera, ciertas “rutas de tránsito”, dentro de estas rutas el movimiento exploratorio se da más bien de manera aleatoria. Recuérdese el algoritmo de *stochastic gradient descent*, empleado junto al *backpropagation*, en el ejemplo mencionado del problema del plegamiento de las proteínas.

---

<sup>71</sup> Recuérdese por ejemplo los dos criterios establecidos para la exploración de las consecuencias en la dinámica de Acetabularia, con relación a la formación de parasoles, dadas ciertas alteraciones en los valores que podían tomar los parámetros contemplados en el modelo que desplegaron computacionalmente.

Lo que se está intentando plantear es que, en los métodos de simulación computacional, o más concretamente en las redes neuronales artificiales de aprendizaje profundo -que son las de nuestro mayor interés-, se presenta una situación epistemológica novedosa, y es la gran *opacidad* frente a lo que está ocurriendo en detalle en la búsqueda exploratoria de los valores de ciertos parámetros que se correlacionan con el fenómeno complejo/emergente a abordar. Con las redes neuronales artificiales se sabe qué es lo que está ocurriendo en términos generales en el despliegue del modelo subyacente que está siendo ejecutado, pero se desconocen completamente los “pasos” exactos que se van dando hasta encontrar dicha correlación. Se desconocen los “pasos” exactos que va dando la red en su entrenamiento. Recuérdese una vez más que, en la búsqueda exploratoria de los valores de ciertos parámetros que permitan establecer una correlación con el fenómeno que se intenta dar cuenta, la red neural sigue pasos aleatorios. Y, sobre todo, que no son pocos.

Es decir, no es sólo por ejemplo que en las redes neuronales artificiales de aprendizaje profundo se sigan pasos aleatorios en su ejecución o entrenamiento, sino que, además, no son pocos los pasos que se dan hasta llegar al resultado deseado. Por lo que, un intento de supervisión de estos, así no sean aleatorios, resulta también improcedente. Son muchísimas las iteraciones que se dan en el entrenamiento de una red como para intentar mínimamente un tipo de supervisión detallada de dicho entrenamiento. Sencillamente resultaría una tarea de nunca acabar.

Sin embargo, frente a esta fuerte opacidad epistémica, justamente las imágenes que se van generando en sus iteraciones juegan un papel muy importante, no contemplado infortunadamente en el interesante artículo de Frigg y Reiss. Llegados a este punto conviene diferenciar dos tipos de situaciones que se dan con las imágenes generadas a partir de simulación computacional, particularmente con aquellas generadas por redes neuronales

artificiales. En primer lugar, es importante señalar que la cantidad de imágenes que se suelen generar por métodos de simulación computacional en general es sumamente grande<sup>72</sup>, por lo que intentar identificar por parte nuestra algún tipo de patrón a partir de ellas es también una tarea improcedente. Sería una tarea de nunca acabar.

En el caso, por ejemplo, de las redes neuronales artificiales, la identificación de patrones a partir de esa enorme cantidad de imágenes ocurre regularmente por parte de ellas mismas. Recuérdese lo que ocurre en una de las etapas en el entrenamiento de *AlphaFold*. La red genera, basándose en ciertas técnicas propias de la biología, una enorme cantidad de imágenes que representan mapas de distancia entre los aminoácidos constituyentes de la proteína de la cual se pretenda predecir su estructura tridimensional específica. Sobre estos mapas de distancia la red se entrena en la identificación de ciertos patrones. Con este entrenamiento, entre otras cosas, la red autoajusta ciertos parámetros para generar nuevos mapas de distancia más discriminados. Siendo todo este proceso opaco para la supervisión humana. Lo que quiere decir que, dentro de las simulaciones computacionales, se da el caso de la generación de imágenes y procesamiento de estas sin supervisión humana. Y esto es algo que no ocurre en otros contextos de investigación científica.

La otra situación que cabe resaltar en cuanto a las imágenes generadas por simulación computacional es aquella en la que éstas figuran como especie de puentes de acceso o de observación para los investigadores a lo que ocurre “semiautónomamente” en la ejecución de estos modelos que involucran entrenamiento y búsqueda exploratoria con cierta dosis de aleatoriedad. Es decir, las imágenes en esta situación terminan siendo nuestro vínculo

---

<sup>72</sup> Un ejemplo de esto puede darse con lo sucedido en la experimentación subyacente a la famosa “fotografía” de un agujero negro que dio la vuelta al mundo hace muy pocos años. De ella hablaremos un poco más adelante.

expedito con el modelo computacional en ejecución. Básicamente, nos permiten *interactuar* con el modelo y esta suerte de interacción arroja un poco de luz sobre la gran opacidad epistémica que subyace en la obtención de ciertos resultados mediante el despliegue de modelos computacionales. En tanto que, es primera vez en nuestra historia que contemplamos una especie de autoridad epistémica *artificial* luego, por ejemplo, del éxito abrumador que vienen sosteniendo las redes neuronales artificiales, y, debido a que las imágenes que se generan a partir de sus ejecuciones terminan siendo en varios casos nuestro punto de acceso a ellas, podría considerarse que esta situación en términos de visualización es novedosa. Esta situación particular es propia de las simulaciones computacionales. Por lo que, junto con Humprheys, Winsberg, Galison y Lenhard, por mencionar algunos, se considera aquí que las simulaciones computacionales sí representan aspectos novedosos en materia de discusión filosófica.

En el análisis que realizaron Frigg y Reiss, cuando examinan la posibilidad de alguna novedad en cuanto a lo epistémico de las simulaciones computacionales, se concentran principalmente en las características que Winsberg bajo ese aspecto identificó: a saber: de arriba hacia abajo (*downward*), autónomas y abigarradas (*motley*) (Frigg y Reiss, 2009, p. 598). Sin embargo, no realizaron un análisis bajo este respecto de la opacidad que se ha venido discutiendo en relación con la visualización y manipulación. A continuación, bajo la postura de que las simulaciones computacionales sí representan alguna novedad en el terreno de la filosofía, se hablará un poco más sobre manipulacionismo, visualización, opacidad epistémica, experimentación iterada y consideraciones prácticas como posibles aspectos novedosos de las simulaciones computacionales.

#### IV.3.1 Visualización y manipulación

Uno de los posibles aspectos epistémicos disruptivos, que se nos presenta a partir del uso de métodos de simulación computacionales en la investigación científica, es el enorme poder de visualización que suele venir con ellos acompañado. Un claro ejemplo de esto podemos encontrarlo el 10 de abril del año 2019. Ese día, la atención de no pocas personas en todo el mundo estuvo fijada en la imagen, lograda por un amplio grupo conformado por varios equipos de trabajo, de un objeto a unos 50.000.000 de años luz del planeta que habitamos (Landau, 2019). Es decir, la imagen que se logró, bajo la red internacional colaborativa de radio telescopios llamada *Event Horizon Telescope* (EHT) fue de un objeto a una distancia extremadamente lejana de donde nos encontramos. Pero, lo hitico de este suceso, va un poco más allá de haber logrado una imagen de un objeto a una distancia sumamente abrumadora; se halla más bien, dada las particularidades de ese objeto, justamente en lograr la primera representación visual a partir de una especie de referencia directa, a saber: fotones de hace unos cincuenta millones de años.

