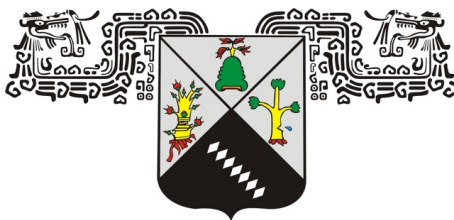


LIC. EN C. RAFAEL VERDUZCO VÁZQUEZ

MODELOS SEMICONTINUOS DE REDES DE
SEÑALIZACIÓN EN EL ESPERMATOZOIDE DE
ERIZO DE MAR



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

FACULTAD DE CIENCIAS
FACULTAD DE CIENCIAS QUÍMICAS E INGENIERÍA
CENTRO DE INVESTIGACIONES QUÍMICAS

MODELOS SEMICONTINUOS DE REDES DE SEÑALIZACIÓN EN
EL ESPERMATOZOIDE DE ERIZO DE MAR

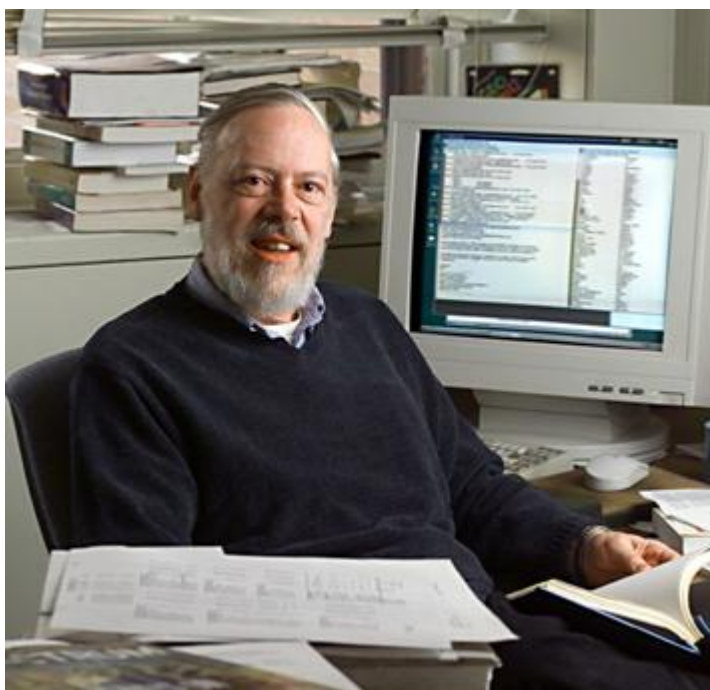
TESIS, QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS
MODELACIÓN COMPUTACIONAL Y CÓMPUTO CIENTÍFICO
PRESENTA

LIC. EN C. RAFAEL VERDUZCO VÁZQUEZ

ASESOR
DR. GUSTAVO MARTÍNEZ MEKLER

CUERNAVACA, MORELOS

MAYO DE 2012



Dedicado a la memoria de Dennis Ritchie

1941 – 2011

RESUMEN

Se presenta un modelo semicontinuo para una vía de señalización presente en el flagelo del espermatozoide de erizo de mar, que ha sido relacionada con la regulación de la motilidad del espermatozoide. El modelo semicontinuo fue construido a partir de un modelo discreto propuesto previamente por [Espinal et al. \[14\]](#), con el objetivo de realizar comparaciones cuantitativas entre comportamientos *in silico* e *in vivo*, a diferencia del modelo anterior, donde las comparaciones son a nivel cualitativo. Debido a la mayor complejidad de las reglas lógicas del modelo discreto antes mencionado en comparación con otros modelos discretos desarrollados para redes de regulación genética, se muestra un caso de estudio simple basado en la transformación de un modelo discreto compuesto de pocos nodos en uno basado en el formalismo semicontinuo aquí presentado. Si bien hay estudios que han sido exitosos en construir modelos semicontinuos a partir de modelos discretos, en el presente trabajo se requiere de una elección de parámetros más compleja para la cual es necesario contar con herramientas más elaboradas. A pesar de que en esta tesis no se logra establecer una metodología adecuada para esta transformación, se exploran pasos en esta dirección.

ORGANIZACIÓN DE LA TESIS

El capítulo 1 introduce la vía de señalización relacionada con la motilidad del espermatozoide de erizo de mar, un formalismo discreto para la modelación de vías de señalización, así como un modelo desarrollado bajo este formalismo por [Espinal et al. \[14\]](#). Tras discutir sus alcances y limitaciones, se introduce un formalismo semicontinuo que permite retomar modelos discretos y transformar el estado de los componentes de la red y el tiempo en variables continuas. Se discuten también algunos aspectos necesarios para la construcción de modelos basados en el formalismo semicontinuo. Finalmente, se presentan la justificación y objetivos del presente trabajo.