El 10 de abril de 2019, *The Washington Post*, por ejemplo, titulaba su sección de Ciencia con la siguiente noticia: “*See a black hole for the first time in a historic image from the Event Horizon Telescope*”<sup>73</sup>. La NASA, por su parte, titulaba en su página web que la imagen de un agujero negro hacía historia<sup>74</sup> (*Black Hole Image Makes History; NASA Telescopes Coordinated Observations*). Ese día entonces se tenía la primera “fotografía” de un agujero negro. Sin embargo, esta especie de “fotografía” podría concebirse más bien como una suerte de organización visual de un flujo de datos de proporciones astronómica (*petabytes*), donde

---

<sup>73</sup><https://www.washingtonpost.com/science/2019/04/10/see-black-hole-first-time-images-event-horizon-telescope/>

<sup>74</sup> [https://www.nasa.gov/mission\\_pages/chandra/news/black-hole-image-makes-history](https://www.nasa.gov/mission_pages/chandra/news/black-hole-image-makes-history)

confluían, además de muchísimo conocimiento en mecánica relativista y electromagnetismo -entre otras áreas de la física-, diversas técnicas de optimización y generación de imágenes dentro del *Machine Learning*, como por ejemplo el método RML (*Regularized Maximum Likelihood*)

La fotografía de M87 -el agujero negro en cuestión-, en cuya captación o recepción de fotones intervinieron sincronizadamente alrededor del mundo ocho potentes radio telescopios (Landau, 2019), podría considerarse como una “tomada” computacionalmente. Es decir, la imagen se generó básicamente por computadora. Y en el proceso de generación, mediante simulación computacional, que permitía la manipulación y exploración de los resultados visuales de las variaciones en los parámetros considerados<sup>75</sup>, hubo muchísima capacidad de procesamiento. Capacidad que se logra, como se había mencionado en el capítulo anterior, a partir de ordenadores con arquitecturas GPUs (*Graphics Processing Unit*)

---

<sup>75</sup> Para una exposición detallada de los parámetros considerados Cf. (Akiyama, et. al, 2019)

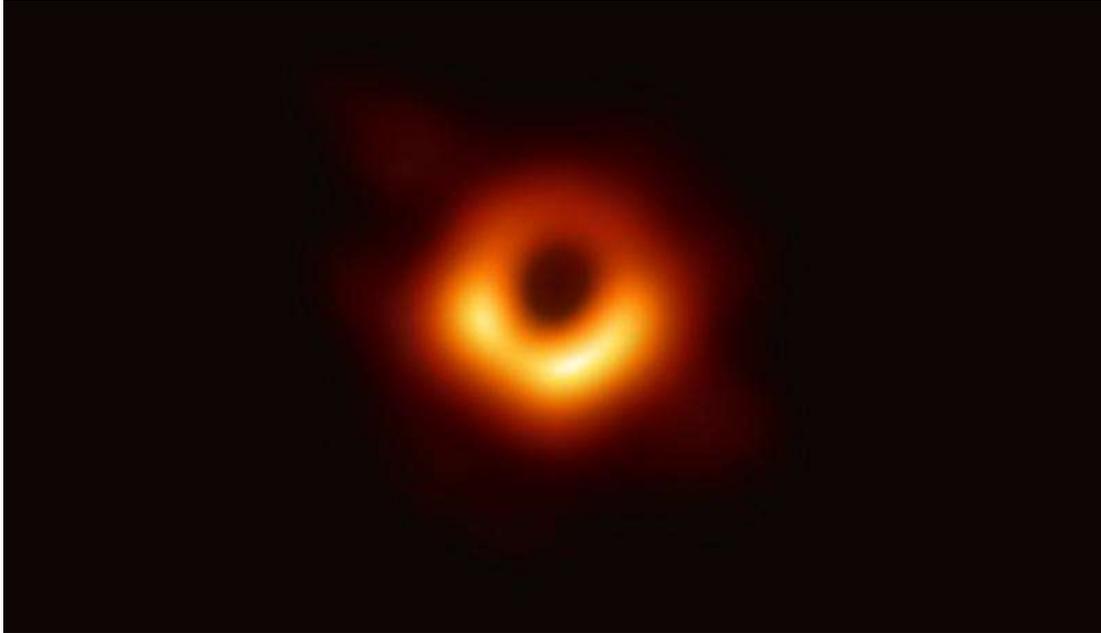


Imagen tomada de la página de la NASA  
[https://www.nasa.gov/mission\\_pages/chandra/news/black-hole-image-makes-history](https://www.nasa.gov/mission_pages/chandra/news/black-hole-image-makes-history)

“El resultado principal de los datos arrojados por los radio-interferómetros como el *Atacama Large Millimeter Array* (ALMA<sup>76</sup>) es un conjunto de visibilidades, o muestras de la transformada de Fourier de la distribución del brillo del cielo astronómico. Para estudiar una fuente astronómica en el plano de la imagen, debemos invertir esta transformación para identificar la imagen más consistente con los datos. Los métodos tradicionales como CLEAN (implementado en el popular paquete de software de instalaciones CASA) logran esto realizando primero una transformada inversa de Fourier de las muestras de visibilidad y luego eliminando los efectos de la PSF subóptima a través de la deconvolución del plano de la imagen. Por otro lado, las técnicas de imagen de Máxima Verosimilitud Regularizada (RML), como las utilizadas por EHT para producir las imágenes recientes de la sombra del agujero negro

---

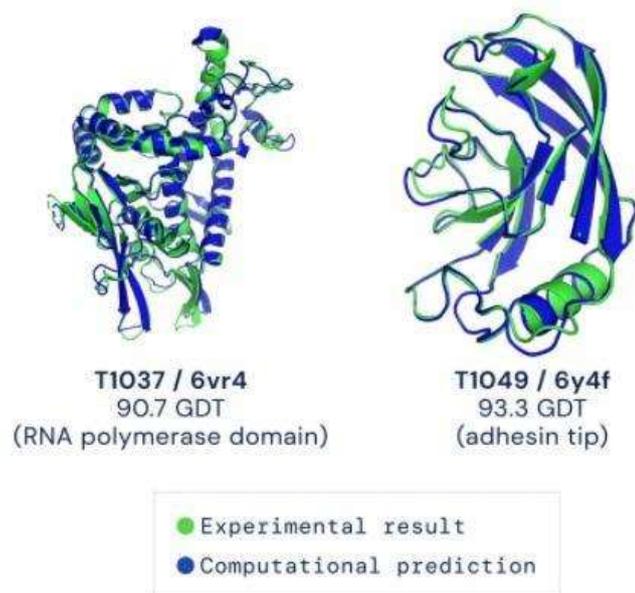
<sup>76</sup> ALMA es uno de los ocho radiotelescopios que integra EHT.

supermasivo M87, funcionan modelando hacia adelante los datos de visibilidad. Aunque históricamente las técnicas CLEAN han sido más populares que los flujos de trabajo RML, los desarrollos recientes en arquitecturas informáticas (GPU) y software (marcos de autodiferenciación de aprendizaje automático) han hecho posible la creación rápida de imágenes RML, incluida la exploración eficiente y la validación cruzada de la configuración de hiperparámetros. Emocionantemente, la naturaleza componible y probabilística de las imágenes RML abre muchas oportunidades para flujos de imágenes de trabajo avanzados, incluida la regularización del espacio de velocidad y la autocalibración simultánea de datos.”<sup>77</sup> (Czekala et. al, 2021)

---

<sup>77</sup> *The core data product of radio interferometers like the Atacama Large Millimeter Array (ALMA) is a set of visibilities, or samples of the Fourier transform of the astronomical sky brightness distribution. To study an astronomical source in the image-plane, we must invert this transformation to identify the image most consistent with the data. Traditional methods like CLEAN (implemented in the popular CASA facility software package) accomplish this by first performing an inverse Fourier transform of the visibility samples and then removing the effects of the suboptimal PSF through image-plane deconvolution. On the other hand, Regularized Maximum Likelihood (RML) imaging techniques, like those used to produce the recent EHT images of the M87 supermassive black hole shadow, work by forward-modeling the visibility data. Though CLEAN techniques have historically been more popular than RML workflows, recent developments in computing architectures (GPUs) and software (machine learning autodifferentiation frameworks) have made rapid RML imaging possible, including efficient exploration and cross-validation of hyperparameter settings. Excitingly, the composable and probabilistic nature of RML imaging opens up many opportunities for advanced imaging workflows, including velocity-space regularization and simultaneous data self-calibration. (Czekala et. al, 2021)*

Otro ejemplo significativo, concerniente al poder de visualización que suele venir acompañado con las simulaciones computacionales, es sin duda el que podemos encontrar en *AlphaFold*. Como se ha dicho, *AlphaFold* predice la estructura o configuración tridimensional específica de una proteína, dada su secuencia de aminoácidos constituyente. En dicha predicción interviene fuertemente un algoritmo de optimización que va haciendo ajustes, siguiendo rutas aleatorias, sobre un conjunto de valores paramétricos a fin de lograr una gran coincidencia entre la estructura predicha y la estructura esperada. En este caso la simulación permite ir visualizando si el resultado de esos ajustes permite un acercamiento o no al resultado esperado.



Two examples of protein targets in the free modelling category. AlphaFold predicts highly accurate structures measured against experimental result.

Imagen tomada de la página de DeepMind:

<https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology>

El papel de las imágenes en la investigación científica sin duda que cobra una gran relevancia. Para ilustrarlo se puede citar por ejemplo la obra de Peter Galison, '*Image and Logic*' (1997). Sin embargo, hay un agregado con las imágenes que se logran a partir de una simulación computacional, y es que pueden ser imágenes dinámicas. Las técnicas de visualización incorporadas a las simulaciones computacionales ofrecen imágenes que permiten ir contemplando inmediatamente los resultados que se desprenden de la exploración y manipulación de los parámetros propuestos como relevantes en la dinámica del sistema objeto de estudio. Sin contar además que permiten un considerable control en la variación de colores, perspectivas, texturas, etc. (Lenhard, 2016, p. 8)

En un *paper* de 2008, '*The philosophical novelty of computer simulation methods*', Paul Humphreys nos planteaba que de alguna manera el auge de las simulaciones computacionales nos desplazaba un poco del centro de la escena epistemológica, en tanto que:

“[p]ara un número creciente de campos de la ciencia, una epistemología exclusivamente antropocéntrica ya no es apropiada, porque ahora existen autoridades epistémicas superiores que no son humanos. Así que ahora nos enfrentamos a un problema, que podemos llamar el *predicamento antropocéntrico*, de cómo nosotros, como humanos, podemos entender y evaluar métodos científicos basados en computación que trascienden nuestras propias habilidades.”<sup>78</sup> (Humphreys, 2008, p. 617).

Frente a este problema que nos plantea Humphreys, las técnicas de visualización implementadas en las simulaciones computacionales llegan en nuestro auxilio. Pues la

---

<sup>78</sup> *For an increasing number of fields in science, an exclusively anthropocentric epistemology is no longer appropriate because there now exist superior, non-human, epistemic authorities. So we are now faced with a problem, which we can call the anthropocentric predicament, of how we, as humans, can understand and evaluate computationally based scientific methods that transcend our own abilities.* (Humphreys, 2008, p. 617)

enorme capacidad que posee el ser humano de identificar patrones a partir de estímulos visuales lo conecta poderosamente con el modelo dinámico que está siendo ejecutado computacionalmente y con los datos que dicha ejecución va generando.

“[L]os modelos de simulación explotan la capacidad de las computadoras para realizar un gran número de cálculos simples (lógicos). Los seres humanos no son particularmente buenos para prever el resultado de este tipo de procedimiento, por lo que se necesita experimentación. Sin embargo, los seres humanos están equipados con una gran capacidad para procesar datos visuales. Es aquí donde entra la visualización. Por ejemplo, una simulación de un patrón de circulación meteorológica podría proporcionar una gran variedad de datos multidimensionales (simulados). La visualización puede condensar esta gran matriz en una imagen y hacer comprensible qué tan bien se ajusta la simulación a fenómenos ya conocidos, como celdas de circulación global. Las visualizaciones crean así un vínculo de regreso a la cognición humana.”<sup>79</sup> (Lenhard, 2016, p. 9)

En definitiva, las visualizaciones son las que permiten a las personas que están llevando a cabo una investigación científica, por medio de simulación computacional, *interactuar* con el modelo dinámico que está siendo ejecutado computacionalmente. Y es gracias a ellas también que se explota la posibilidad de manipulación de la que se ha venido hablando a lo largo de la tesis. Para el caso de *AlphaFold*, por ejemplo, era de suma importancia poder

---

<sup>79</sup> *We have already seen that simulation models exploit the capability of computers for performing a large number of simple (logical) calculations. Human beings are not particularly good in foreseeing the outcome of this sort of procedure therefore experimentation is needed. However, human beings are equipped with a strong capability to process visual data. This is where visualization comes in. For example, a simulation of a meteorological circulation pattern might provide a large array of multidimensional (simulated) data. Visualization can condense this large array into an image and make graspable how well the simulation fits already known phenomena, like global circulation cells. Visualizations thus create a link back to human cognition.* (Lenhard, 2016, p. 9)

visualizar cómo se comportaba el modelo y cómo respondía a variaciones en los valores de ciertos parámetros. La prueba crucial consistía en observar si había una coincidencia significativa entre la estructura predicha y la estructura esperada. Igualmente, para el caso de EHT (*Event Horizon Telescope*). La visualización de alrededor unas 60 mil imágenes les permitió ir ajustando los parámetros hasta lograr la imagen que dio la vuelta al mundo.

#### *IV.3.1.1 Manipulacionismo*

Visualización y manipulación se encuentran tan estrechamente relacionadas en las simulaciones computacionales que resultan prácticamente inseparables. El gran poder que nos representan las técnicas de visualización en las simulaciones computacionales se explota justamente en la exploración manipulacionista del comportamiento del modelo dinámico. Por su parte, teniendo en cuenta la enorme cantidad de elementos involucrados en la dinámica de los sistemas complejos en consideración, y su gran sensibilidad a ligeras variaciones en las condiciones iniciales, al ser sistemas que presentan comportamientos caóticos, un manipulacionismo a ciegas en esta clase de sistemas sería un aspecto epistémico con muy poco sentido. Piénsese, por ejemplo, en una simulación computacional llevada a cabo mediante redes neuronales artificiales de tipo *Deep learning*. Sencillamente resulta impensable intentar una estrategia manipulacionista sin la posibilidad de visualización, ya que sin ella no tendríamos capacidad de vislumbrar siquiera, con relación al *output* pretendido, mínimamente los resultados de dicha manipulación.

#### *IV. 3.2 Experimentación*

Algunos de los propósitos epistémicos generales que pueden identificarse del empleo de las simulaciones computacionales en la investigación científica serían los siguientes:

*heurístico, predictivo y asimilativo.* En palabras de Eric Winsberg (2013, p. 7): “[I]as simulaciones se pueden utilizar con fines heurísticos, con el fin de predecir datos que no tenemos y para generar comprensión de los datos que ya tenemos.”<sup>80</sup>

Cuando se habla de un propósito heurístico en las simulaciones computacionales se está haciendo alusión al propósito de resolver un problema para el que no existe una solución de manera algorítmica, donde se usan más bien métodos que involucran muchísimo ensayo y error (Humphreys, 1995, p. 1). Es decir, cuando se habla de heurística en las simulaciones computacionales las palabras claves con que las que habría que asociarla inmediatamente son ‘ensayo y error’ y ‘solución de un problema’ (Humphreys, 1995, p. 1).

Regularmente, cuando se habla de ‘solución de un problema’ en este contexto, se está haciendo alusión a una intratabilidad analítica del sistema de ecuaciones con que se pretende modelar la dinámica de un sistema. Es por eso por lo que cuando se habla de heurística en este contexto se está hablando, regularmente, de métodos de solución o aproximación numérica para un sistema, frecuentemente de ecuaciones diferenciales parciales, que representan, en forma matemática, un modelo científico subyacente (ibid., 1995, p. 1).

Para lograr hacerle frente entonces a la intratabilidad analítica que representan estos modelos, con los que se suelen abordar por ejemplo los estudios de sistemas complejos, se vienen desarrollando una clase de experimentos con muy buenos resultados; experimentos en los que se despliegan una serie de técnicas para arribar a una gran colección de datos mediante procesos iterativos que pueden ser ejecutados en el ordenador (Lenhard, 2016, p. 8); como sucede por ejemplo con las redes neuronales artificiales.