El capítulo 2 presenta los datos y técnicas usadas para la transformación de los modelos discretos en semicontinuos.

El capítulo 3 muestra los modelos semicontinuos que se obtuvieron a partir de modelos discretos y discute su validez de acuerdo a los comportamientos observados y el conocimiento biológico previo.

El capítulo 4 contiene las conclusiones del presente trabajo y perspectivas de trabajo a futuro.

El apéndice A describe el modelo de tres nodos usado como caso de estudio simple.

Finalmente, se presenta la bibliografía consultada para este proyecto.

*[C has] the power of assembly language,
and the convenience of... assembly language.*

— dmr, Dennis Ritchie.

AGRADECIMIENTOS

Gracias a todos, yadda yadda.

Este trabajo contó con una beca de maestría por parte de CONACYT con número 241439. Además, se contó con el apoyo del proyecto PAPIIT IN-109210 de la DGAPA-UNAM.

ÍNDICE GENERAL

1	INTRODUCCIÓN	1
1.1	La vía de señalización de speract	1
1.2	Modelos discretos de vías de señalización	4
1.3	Modelo discreto de la vía de señalización inducida por speract	5
1.4	Alcances y limitaciones del modelo discreto	9
1.5	Ecuaciones de Glass	10
1.5.1	Construcción de las ecuaciones de Glass	10
1.5.2	Ventajas y desventajas	12
1.6	Estimación de parámetros	12
1.6.1	Funciones de comparación	13
1.6.2	Estrategias de exploración	14
1.7	Justificación y objetivo	15
1.8	Resumen	17
2	MATERIALES Y MÉTODOS	19
2.1	Datos experimentales	19
2.2	Funciones objetivo	19
2.3	Estrategias de Búsqueda	20
2.4	Búsqueda de parámetros	22
2.5	Resumen	23
3	RESULTADOS	25
3.1	Modelo de tres nodos	25
3.2	Modelo de la red de señalización	26
3.3	Discusión de los Resultados	29
3.4	Resumen	31
4	CONCLUSIONES Y PERSPECTIVAS	33
4.1	Conclusiones	33
4.2	Perspectivas y trabajo futuro	34
A	MODELO AUXILIAR DE 3 NODOS	35
	BIBLIOGRAFÍA	37

ÍNDICE DE FIGURAS

Figura 1	Medición experimental de calcio intracelular . .	2
Figura 2	Red de señalización	3
Figura 3	Modelo Discreto de la red de señalización . . .	6
Figura 4	Applet del Modelo Discreto	7
Figura 5	El bloqueo del canal canal de Calcio dependien- te de alto voltaje (<i>HVA</i>) retrasa el inicio de las oscilaciones de calcio (Ca^{2+})	9
Figura 6	Dinámica original σ_z	26
Figura 7	Dinámicas original y estimada σ_z	27
Figura 8	Dinámica del nodo de calcio intracelular ($[Ca^{2+}]_i$) del modelo y medición experimental de Ca^{2+} . .	29
Figura 9	Dinámica de $[Ca^{2+}]_i$, potencial de membrana (V) y otros componentes del modelo, y medición experimental de Ca^{2+}	30
Figura 10	Modelo Booleano de 3 nodos	36

ABREVIATURAS

Ca^{2+} calcio

SR receptor de speract

GC guanilato ciclasa

cGMP guanosín monofosfato cíclico

KCNG canal de potasio regulado por cGMP

V potencial de membrana

NCE intercambiador sodio-calcio

$[Ca^{2+}]_i$ calcio intracelular

NHE intercambiador sodio-protones

pH_i pH intracelular

HCN canal activado por hiperpolarización y regulado por
nucleótidos cíclicos

HVA canal de Calcio dependiente de alto voltaje

LVA canal de Calcio dependiente de bajo voltaje

AC adenilato ciclasa

cAMP Adenosín Monofosfato Cíclico

cAMPCC canal de calcio dependiente de cAMP

CaCC canal de calcio dependiente de cloro

CaKC canal de calcio dependiente de potasio

CaP bombas de calcio

dK permeabilidad de potasio

PDE fosfodiesterasa

EDOs ecuaciones diferenciales ordinarias

INTRODUCCIÓN

Este capítulo inicia presentando una vía de señalización relacionada con la motilidad del espermatozoide de erizo de mar. Posteriormente presenta un modelo discreto para dicha vía de señalización, señalando sus alcances y limitaciones. Seguidamente se muestra un formalismo semicontinuo que promete llevar las comparaciones entre modelo y experimento a un nivel más cuantitativo, así como algunos detalles relevantes para la transformación de un modelo discreto en el formalismo semicontinuo antes mencionado.

1.1 LA VÍA DE SEÑALIZACIÓN DE SPERACT

La fertilización es un proceso crucial para la transmisión de la vida, que requiere el encuentro y fusión de los gametos. Para que este encuentro tenga lugar, el espermatozoide debe valerse de una intrincada maquinaria en su flagelo que le permita nadar en busca del óvulo. En algunas especies, el óvulo secreta una sustancia quimioatrayente que guía al espermatozoide hacia él. En el caso particular del erizo de mar, esta sustancia es un decapeptido llamado *speract*, [Suzuki \[52\]](#), [Hansbrough and Garbers \[28\]](#), el cual se une a un receptor específico en la membrana del flagelo del espermatozoide. La unión de *speract* con su receptor activa una red de señalización que produce oscilaciones en la concentración interna de algunos iones, de los cuales el principal es el Ca^{2+} [Darszon et al. \[12\]](#).

La vía de señalización inducida por *speract* provoca la apertura de distintos canales que hiperpolarizan la membrana, es decir, que disminuyen drásticamente la polaridad de la membrana [Kaupp et al. \[35\]](#). A su vez, esta hiperpolarización cancela la desactivación de canales regulados por voltaje, que al abrirse despolarizan la membrana [Strünker et al. \[50\]](#), [Granados-Gonzalez et al. \[24\]](#). La despolarización se refiere al proceso de regresar la polaridad de la membrana a sus niveles basales. El proceso repetido de hiper y despolarización se ha relacionado con cambios transitorios en la concentración de calcio intracelular que a su vez han sido relacionados con cambios en la curvatura del flagelo del espermatozoide que resultan en modificaciones abruptas de su trayectoria [Darszon et al. \[12\]](#). Estos cambios de trayectoria son una parte esencial para la motilidad y reorientación del espermatozoide. De ahí

la importancia de entender los mecanismos bioquímicos que la generan.

Las oscilaciones de $[Ca^{2+}]_i$ se caracterizan por presentar un incremento sostenido (*tónico*) y fluctuaciones superimpuestas (*supratónico*). Las oscilaciones tónicas corresponden al incremento sostenido en $[Ca^{2+}]_i$, y tienen una forma sigmoidal. Este tipo de oscilaciones de $[Ca^{2+}]_i$ se presentan tanto en la cabeza como en el flagelo del espermatozoide tras la activación de la vía de señalización mediante la adición de speract al medio. La figura 1 muestra una medición experimental de las oscilaciones, obtenida al añadir un marcador fluorescente al Ca^{2+} y liberar moléculas de speract en el medio. Los detalles experimentales se describen a profundidad en Darszon et al. [12] y Wood et al. [59]. En la figura, la serie en rojo corresponde a la concentración de $[Ca^{2+}]_i$ en la cabeza, mientras que los datos en color negro corresponden a la concentración de $[Ca^{2+}]_i$ en el flagelo, ambas medidas a través del tiempo. Las oscilaciones supratónicas se observan solamente en el flagelo y se distinguen por ser las fluctuaciones que parecen "montarse" sobre la curva con forma sigmoidal. La barra de color rojo en la parte inferior, que inicia poco después de la marca de los 2 segundos muestra el momento en el que el speract comienza a estimular la vía de señalización y el tiempo en la que ésta se mantiene activa.