---

<sup>80</sup> *Simulations can be used for heuristic purposes, for the purpose of predicting data that we do not have, and for generating understanding of data that we do already have.* (Winsberg, 2013, p. 7)

Como se ha venido insistiendo:

“Las redes neuronales artificiales (RNA) tienen una estructura genérica donde los nodos (neuronas) están ordenados en capas y los nodos de las capas vecinas están conectados, asemejándose a la estructura de una red neuronal fisiológica. Las señales entrantes a la primera capa se procesan a lo largo de las conexiones hasta llegar a la última capa. Dependiendo de la intensidad o pesos de las conexiones (que regulan cuándo una neurona "dispara"), tales RNA pueden exhibir un comportamiento universal, es decir, pueden generar todos los patrones Turing-computables. El punto es que, al variar sistemáticamente los parámetros involucrados, es decir, las intensidades o los pesos de conexión, se puede entrenar una red para generar patrones *particulares*, como reconocer el habla de una persona (input auditivo, output escrito). Hablar sobre el entrenamiento de una red es otra forma de hablar sobre el uso de la experimentación iterada en una forma sistemática de modificar el comportamiento del modelo.”<sup>81</sup>

(Lenhard, 2016, p. 8)

Un claro ejemplo de esto es justamente la experimentación llevada a cabo por *DeepMind*, con *AlphaFold*, para resolver el problema del plegamiento de proteínas. La simulación llevada a cabo por este prestigioso laboratorio de inteligencia artificial logra, significativamente, *predecir* la estructura tridimensional de una proteína a partir de su

---

<sup>81</sup> *Artificial neural networks (ANN) have a generic structure where nodes (neurons) are ordered in layers and nodes of neighboring layers are connected, resembling the structure of a physiological neural network. Incoming signals to the first layer are processed along the connections until they reach the last layer. Depending on the strength of the single connections (regulating when a neuron “fires”), such ANNs can exhibit universal behavior, i.e. can generate all Turing-computable patterns. The point is that by systematically varying parameters, i.e. connection strengths, one can train a network to generate **particular** patterns, like recognizing the speech of a person (audio input, written output). Talking about training a network, however, is another way of talking about using iterated experimentation in a systematic way for modifying model behavior.* (Lenhard, 2016, p. 8)

secuencia de aminoácidos constitutiva; resuelven un problema, claramente, de forma *heurística* (Recuérdese su intratabilidad analítica, por ejemplo, a partir de la función de onda de Schrödinger. No había un método tan eficiente como el que desarrollaron para lograr tal predicción. Prueba de ello es la abrumadora ventaja que les sacó al resto de grupos y centros investigativos que pretendían lo mismo). Asimismo, permiten una suerte de *asimilación* o *comprensión* de la información que se dispone sobre algunas propiedades emergentes para poder, a partir de esa comprensión, abordarlas.

Se ha venido planteando que el quid de las propiedades emergentes está básicamente en la identificación de una configuración estructural bien específica, que se da a partir de la interacción de las múltiples partes constituyentes de un sistema complejo. Siendo quizás la principal fuente de dificultad para abordarlas justamente la enorme cantidad de elementos que se hallan en interacción para dar lugar a ella; interacción que además se da de manera no lineal. En este caso, en la implementación de las redes neuronales artificiales, se hace uso de esa información y se exploran alternativas para lograr abordar la no linealidad de las múltiples interacciones que subyacen a la emergencia de una propiedad o comportamiento. Y, como se ha visto, tuvieron éxito. El problema quizás está en que dicho éxito se alcanza pagando un precio no menor: el predicamento antropocéntrico (*anthropocentric predicament*), como le llama Humphreys.

A partir, por ejemplo, del modelo computacional que constituye *AlphaFold*, se tiene una poderosa herramienta algorítmica para llegar a la configuración estructural específica que se da de la interacción no lineal de los elementos involucrados (aminoácido) en la propiedad emergente a abordar (funcionalidad), a partir de dichos elementos involucrados; que es de lo que se dispone. Esto, como se ha dicho, representaba un gran reto. Si el número de configuraciones estructurales posibles es tan gigantesco, ¿cómo encontrar justo la que está

ligada a la propiedad emergente en cuestión? *DeepMind* lo consiguió gracias a un grupo de trabajo de mentes seguramente muy brillantes. Sin embargo, aún para un grupo de mentes brillantes es prácticamente imposible llevar a cabo el tipo de procesamiento de información que se lleva a cabo mediante los ordenadores que emplean para la simulación computacional.

#### IV.3.3 *Opacidad epistémica y consideraciones prácticas en las simulaciones computacionales*

“Muchas de las características -quizás todas- que son especiales de las simulaciones computacionales son el resultado de esta incapacidad de las habilidades cognitivas humanas para conocer y comprender los detalles del proceso computacional. Los cálculos involucrados en la mayoría de las simulaciones son tan rápidos y complejos que ningún humano o grupo de humanos puede en la práctica reproducir o comprender los procesos involucrados en dichos cálculos.”<sup>82</sup> (Humphreys, 2008, pp. 618, 619)

Esto introduce, no sólo para el caso de interés, sino en general para aquellos resultados a los que se haya llegado por medio de simulaciones computacionales, cierta *opacidad epistémica* entre el input y el output; en el caso de *AlphaFold*: entre la secuencia de aminoácidos y la estructura tridimensional específica que la hace funcional. La enorme cantidad de pasos computacionales involucrados en una simulación hacen del modelo subyacente uno completamente irrelevante ateniéndonos a nuestras capacidades.

“La opacidad epistémica es una característica típica de las simulaciones computacionales. Sin embargo, no es una limitación. Los modelos de simulación pueden proporcionar medios para controlar una dinámica, *aunque* estos modelos sean

---

<sup>82</sup> *Many, perhaps all, of the features that are special to simulations are a result of this inability of human cognitive abilities to know and understand the details of the computational process. The computations involved in most simulations are so fast and so complex that no human or group of humans can in practice reproduce or understand the processes.* (Humphreys, 2008, pp. 618, 619)

epistémicamente opacos. Los investigadores y profesionales en el área pueden emplear una serie de experimentos de simulación iterados y visualizaciones para probar cómo se relacionan las variables de entrada con las variables de salida. De esta manera pueden [interactuar] con el modelo, incluso si la dinámica de la simulación permanece (al menos parcialmente) opaca. Es cierto que tal tipo de familiaridad con el modelo no alcanza los altos estándares epistémicos que generalmente se atribuyen a las matemáticas. Sin embargo, este estándar más bien bajo sigue siendo suficiente cuando el objetivo consiste en una intervención [o *manipulación*] *controlada*<sup>83</sup>.”<sup>84</sup> (Lenhard, 2016, p. 19)

Esto es precisamente una de las cosas que se ha venido tratando de sostener de manera argumentada en esta investigación. Básicamente, que las simulaciones computacionales permiten, gracias a las poderosas técnicas de visualización que vienen con ellas incorporadas, un tipo de interacción *manipulacionista* entre el modelo subyacente y los datos que computacionalmente se van generando a partir de su ejecución, incluso cuando la dinámica del modelo permanezca (parcialmente) opaca (Lenhard, 2016, p. 20). Asimismo, se podría plantear que las simulaciones computacionales cambian de alguna manera la concepción subyacente al modelado matemático como tal, basada en una suerte de transparencia epistémica. Por ejemplo, podría cuestionarse si lo que se logra mediante métodos de aproximación numérica o métodos de discretización constituyen propiamente una solución.

---

<sup>83</sup> Cursiva introducida aquí.

<sup>84</sup> *Epistemic opacity is a typical feature of simulations. However, it does not block simulation modeling. Simulation models can provide means to control a dynamics, **although** these models are epistemically opaque. Researchers and practitioners can employ a series of iterated simulation experiments and visualizations for testing how varying inputs are related to outputs. In this way they can orient themselves in the model – even if the dynamics of the simulation remain (at least partly) opaque. Admittedly, such a kind of acquaintance with the model falls short of the high epistemic standards that usually are ascribed to mathematical models. This lower standard is still sufficient, however, when the aim is controlled intervention.* (Lenhard, 2016, p. 19)

“Cuando la simulación se caracteriza como una solución numérica, esto invoca un sentido pragmático de solución que difiere significativamente del sentido matemático habitual.

Cuando hablamos de solución en el sentido propio, normalmente nos referimos a la noción matemática de solución. Esta es una noción estricta bastante relevante en las ingenierías y en una amplia clase de ciencias naturales, donde se emplea el lenguaje del cálculo diferencial e integral, y -aún más general- donde el lenguaje de las ecuaciones matemáticas se emplea para formular modelos teóricos. Si se tiene una ecuación con un término desconocido, el término matemático que satisface la igualdad es una solución.