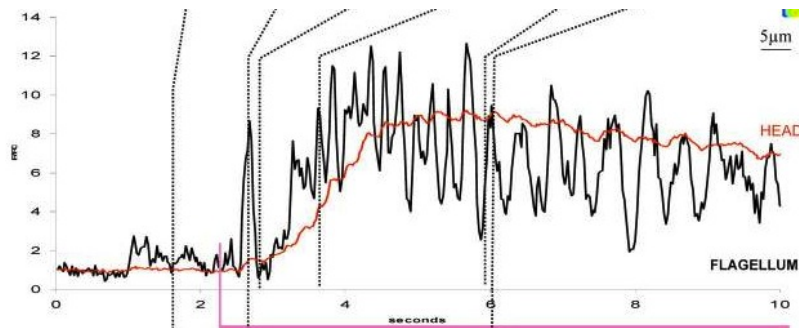


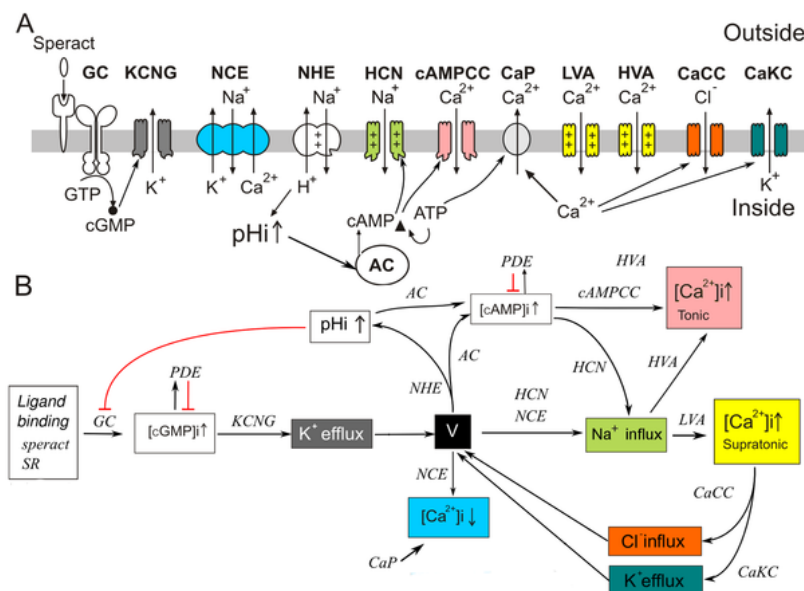
Figura 1: Medición experimental de fluorescencia de calcio intracelular en espermatozoides. Se muestran mediciones para la cabeza (rojo) y el flagelo (negro) del espermatozoide. Figura modificada tomada de Wood et al. [57].

La figura 2 A muestra el diagrama de la vía de señalización de speract, que consta del receptor de speract (SR), así como de varios canales e intercambiadores iónicos, entre otros componentes. La parte superior representa el exterior del flagelo, la parte gris es la membrana celular y debajo de esta se encuentra el interior de la célula. En esa misma figura, el apartado B muestra de manera esquemática los eventos de la vía de señalización, mismos que se detallan a continuación.

La vía de señalización inicia con la unión de speract a su receptor SR Darszon et al. [12], que interacciona con una guanilato ciclasa (GC) Garbers [19], la cual a su vez produce guanosín monofosfato cíclico (*cGMP*). El aumento en la concentración de *cGMP* abre el canal de

potasio regulado por cGMP (*KCNGB*) Galindo et al. [18]. La apertura de *KCNGB* resulta en la hiperpolarización del *v* Strünker et al. [50]. Como consecuencia de la hiperpolarización suceden cuatro eventos importantes:

1. se activa un intercambiador sodio-calcio (*NCE*) que disminuye los niveles de $[Ca^{2+}]_i$; Nishigaki et al. [43]
2. se activa un intercambiador sodio-protones (*NHE*) que incrementa el pH intracelular (pH_i) Lee and Garbers [38] Wang et al. [54], Su and Vacquier [51], Rodríguez and Darszon [47]
3. se activa un canal activado por hiperpolarización y regulado por nucleótidos cíclicos (*HCN*) Nishigaki et al. [43], Rodríguez and Darszon [47], Gauss et al. [20], Galindo et al. [17]
4. finalmente, la sustracción de la inactivación del *HVA* y el canal de Calcio dependiente de bajo voltaje (*LVA*) Strünker et al. [50], Granados-Gonzalez et al. [24], Wood et al. [57], Wood et al. [58], Darszon et al. [11], Perez-Reyes [45]



La figura 2 B muestra los eventos de la vía de señalización usando una notación estándar en los esquemas de regulación biológica: los arcos o aristas dirigidas de color negro que acaban en punta representan activación o regulación positiva, mientras que los arcos rojos que acaban en una línea llana representan inhibición o regulación negativa.

Figura 2: Red de señalización Espinal et al. [14].

A) Componentes principales involucrados en la vía de señalización de speract. La unión de speract a su receptor en la membrana del flagelo dispara la cascada que produce cambios en $[Ca^{2+}]_i$ a través de la entrada y salida de iones a través de intercambiadores, bombas y canales.

B) Esquema de eventos producidos por la vía de señalización.

El incremento del pH_i disminuye la actividad de *GC* a la vez que activa una adenilato ciclasa (*AC*), con la consecuente producción de Adenosín Monofosfato Cíclico (*cAMP*), Nishigaki et al. [43], Cook and Babcock [10], Beltrán et al. [6]. Este último estimula un canal de calcio