Lo que cuenta como una solución en sentido estricto depende únicamente de la ecuación matemática que se intenta resolver (incluidas las definiciones del espacio matemático subyacente). Una ecuación cuadrática, por ejemplo, puede ser escrita como

$$a \cdot x^2 + bx + c = 0$$

y la solución son exactamente esos valores de  $x$  que satisfacen esta igualdad (con números  $a$ ,  $b$  y  $c$  dados). Este sentido estricto de una solución se generaliza fácilmente a problemas más complicado que una sola ecuación. Si se tiene un sistema de ecuaciones diferenciales parciales, por ejemplo, un término dado lo satisfará (resolverá) o no. Hablando de la noción de solución, no importa si uno realmente puede encontrar la solución por integración. Los sistemas de ecuaciones diferenciales

parciales (EDP) son endémicos en muchas ciencias, pero por lo general no se pueden resolver analíticamente.”<sup>85</sup> (Lenhard, 2016, pp. 16, 17)

Esto ha llevado a que se plantee una distinción importante en este contexto. Y es la distinción entre ‘*soluciones en principio*’ o soluciones en teoría, y ‘*consideraciones prácticas*’. Lo que se logra generalmente con métodos de simulación computacional son consideraciones prácticas frente a la regular intratabilidad analítica del esquema matemático con el que se intenta modelar la dinámica del sistema objeto de estudio, más allá de que no se esté resolviendo propiamente el sistema de ecuaciones. Pero esas consideraciones prácticas vienen avaladas en gran medida por una correspondencia visual entre lo que se pretende generar y lo que efectivamente se genera al partir del modelo empleado. Es decir, en gran medida el criterio epistémico relevante para evaluar la validez del modelo consiste en la visualización de los datos generados a partir de la ejecución computacional del modelo. En la adecuación del modelo propuesto para el abordaje de la dinámica del sistema en consideración la visualización juega un papel muy importante.

---

<sup>85</sup> *When simulation is characterized as numerical solution, this invokes a pragmatic sense of solution that differs significantly from the usual mathematical sense.*

*When we talk of solution in the proper sense, then we normally refer to the mathematical notion of solution. This is a rather strict notion relevant in a broad class of natural and engineering sciences, namely where the language of differential and integral calculus is employed, and – even more general – where the language of mathematical equations is employed to formulate theoretical models. If you have an equation with one unknown term, the mathematical term that satisfies your equation is a solution.*

*What counts as a solution in the strict sense solely depends on the mathematical equations that get solved (including definitions of the underlying mathematical space). A quadratic equation, for example, can be written as*

$$a \cdot x^2 + b \cdot x + c = 0$$

*and solutions are exactly those numbers  $x$  that satisfy this equation (with given numbers  $a$ ,  $b$ ,  $c$ ). This strict sense of a solution easily generalizes to problems more complicated than one single equation. If you have a system of partial differential equations, for instance, a given term will either satisfy (solve) it or not. Talking about the notion of solution, it does not matter whether one can actually calculate the solution by integration. Systems of partial differential equations (PDEs) are endemic in many sciences, but they are usually not analytically solvable. (Lenhard, 2016, pp. 16, 17)*

Pretensiones epistémicas como la transparencia podrían considerarse un poco desplazadas cuando de abordar la complejidad se trata. Frente al estudio de sistemas complejos una comprensión o entendimiento analítico es reemplazado en gran medida por una comprensión o entendimiento manipulacionista enfocado quizás más bien en lograr resolver un problema heurísticamente y ofrecer predicciones. Como ocurre perfectamente con *AlphaFold* y el problema del plegamiento de las proteínas. Las simulaciones computacionales permiten hacerles frente a las barreras de la complejidad antes que eludirlas. Presentan una alternativa frente a la frustración de una comprensión analítica de la complejidad. (Lenhard, 2016, p. 20)

## CONCLUSIONES

En esta investigación doctoral me he propuesto, en primer lugar, analizar algunos de los rasgos estimados distintivos de los sistemas complejos. Esto, con el objetivo expreso de poner de manifiesto las serias dificultades aprehensivas, bajo un único esquema general, que este tipo de sistemas acarrea. Los sistemas complejos son, entre otras cosas, sistemas cuyas dinámicas suelen presentar comportamientos caóticos, constituidos de muchísimas partes. Esto representa una gran posibilidad siempre latente de que en ellos emerjan propiedades, y comportamientos enteramente novedosos, sin contemplar en el esquema o en los esquemas previos para su clasificación o intento aproximativo de definición.

Al ser sistemas con ciertos comportamientos caóticos, son sistemas susceptibles de experimentar cambios en sus comportamientos, o presentar nuevas propiedades, debido a ligeras variaciones respecto a algunos de sus muchísimos ítems o elementos constitutivos; ligeras variaciones que pueden ocurrir, en gran medida, debido a la alta exposición a mecanismos de cambio, donde el azar juega ciertamente un papel no menor. Las mutaciones son un claro ejemplo de esto.

Básicamente, lo que se intentó en un primer momento en esta investigación doctoral fue plantear que pretensiones epistemológicas unificacionistas resultan fuertemente frustradas cuando del fenómeno de la complejidad se trata. La complejidad, debido a la naturaleza de los aspectos que hasta el momento se han logrado identificar como distintivos de los sistemas en los que se hace manifiesta (emergencia, interacción no lineal entre sus muchísimas partes constitutivas, adaptación, etc.) parece frustrar, por lo menos hasta el momento, pretensiones epistemológicas de unificación bajo un único esquema general con el que dar cuenta de ella;

pretensiones exitosamente alcanzadas por ejemplo en los casos de Maxwell, Newton, Darwin, Schrödinger, por mencionar sólo algunos.

El proceso de evolución biológica -permítaseme la expresión- no da “productos” acabados. Nuevas formas, funciones, estructuras, organizaciones se van generando de los mecanismos de variación subyacentes en dicho proceso. Esto abre la posibilidad a que surjan posibles nuevos aspectos dentro del panorama intuitivo en el que observamos y tratamos de estudiar la complejidad.

La complejidad, podría considerarse, surge a partir de los mecanismos de cambio involucrados en la dinámica de adaptación de los organismos biológicos en su interacción con su entorno. Al ser una dinámica no determinista, resulta en extremo difícil de predecir cuáles podrían ser esas posibles novedades, bien sea en cuanto a funcionalidad, organización, estructura, forma. Estas novedades pueden significar a su vez nuevos aspectos con los que estimar complejidad, dejando rezagado, o rezagados, al esquema, o a los esquemas previos, con el que, o con los que se intenta dar cuenta de ella. Sin embargo, esto no constituye una invitación a que se dejen de proponer alternativas teóricas para un abordaje general de la misma.

Intentos teóricos como el de Murray Gell-Mann, con la noción de complejidad efectiva; el de Brian Goodwin, con la noción de complejidad morfológica mediante el número de simetrías rotas durante el proceso de desarrollo; o el de Miguel Fuentes, con la noción de la complejidad y la emergencia modelo paramétricas, por mencionar sólo algunos, representan avances significativos para su abordaje. Particularmente, para un abordaje de carácter métrico, y no propiamente uno de carácter definitorio.

Justamente, el segundo de los objetivos aquí planteados consistió en analizar algunos de los marcos teóricos con los que se han intentado abordar algunos de los rasgos estimados

distintivos de los sistemas complejos, destacando la propuesta de Miguel Fuentes, por hacerle frente quizás a la característica más distintiva de estos sistemas; a saber, la emergencia. Esto, haciendo especial énfasis en el carácter dinámico y caótico de los sistemas complejos, y en la importancia del modelado matemático como una poderosa herramienta de estudio frente a dicho carácter.

En el modelado de sistemas dinámicos no lineales y caóticos regularmente sistemas de ecuaciones diferenciales parciales juegan un papel muy importante. En la propuesta de Miguel Fuentes, dentro de dichos sistemas de ecuaciones, se hace además necesaria la identificación de ciertos parámetros, o de cierto parámetro, que al presentar ligeras variaciones en los valores que puede tomar, traiga consigo una alteración significativa en la dinámica del sistema objeto de estudio. Al ser sistemas caóticos, su dinámica es muy sensitiva a ligeras variaciones en los valores de ciertos o de cierto parámetro. La idea consiste entonces, como se pudo observar, en aplicar la complejidad de Kolmogorov, discriminando señal de ruido – lo que en últimas coincidiría básicamente con la aplicación de la complejidad efectiva- sobre los valores de dicho parámetro, y a partir de esta aplicación estimar complejidad y emergencia.

No todo sistema caótico es un sistema complejo. Pero todo sistema complejo presenta, entre otras cosas, comportamientos caóticos. Complicación matemática y caos no siempre van de la mano. Existen dinámicas caóticas que se pueden modelar con cierta adecuación mediante el empleo de esquemas matemáticos muy sencillos de resolver o ejecutar, como la ecuación o mapa logístico. Sin embargo, en los sistemas complejos, se suele dar el caso, como vimos por ejemplo con *Acetabularia*, que complicación matemática y caos sí van de la mano. De hecho, la complicación matemática es tal, que regularmente se requieren métodos aproximativos, que no analíticos, de “solución”. Esto hace que la propuesta de Miguel

Fuentes sea una, si bien sumamente interesante, difícil de aplicar en varios casos sin el auxilio del altísimo poder de cálculo de los ordenadores de los que disponemos hoy día. Pues, para aplicar la complejidad de Kolmogorov, sobre ciertos valores del parámetro de control, hay que resolver el sistema de ecuaciones. Es decir, habría que ver para qué valores de dicho parámetro, al resolverse el sistema, arroja una discontinuidad de tipo escalonada. El problema de fuste radica en “resolver” el sistema.

Precisamente, en concomitancia con el resultado principal esperado del análisis propuesto como segundo objetivo de este trabajo, que consistió en poner de manifiesto la gran relevancia métrica que cobra la propuesta de la complejidad y emergencia modelo paramétricas, haciendo especial énfasis en su aplicación en el ámbito de la biología, se intentó además evidenciar la enorme importancia de la simulación computacional en los intentos de abordaje de la complejidad. Simulación computacional y complejidad se encuentran estrechamente relacionadas; al punto tal de representar la primera una herramienta casi que necesaria para el abordaje de la segunda, como bien lo habían notado Horgan y Vicsek. Las simulaciones computacionales representan una poderosa herramienta para el despliegue de modelos que involucran esquemas matemáticos de muy difícil resolución analítica.

Por medio de simulación computacional no sólo se explota el altísimo poder de cómputo de los ordenadores para “resolver”, regularmente por métodos de aproximación numérica, el sistema de ecuaciones diferenciales parciales subyacente, sino que, además, y esto como valor agregado fundamental para un exitoso abordaje de la complejidad, se logra una exploración visual de las posibles consecuencias de manipular los valores de ciertos parámetros hasta encontrar una correlación entre estos y el fenómeno complejo/emergente a abordar. Simulación computacional y manipulacionismo vienen de la mano, enlazados mediante poderosas técnicas de visualización dinámicas.

En continuidad con dicho intento de evidenciar la enorme importancia de la simulación computacional para el abordaje de la complejidad, en el tercer objetivo de este trabajo, además de enfatizar en el giro predictivo que se ha venido dando recientemente en ciencias - como lo dejan saber por ejemplo Chirimuuta y Weiskopf- se intentó analizar un problema concreto de suma importancia en el ámbito de la biología y de la biomedicina de expresa manifestación de la complejidad y su asombrosa resolución por un método de simulación. Más concretamente, por medio de una red neuronal artificial de aprendizaje profundo. Mediante el ejemplo de AlphaFold, no sólo se enfatizó en la gran relevancia de las redes neuronales artificiales de aprendizaje profundo para un abordaje exitoso (predictivo) de la complejidad, sino que además se enfatizó en el carácter manipulacionista subyacente en dicho abordaje. Para el caso concreto hubo una extensa manipulación de los valores posibles de los ángulos de torsión de los aminoácidos involucrados en las proteínas consideradas hasta encontrar una gran coincidencia entre la estructura proteica predicha y la estructura proteica esperada. Esta exploración manipulacionista, que se suele llevar a cabo en las simulaciones computacionales, posibilitada por las técnicas de visualización regularmente con ellas incorporadas, en el caso concreto de las redes neuronales artificiales de aprendizaje profundo viene optimizada por los algoritmos de *back propagation* y *stochastic gradient descent*.

Estos algoritmos, dada la enorme cantidad de elementos sobre los que estimar ciertos valores muy específicos con los que se establece una correlación con el fenómeno emergente/complejo a abordar, permiten una disminución sumamente significativa del número de alternativas posibles de valores a explorar. El número de alternativas posibles es prácticamente inconmensurable, de modo que sin ellos sería una tarea de nunca acabar encontrar el conjunto de valores paramétricos correlacionados con el fenómeno que se intenta dar cuenta. Aun así, el número de combinaciones de los valores posibles a explorar sigue

siendo extremadamente grande, sobrepasando enormemente las capacidades de cómputo del ser humano. El entrenamiento de estas redes requiere de muchísimas iteraciones, anulando esta situación la posibilidad de un seguimiento supervisado para las capacidades humanas. Esto representa una suerte de opacidad epistémica que abre nuevos escenarios de discusión filosófica en materia de metodología, y principalmente en materia de justificación.

Discutir entonces la opacidad epistémica como posible aspecto novedoso que trae consigo el empleo extensivo de las simulaciones computacionales ocupó nuestra atención en el cuarto y último objetivo de esta investigación. Lo planteado aquí, pese a las consideraciones importantes de Frigg y Reiss de que las simulaciones computacionales no introducen estrictamente aspectos novedosos en la discusión filosófica sobre modelos, es que la opacidad epistémica, junto con el manipulacionismo ligado a las poderosas técnicas de visualización, y la experimentación iterada sí introducen novedades epistemológicas. Novedades que se ligan positivamente con el marcado giro predictivo que se viene dando en ciencias en tiempos recientes.

Sobre la base del cumplimiento de estos objetivos, concluimos que un abordaje de tipo manipulacionista, posibilitado por el enorme poder de visualización que nos representan las simulaciones computacionales que nos permiten hoy día los ordenadores de los que se disponen, sobre todo de aquellos con *Graphics Processing Units* (GPU), y, con el apoyo, por ejemplo, de redes neuronales artificiales de tipo *Deep learning*, cada vez más poderosas en cuanto su capacidad de procesamiento de datos en paralelo, permite un abordaje exitoso de la complejidad; entendiéndose aquí por éxito la capacidad de resolver un problema concreto respecto de la identificación de un patrón de comportamiento del sistema complejo, y la replicación representacional de mismo mediante simulación computacional de un aspecto particular de ese comportamiento.

Este tipo de abordaje, sin embargo, implica una renuncia a pretensiones epistemológicas normativas, asumir cierta opacidad epistémica y una gran necesidad de recursos heurísticos para aproximarnos a una comprensión de los fenómenos complejos. Esta renuncia parcial a la normatividad epistémica no es, sin embargo, una derrota para la epistemología. Solamente supone modificar los fines de esta y su enriquecimiento con nuevas herramientas teórico-prácticas.

## BIBLIOGRAFÍA

- Akiyama, K. et al (2019), “First M87 Event Horizon Telescope Results. III. Data Processing and Calibration.” *The Astrophysical Journal Letters*, **875**: L3
- Alberch, P. (1998), “El concepto de progreso y la búsqueda de teorías generales en la evolución”, en: J. Wagensberg y J. Agustí (eds.), *El progreso: ¿un concepto acabado o emergente?* Barcelona: Tusquets, 107-130.
- Alzogaray, R. (2008), *Una tumba para los Romanov (y otras historias con ADN)*. Buenos Aires: Siglo Veintiuno.
- Anderson, P. (1972), “More is Different. Broken symmetry and the nature of the hierarchical structure of science”, *Science* vol. **177**: 393-396
- Anderson, P. (1994), *Physics: The Opening to Complexity*, Princeton University Press (Preface to the NAS Proceedings of the Colloquium on Physics: The Opening of Complexity, June 27-28)
- Beatty, J. (1990), “Evolutionary anti-reductionism: historical reflections”, *Biology and Philosophy*, **5**:197-210.
- Bechtel, W. y Richardson, R. (1993), *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, Princeton: Princeton University Press.

- Bedau, M. (2003), "Artificial life: organization, adaptation and complexity from the bottom up", *TRENDS in Cognitive Science*, 7(11).
- Bickle, J. (2016), "Revolutions in Neuroscience: Tool Development, *Frontiers in system neuroscience*, 10(24) <https://doi.org/10.3389/fnsys.2016.00024>
- Bonner, J. (1988), *The Evolution of Complexity*, Princeton NJ: Princeton University Press.
- Boyd, R. (1999), "Kinds, complexity, and multiple realization", *Philosophical Studies*, 95: 67-98.
- Brandon, R. (1996), "Reductionism versus holism versus mechanism", *Concepts and Methods in Evolutionary Biology*, Cambridge: Cambridge University Press, 179-205.
- Bunge, M. (2004), *Emergencia y Convergencia. Novedad Cualitativa y Unidad del Conocimiento* (Rafael González del Solar, Trad.), Barcelona: Gedisa. (Obra original publicada en 2003)
- Cadevall, M. (2005), "La Genética del Desarrollo: ¿Derribo o Ampliación del Darwinismo?", en: Anna Estany (ed.), *Filosofía de las Ciencias Naturales, Sociales y Matemáticas*, Madrid: Editorial Trotta, 177-195.
- Caponi, G. (2011), *La segunda agenda darwiniana. Contribución preliminar a una historia del programa adaptacionista*. México: Centro de estudios filosóficos, políticos y sociales Vicente Lombardo Toledano.
- Casanueva, M. y Méndez, G. (2005), "Tres Teorías y Tres Niveles en la Genética del Siglo XX", en: Anna Estany (ed.), *Filosofía de las Ciencias Naturales, Sociales y Matemáticas*, Madrid: Editorial Trotta, 197-224.