dependiente de cAMP (*cAMPCC*) y el canal *HCN* previamente activado, lo cual tiende a repolarizar el potencial de membrana *V* [Strünker et al. \[50\]](#), [Kaupp et al. \[34\]](#), [Nishigaki et al. \[43\]](#), [Rodríguez and Darszon \[47\]](#), [Gauss et al. \[20\]](#). Esta repolarización abre los canales *HVA* y *LVA* causando una despolarización y un incremento del $[Ca^{2+}]_i$, [Strünker et al. \[50\]](#), [Granados-Gonzalez et al. \[24\]](#), [Kaupp et al. \[34\]](#), [Böhmer et al. \[7\]](#), [Wood et al. \[57\]](#), [Nishigaki et al. \[43\]](#), [Wood et al. \[58\]](#). Finalmente, la vía de señalización se inicia de nuevo, posiblemente a través de un canal de calcio dependiente de cloro (*CaCC*) y un canal de calcio dependiente de potasio (*CaKC*) [Wood et al. \[57\]](#), [Wood et al. \[59\]](#), [Greenwood and Leblanc \[25\]](#), que se abren cuando la concentración de $[Ca^{2+}]_i$ es alta. Los mecanismos pasivos y constantes de extrusión de Ca^{2+} , tales como bombas de calcio (*CaP*) y *NCE*, mantienen los niveles basales de $[Ca^{2+}]_i$ [Nishigaki et al. \[43\]](#), [Su and Vacquier \[51\]](#), [Rodríguez and Darszon \[47\]](#), [Okunade et al. \[44\]](#). El mecanismo anteriormente descrito es repetido cíclicamente, generando un tren de oscilaciones de Ca^{2+} que produce una secuencia repetitiva de cambios de dirección en el espermatozoide [Guerrero et al., \[26\]](#), [Guerrero et al. \[27\]](#), [Böhmer et al. \[7\]](#), [Wood et al. \[58\]](#), [Wood et al. \[59\]](#). En un experimento típico, como el mostrado en la figura 1, la vía se mantiene activa un promedio de 10 segundos [Wood et al. \[57\]](#).

1.2 MODELOS DISCRETOS DE VÍAS DE SEÑALIZACIÓN

Una red Booleana se define como un conjunto de nodos conectados entre sí. En una red de n nodos, cada nodo i puede tener k_i aristas adyacentes, donde $k_i \leq n$. Los nodos con aristas adyacentes al nodo i se denominan *reguladores* del nodo i . En cada instante de tiempo t , cada nodo i tiene asociado un valor del conjunto $\mathbb{B} = \{\text{verdadero}, \text{falso}\}$. Este valor se denomina *estado del nodo*. La función $\sigma_i(t)$ devuelve el estado del nodo i al tiempo t . El estado de cada nodo cambia a través de intervalos discretos en el tiempo mediante el mapeo discreto

$$\sigma_i(t+1) = F_i[\sigma_{i_1}(t), \sigma_{i_2}(t), \dots, \sigma_{i_k}(t)] \quad (1)$$

donde F_i es una función Booleana definida como $F_i : \mathbb{B}^{k_i} \rightarrow \mathbb{B}$, es decir, el estado del nodo i al tiempo t depende de sus reguladores, [Kauffman \[33\]](#), [Gershenson \[21\]](#).

Las redes Booleanas fueron propuestas como modelos para entender la dinámica de redes de regulación genética por Stuart Kauffman [Kauffman \[33\]](#), haciendo corresponder a cada nodo de la red con un componente de una vía de señalización bioquímica o red regulatoria genética.

Si a un tiempo t un nodo se encuentra en el estado *falso*, se considera que éste tiene actividad basal o inactividad, mientras que un estado

verdadero representa actividad o expresión. En general, se pueden establecer distintos niveles de expresión o actividad si se considera que los nodos puedan tomar valores discretos en lugar de valores Booleanos, i.e. 0, 1, -1, 2, etc. en vez de *verdadero* o *falso*.

Al iterar la regla de evolución para cada nodo de la red se obtiene una descripción por pulsos de la dinámica del sistema. Cada pulso puede ser considerado como el promedio discretizado de la variable de estado de cada nodo de la red durante un intervalo de tiempo dado. Si bien el $(t + 1)$ en la función de evolución se refiere a un tiempo de máquina o tiempo de simulación, es posible hacer comparaciones entre este tiempo de máquina y el tiempo de los experimentos.

Uno de los aspectos fundamentales de este tipo de modelos discretos es que cualquier configuración inicial posible de estados del sistema llega a un conjunto de configuraciones que se repiten a lo largo del tiempo tras iterar un cierto número de veces la reglas de evolución de la red. Estas configuraciones repetidas pueden ser puntuales, es decir, la misma configuración se repite *ad eternum*; o bien cíclicas, es decir, el sistema vuelve a la misma configuración después de un cierto número de pasos de tiempo. En ambos casos, el conjunto de los estados que se repiten una y otra vez, se conoce como atractor. El conjunto de todas las configuraciones iniciales que tras iterar la regla de evolución del sistema alcanzan un atractor, se conoce como cuenca de atracción. Las configuraciones que se encuentran entre una condición inicial y un atractor, se conocen como transitorios [Bornholdt \[8\]](#) [Gershenson \[21\]](#).

Este tipo de modelos han sido utilizados para modelar redes de regulación genética dando una descripción cualitativa del sistema, como se muestra en [Espinosa-Soto et al. \[15\]](#), [Albert and Othmer \[2\]](#). Además, recientemente se han usado para modelar vías de señalización [Morris et al. \[42\]](#).