- Chaitin, G. (2002), “Información y Azar”, *Boletín de la Asociación Matemática Venezolana* **IX** (1): 55-81
- Chaitin, G. (2007), “Speculations on biology, information and complexity”, *EATCS Bulletin*, **91**: 231-237.
- Chirimuuta, M. (2020), “Prediction versus understanding in computational enhanced neuroscience”, *Synthese*, <https://doi.org/10.1007/s11229-020-02713-0>
- Czkala, I. et al. (2021), “Regularized Maximum Likelihood Imaging for ALMA, *American Astronomical Society meeting # 238*, id. 132. 09.
- Darwin, Ch. (1992), *El Origen de las Especies*, Barcelona: Editorial Planeta – De Agostini.
- Díez, J. (2005), “La Explicación Científica: Causalidad, Unificación y Subsunción Teórica”, en: Luis Eduardo Hoyos (ed.), *Relativismo y Racionalidad*, Bogotá: Universidad Nacional de Colombia., 383-414
- Díez, J. y Moulines, C. (1997), *Fundamentos de Filosofía de la Ciencia*, Barcelona: Editorial Ariel, S.A.
- Ericsson, K. A., & Kintsch, W. (1995), “Long-term working memory”, *Psychological Review*, **102**(2): 211–245
- Evans et al. (2018), *De Novo Structure Prediction with Deep-Learning Based Scoring*, <https://deepmind.com/blog/article/alphafold-casp13>
- Folguera, G. (2009), *Multiplicidad de procesos evolutivos y de entidades biológicas: Análisis crítico de las jerarquías genealógicas en la teoría de la evolución del siglo XX*. Tesis

de Licenciatura en Filosofía. Universidad de Buenos Aires. Facultad de Filosofía y Letras.

Ford, J. (1983), “How random is a coin toss?”, *Physics Today*, **36**: 40-47

Fontanelli, et al (2020), “Distribuciones de probabilidad en las ciencias de la complejidad: una perspectiva contemporánea”, *Inter disciplina* [online], **8**(22)

Frigg, R. y Reiss, J. (2009), “The philosophy of simulation: hot new issues or same old stew?”, *Synthese*, **169**: 593-613.

Fuentes, M. (2018), *Complejidad, emergencia y cambio teórico*. Tesis de posgrado. Universidad Nacional de la Plata. Facultad de Humanidades y Ciencias de la Educación.

Galison, P. (1996), “Computer Simulations and the Trading Zone”, en: Peter Galison y David J. Stump (eds.) *The Disunity of Science: Boundaries, Context and Power.*, Stanford University Press, pp 118-157.

Galison, P. (1997), *Image and Logic. A Material Culture of Microphysics*, Chicago: The University of Chicago Press.

Gell-Mann, M. (1998), *El Quark y el Jaguar: Aventuras en lo Simple y lo Complejo* (Ambrosio García y Romualdo Pastor, Trad.), Barcelona: Tusqued. (Obra original publicada en 1994).

Gilbert, S. F. y Sarkar, S. (2000), “Embracing complexity: organicism for the 21st century”, *Developmental Dynamics*, **219**: 1-9.

Ginobili, S. (2014), “La inconmensurabilidad empírica entre la teoría de la selección natural darwiniana y el diseño inteligente de la teología natural”, *THEORIA*, **81**: 375-394

Godfrey-Smith, P. (2009), *Darwinian Populations and Natural Selection*, Oxford Scholarship online.

Goldenfeld, N. y Kadanoff, L. P. (1999), "Simple lessons from complexity", *Science*, **284**: 87-89.

Goodwin, B. (1994), *How the Leopard Changed Its Spots: The Evolution of Complexity*, New York: Charles Scribner's Sons.

Goodwin, B. (1998), "Forma y transformación: la lógica del cambio evolutivo", en: J. Wagensberg y J. Agustí (eds.), *El progreso: ¿un concepto acabado o emergente?* Barcelona: Tusquets, 139-163.

Goodwin, B. y Solé, R. (2000), *Signs of Life: How Complexity Pervades Biology*, New York: Basic Books.

Gould, S. J. y Lewontin, R. (1979), "The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist program", *Proceedings of the Royal Society of London B*, **205**: 581-598.

Gould, S. J. (2002), *The Structure of Evolutionary Theory*, Harvard University Press.

Grene, M. (1987), "Hierarchies in biology", *American Scientist*, **75**: 504-510.

Griffiths, P. E. (2001), "Genetic information: a metaphor in search of a theory", *Philosophy of Science*, **68**: 394-412.

Hadamard, J. (1945), *An Essay on The Psychology of Invention in the Mathematical Field*, New York: Princeton University Press.

- Harman, G. y Kulkarni, S. (2007), *Reliable Reasoning. Induction and Statistical Learning Theory*, Cambridge: MIT press.
- Harms, W. (2004), *Information and Meaning in Evolutionary Processes*, Cambridge: Cambridge University Press.
- Hartmann, S. (1996), “The World as a Process: Simulations in the Natural and Social Sciences,” en: R. Hegselmann, et al. (eds.), *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*, Dordrecht: Kluwer.
- Hawkins, J. y Blakeslee, S. (2005), *On Intelligence: How a New Understanding of the Brain will Lead to the Creation of Truly Intelligent Machines*, St. Martin’s Griffin.
- Hidalgo, C. (2015), *Why Information Grows. The Evolution of Order, from Atoms to Economies*, New York: Basic Book.
- Holland, J., et al. (1993), *Induction. Processes of Inference, Learning, and Discovery*. MIT press (Primera edición 1989)
- Horgan, J. (1995), “From Complexity to Perplexity”, *Scientific American*, **272**:104-109.
- Hull, D. (1972), “Reductionism in genetics—biology or philosophy?”, *Philosophy of Science*, **39**: 491-499.
- Hull, D. (1974), *Philosophy of biological science*, New Jersey: Prentice-Hall.
- Hull, D. (1998), “Progreso panglossiano”, en: J. Wagensberg y J. Agustí (eds.), *El progreso: ¿un concepto acabado o emergente?* Barcelona: Tusquets, 107-130.
- Humphreys, P. (1995), “Computational Science and Scientific Method”, *Mind and Machines*, **5**: 499-512.

- Humphreys, P. (1997), “Emergence, Not Supervenience”, *Philosophy of Science*, **64**: S337-S345.
- Humphreys, P. (2004), *Extending Ourselves. Computational Sciences, Empiricism, and Scientific Method*, New York: Oxford University Press.
- Humphreys, P (2009), “The philosophical novelty of computer simulation methods” *Synthese*, **169**: 615-626.
- Johnson, S. (2003), *Sistemas emergentes. O qué tienen en común hormigas, neuronas, ciudades y software*. Madrid: Fondo de Cultura Económica.
- Jumper, J. et al. (2021), “Highly accurate protein structure prediction with AlphaFold”, *Nature*, **596**: 583-589.
- Kauffman, S. A. (1993), *The Origins of Order: Self-organization and Selection in Evolution*, New York: Oxford University Press.
- Kesner, E. y Newman, J. (1988), *Matemáticas e imaginación*, Barcelona: Ediciones Orbis
- Kitcher, P. (1993), *The Advancement of Science. Science Without Legend, Objectivity Without Illusion*, Oxford University Press.
- Kitcher, P. (2001), *Science, Truth and Democracy*, Oxford University Press.
- Kuhn, T. (1962), *The Structure of Scientific Revolutions*, University of Chicago Press.
- Laudan, L. (1978), *Progress and Its Problems: Toward a Theory of Scientific Growth*, University of California Press.