[Huang et al. \[31\]](#) muestran que los atractores dinámicos de un modelo discreto de red de regulación genética, corresponden a patrones estables de expresión genética que determinan estados funcionales estables de una célula.

1.3 MODELO DISCRETO DE LA VÍA DE SEÑALIZACIÓN INDUCIDA POR SPERACT

[Espinal et al. \[14\]](#) presentan un modelo discreto en tiempo y estado para la vía de señalización inducida por speract en el flagelo del espermatozoide de erizo de mar. En ese trabajo la red se compone de veintidós nodos, de los cuales dieciocho toman valores de estado 0 ó 1, mientras que los cuatro nodos restantes tienen un valor de estado ternario, es decir, cada uno de estos nodos puede encontrarse en el estado 0, 1 ó 2. Los nodos ternarios son los correspondientes al potencial de membrana (V), calcio (Ca^{2+}), canal de Calcio dependiente de alto voltaje (HVA) y canal de Calcio dependiente de bajo voltaje (LVA). Esta

extensión a tres valores posibles es necesaria para capturar los estados posibles en los que se puede encontrar un componente de esta red de señalización. La figura 3 muestra un diagrama de las conexiones de la misma vía de señalización presentada en la sección 1.1 e ilustrada en la figura 2. Los cuadros amarillos y verdes indican nodos binarios y ternarios, respectivamente. Las flechas negras indican activación, las líneas rojas inhibición y las flechas amarillas pueden representar activación o inhibición, dependiendo del valor del nodo correspondiente al potencial de membrana (V). Los números sobre las flechas corresponden a las referencias con las que Espinal et al. [14] construyeron la red y sustentan cada interacción. A manera de ejemplo, la función reguladora o tabla de verdad del nodo de $cAMP$ se muestra en la esquina inferior izquierda. Las primeras 3 columnas en esta tabla contienen todos los valores posibles de activación de los reguladores de $cAMP$: AC , fosfodiesterasa (PDE) y $cAMP$; la cuarta columna muestra los valores para la función que corresponden a cada combinación de los reguladores.

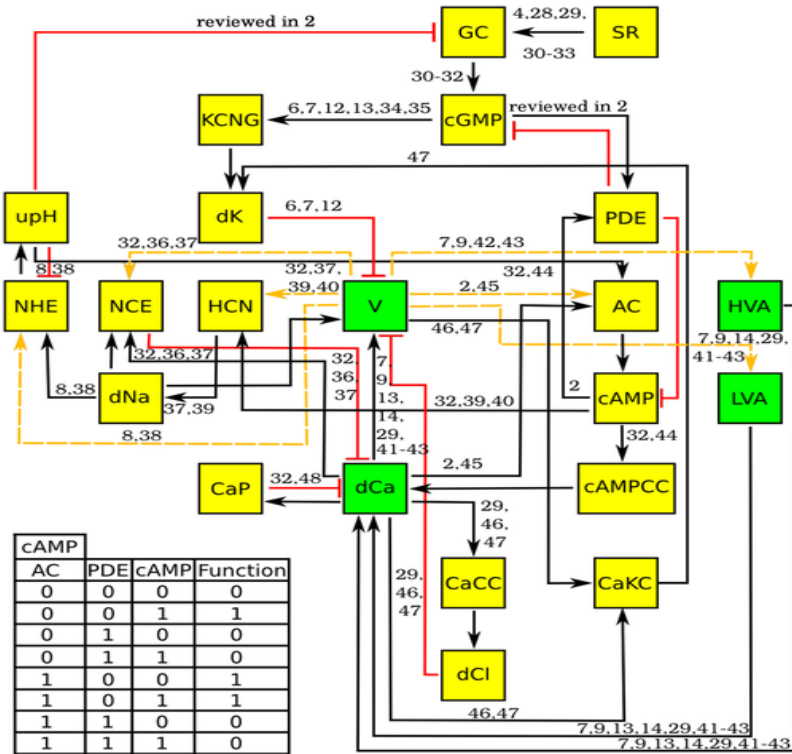


Figura 3: Modelo Discreto de la red de señalización Espinal et al. [14]. La figura muestra el diagrama de conexiones de la vía de señalización. Las flechas negras y rojas representan activación e inhibición, respectivamente. Las flechas naranjas son de activación o inhibición en función del valor del potencial de membrana (V). Los números sobre las flechas corresponden a las referencias con las que Espinal et al. [14] construyeron la red y sustentan cada interacción. En la parte inferior izquierda se muestra la tabla de la función de regulación para $cAMP$.

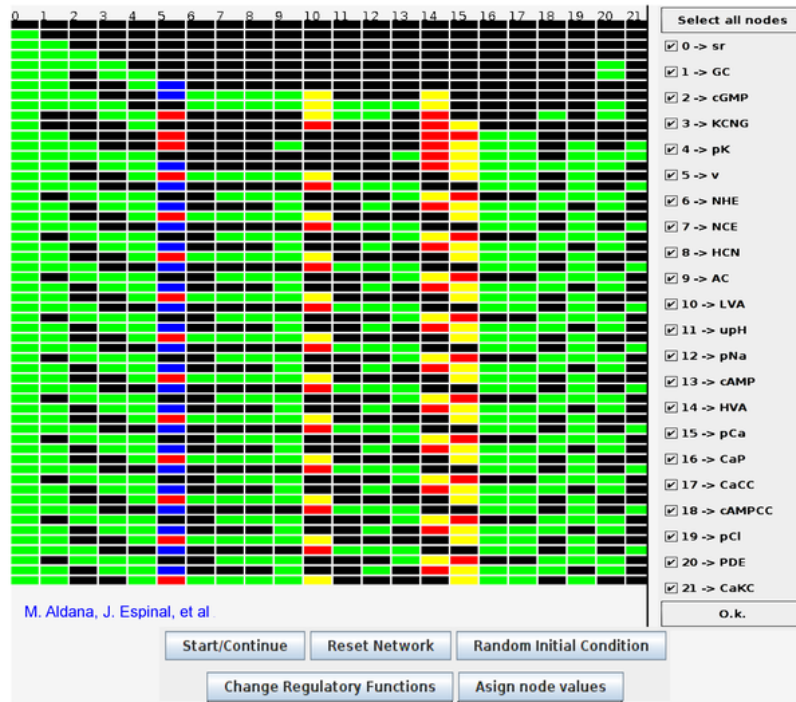


Figura 4: Applet del Modelo Discreto de la red de señalización Espinal et al. [14]. Serie temporal de los patrones de activación de la red de señalización. Los nodos binarios se presentan en negro y verde, correspondientes a *apagado*(0) y *encendido* (1). Los nodos ternarios correspondientes a los canales *HVA* y *LVA*, el negro es un estado inactivo (0), el amarillo (1) corresponde a un canal cerrado y el rojo (2) a un canal abierto. El nodo correspondiente al potencial de membrana (*v*) es negro (0) para un potencial en reposo, azul para la hiperpolarización (1) y rojo (2) para la repolarización. El nodo correspondiente al $[Ca^{2+}]_i$ es amarillo (1) para indicar el incremento tónico, rojo (2) para el incremento supratónico y negro (0) para indicar estado basal. Se muestra el comportamiento del espécimen silvestre o *wild type*. El applet fue desarrollado por el Dr. Maximino Aldana y se encuentra disponible en <http://www.fis.unam.mx/research/seaurchin/discrete/>.

Para explorar distintas configuraciones iniciales, cambiar las reglas lógicas así como observar el comportamiento de la red en ausencia de algunos nodos, Espinal et al. [14] presentan un *applet* que permite explorar estas condiciones de manera interactiva. La figura 4 muestra una captura de pantalla de esta aplicación, donde se simula un organismo silvestre, es decir, una simulación en la que no se han realizado alteraciones a las reglas lógicas ni se han borrado nodos. Los nodos se presentan en el eje horizontal y el tiempo corre de arriba hacia abajo. Para los nodos binarios, el negro representa el estado *apagado*, verde *encendido*. Los nodos 10 y 14, correspondientes al *HVA* y *LVA*, el negro es un estado inactivo, el amarillo corresponde a un canal cerrado y el rojo a un canal abierto. El nodo 5, correspondiente al potencial de membrana (*v*) es negro para un potencial en reposo, azul para la

hiperpolarización y rojo para la repolarización. El nodo 15 (dCa), correspondiente al $[Ca^{2+}]_i$, es amarillo para indicar el incremento tónico, rojo para el incremento supratónico y negro para indicar estado basal. Como puede observarse, bajo estas condiciones y para la condición inicial mostrada en la figura, tras un transitorio la red llega a un atractor de período 4.

Espinal et al. [14] determinan que las configuraciones iniciales posibles para este modelo llegan a una de dos configuraciones cíclicas o atractores, uno de período cuatro y otro de período ocho. Estos atractores se relacionaron con las mediciones experimentales de fluorescencia de Ca^{2+} en el flagelo del espermatozoide. Bajo el criterio de Huang et al. [31] mencionado en la sección anterior, los atractores del modelo discreto de la red de señalización de speract determinan las oscilaciones estables de Ca^{2+} que posibilitan la relocalización de los espermatozoides a través de las alteraciones que el Ca^{2+} induce a la curvatura del flagelo. El 80% de las configuraciones iniciales posibles convergen al atractor de período cuatro. El período cuatro de este atractor corresponde a una oscilación del tren de oscilaciones que se observa en las mediciones experimentales como la de la figura 1 Espinal et al. [14].