- Landau, E. (2019), “Black Hole Image Makes History; NASA Telescopes Coordinated Observations,” en: Sarah Loff (ed.) [https://www.nasa.gov/mission\\_pages/chandra/news/black-hole-image-makes-history](https://www.nasa.gov/mission_pages/chandra/news/black-hole-image-makes-history)
- Lenhard, J. (2016), “Computer Simulation”, en: Paul Humphreys (ed.) *The Oxford Handbook of Philosophy of Science*.
- Lenhard, J. (2017), “License to Explore: How Images Work in Simulation Modeling”, en: Sabine Ammon y Remei Capdevila-Werning (eds), *The Active Image. Architecture and Engineering in the Age of Modeling*. Springer.
- Laughlin, R. (2005), *A Different Universe. Reinventing Physics from the Bottom Down*, Basic Books.
- Lombardi, O. (2000b), “Prigogine y el azar de las bifurcaciones”, *Revista de Filosofía de la Universidad de Costa Rica*, **XXXVIII**: 53-63.
- Lombardi, O. y López, C. (2015), “Ergodicidad y Determinismo”, en: Claudia E. Vanney y Olimpia Lombardi (eds.) *Fronteras del determinismo científico. Filosofía y ciencias en diálogo*, Madrid: Editorial Biblioteca Nueva.
- Lloyd, S. (2001), “Measures of Complexity: A nonexhaustive list”, *IEEE Control Systems Magazine*, **21**: 7-8.
- Maynard Smith, J. (1988), “Evolutionary progress and levels of selection”, en: M. H. Nitecki (ed.), *Evolutionary Progress*, Chicago: University of Chicago Press.
- Maynard Smith, J. (2000), “The concept of information in biology”, *Philosophy of Science*, **67**: 177-194.

- McShea, D. W. (1991), "Complexity and evolution: what everybody knows", *Biology and Philosophy*, **6**: 33-324.
- McShea, D. W. (1996), "Metazoan complexity and evolution: Is there a trend?", *Evolution*, **50**: 477-492.
- McShea, D. W. (2001), "The minor transitions in hierarchical evolution and the question of a directional bias", *Journal of Evolutionary Biology*, **14**: 502-518.
- McKinney, M. (1998), "La escalera estadística de la evolución: el desarrollo embrionario como generador de complejidad", en: J. Wagensberg y J. Agustí (eds.), *El progreso: ¿un concepto acabado o emergente?* Barcelona: Tusquets, 107-130.
- Mitchell, M. (2009), *Complexity. A Guided Tour*, New York: Oxford University Press
- Mitchell, S. (2003), *Biological Complexity and Integrative Pluralism*, Cambridge: Cambridge University Press.
- Mosterín, J. (2001), *Ciencia Viva: Reflexiones sobre la aventura intelectual de nuestro tiempo*. Barcelona: Espasa Calpe.
- Mountcastle, V. B. (1997), "The Columnar Organization of the Neocortex", *Brain*, **120**: 701-722.
- Moss, L. (2002), *What Genes Can't Do*, Cambridge, MA: MIT Press.
- Moulines, C. U. (2011), *El Desarrollo Moderno de la Filosofía de la Ciencia (1890-2000)* (Xavier de Donato, Trad.), Universidad Nacional Autónoma de México. (Obra original publicada en 2008)

- Nicolis, G. y Prigogine, I. (1989), *Exploring Complexity. An Introduction*, New York: W. H. Freeman & Company.
- Nielsen, M. (2015), *Neural Networks and Deep Learning*, LibreTexts.
- Pierce, Ch. S. (1892), “The Doctrine of Necessity Examined”, *The Monist* **2**(3): 321-337.
- Poincaré, H. (1910), “Mathematical Creation”, *The Monist*, **20**: 321-335.
- Polanyi, Michael (2009), *The Tacit Dimension*, Chicago: The University of Chicago Press  
(Foreword by Amartya Sen) (Obra original publicada en 1966)
- Popper, R. Karl (2001), *Conocimiento Objetivo. Un enfoque evolucionista*. Madrid: Tecnos  
(4<sup>o</sup> edición).
- Prigogine, I. y Stengers, I. (1988) [1991], *Entre el tiempo y la eternidad*, Alianza Editorial,  
Buenos Aires.
- Quine, W. V. (1969), “Epistemology Naturalized”, *Ontological, Relativity and Other Essays*,  
New York: Columbia University Press.
- Rescher, N. (1998), *Complexity. A Philosophical Overview*, New Jersey: Transaction  
Publishers.
- Ruiz, R. y Ayala J. F. (2002), “Polémicas V. Microevolucion.”, en: *De Darwin al DNA y el origen de la humanidad: la evolución y sus polémicas*. Ediciones Científicas Universitarias.
- Rumelhart, D. et al. (1986), “Learning representations by back-propagation errors”, *Nature*  
**323**: 533-536

- Sanjuán, M. A (2007), “La Física al Encuentro de la Complejidad”, *ARBOR, Ciencia, Pensamiento y Cultura*, **728**: 889-898.
- Sarkar, S. (1996), “Biological information: a skeptical look at some central dogmas of molecular biology”, *The Philosophy and History of Molecular Biology: New Perspectives*, Dordrecht: Kluwer.
- Sarkar, S. (1998), *Genetics and Reductionism*, Cambridge: Cambridge University Press.
- Saunders, P. T. y Ho, M. W. (1976), “On the increase in complexity in evolution”, *Journal of Theoretical Biology*, **63**: 375-384.
- Senior, A. et al. (2020), “Improved protein structure prediction using potentials from deep learning”, *Nature*, **577**: 706-710.
- Shannon, C. (1948), “A Mathematical Theory of Communication”, *The Bell System Technical Journal*, **27**: 379-423
- Simon, H. (1962), “The Architecture of Complexity”, *Proceedings of the American Philosophical Society*, **106**, No. 6: 467-482.
- Slichter, C. (2007), “Superconductivity: So simple, yet so hard to explain!”, *Moment of Discovery* <https://history.aip.org/exhibits/mod/superconductivity/01.html>
- Stegmann, U. (2004), “The arbitrariness of the genetic code”, *Biology and Philosophy*, **19**: 205-222.
- Stotz, K. (2006), “With ‘genes’ like that, who needs an environment? Postgenomics's argument for the ‘ontogeny of information’”, *Philosophy of Science*, **73**: 905-917.

- Suárez, E. (2005), “La Biología Molecular: El Reto de Formular Explicaciones Reduccionistas”, en: Anna Estany (ed.), *Filosofía de las Ciencias Naturales, Sociales y Matemáticas*, Madrid: Editorial Trotta, 225-245.
- Taylor, M. C. (2001), *The Moment of Complexity. Emerging Network Culture*, Chicago: The University of Chicago Press.
- Thagard, Paul (2000), *Coherence in Thought and Action*. Massachusetts: A Bradford Book. MIT Press.
- Valentine, J. W. y May, C. L. (1996), “Hierarchies in biology and paleontology”, *Paleobiology*, **22**: 23-33.
- Vicsek, T. (2002), “The Bigger Picture”, *Nature*, **418**: 131.
- Vrba, E. S. y Gould, S. J. (1986), “The hierarchical expansion of sorting and selection: sorting and selection cannot be equated”, *Paleobiology*, **12**: 217-228.
- Wagensberg, J. (1998), “El progreso, ¿un concepto acabado o emergente?”, en: J. Wagensberg y J. Agustí (eds.), *El progreso: ¿un concepto acabado o emergente?* Barcelona: Tusquets, 107-130.
- Waldrop, M. (1992), *Complexity: The Emerging Science at the Edge of Order and Chaos*, New York: Simon & Schuster.
- Weiskopf, D. (2022), “The Predictive Turn in Neuroscience”, *Philosophy of Science* Forthcoming

Wimsatt, W. C. (1974), “Complexity and organization”, en: K. F. Schaffner and y R. S. Cohen (eds.), *Proceedings of the 1972 Meeting of the Philosophy of Science Association*, Dordrecht: Reidel, 67-86.

Winsberg, E. (2010), *Science in the Age of Computer Simulation*, Chicago: The University of Chicago Press.

Winsberg, E. (2013), “Computer Simulations in Science”, *The Stanford Encyclopedia of Philosophy* <https://plato.stanford.edu/archives/win2019/entries/simulations-science/>

Woodward, James (2003), *Making Things Happen. A theory of causal explanation*. New York: Oxford University Press.

Yablo, Stephen (2014), *Aboutness*. New Jersey: Princeton University Press.