Un paso fundamental al proponer un modelo para un sistema biológico consiste en comparar los resultados de dicho modelo con comportamientos que se hayan observado previamente mediante técnicas experimentales. A este proceso se le conoce como validación del modelo. El modelo presentado en Espinal et al. [14] corrobora en un nivel cualitativo comportamientos observados experimentalmente, por ejemplo:

- En ausencia de speract, todas las condiciones iniciales alcanzan las condiciones basales tras un pequeño transitorio. De esta manera se confirma el papel que juega el speract para disparar las oscilaciones de $[Ca^{2+}]_i$ Espinal et al. [14].
- Si el nodo correspondiente a la permeabilidad de potasio (dK) se elimina, las oscilaciones de $[Ca^{2+}]_i$ se suprimen. Este comportamiento se ha observado experimentalmente en Babcock et al. [5].

El modelo de la vía de señalización de speract predijo dos comportamientos que no se habían observado experimentalmente. En particular, uno de ellos consiste en que al suprimir el canal HVA , aumenta el tiempo para que la concentración de Ca^{2+} alcance los valores de alta intensidad que provoca la unión del speract a su receptor. Esta condición fue el comportamiento observado en promedio para 10^5 condiciones iniciales diferentes. A raíz de este resultado del modelo se realizaron experimentos añadiendo $10nm$ de *verapamil*, un inhibidor del canal HVA . El bloqueo de este canal produce un retraso en la aparición de la primera fluctuación de $[Ca^{2+}]_i$ en el flagelo del espermatozoide,

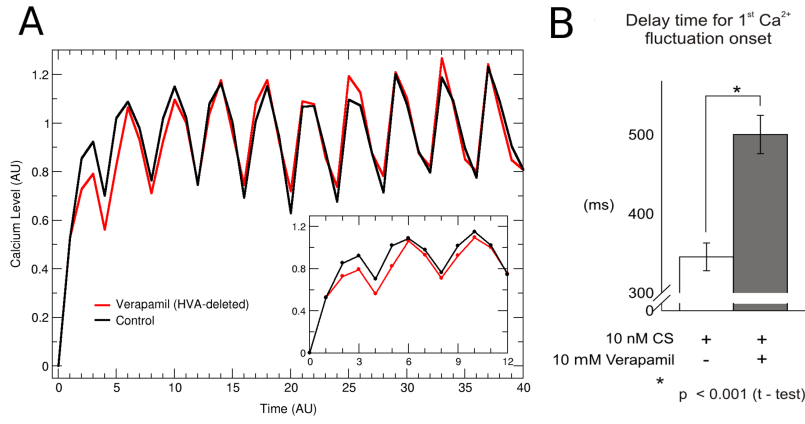


Figura 5: El bloqueo del canal *HVA* retrasa el inicio de las oscilaciones de Ca^{2+} . A) Evolución temporal del nivel promedio de Ca^{2+} tomado a cada paso de tiempo durante 10^5 condiciones iniciales con *HVA* (curva negra) y sin *HVA* (curva roja), usando unidades arbitrarias. Nótese que cuando el canal *HVA* está presente el incremento en los niveles de Ca^{2+} comienza más rápidamente que cuando se suprime el canal *HVA*, es decir, en los primeros pasos de tiempo, la curva negra se incrementa antes que la curva roja. La gráfica más pequeña es un acercamiento a los primeros pasos de tiempo de la dinámica. B) Los experimentos realizados muestran que el verapamil, sustancia que inhibe al canal *HVA*, prolonga el tiempo entre la estimulación de speract y el establecimiento de la primera fluctuación de Ca^{2+} . Espinal et al. [14].

lo cual valida a un nivel cualitativo la predicción hecha por el modelo. Este resultado se muestra en la figura 5

Como se puede observar en el apartado B de la figura 5, el tiempo medio para la aparición de la primera oscilación de Ca^{2+} tras la adición de speract al medio es de alrededor de 500ms en presencia de verapamil, en contraste con los 350ms del tipo silvestre.

1.4 ALCANCES Y LIMITACIONES DEL MODELO DISCRETO

La figura 4 muestra que el modelo discreto divide una oscilación de Ca^{2+} en cuatro subintervalos discretos en el tiempo. El valor discreto de un nodo cualquiera de la red en un subintervalo de tiempo, puede interpretarse como el promedio de los valores continuos que tomó ese nodo en el subintervalo: si dicho valor promedio fue mayor que un umbral, el estado discreto en todo el subintervalo es 1, y 0 en caso contrario. Lo que se quiere poner de manifiesto con esta observación es que las comparaciones entre un modelo discreto y las mediciones experimentales son a un nivel cualitativo y con una resolución temporal lo suficientemente grande como para contener en un paso de tiempo discreto varios cambios de estado continuo. Un modelo discreto es capaz de describir cómo cambia un sistema en una escala de tiempo más grande que aquella en la que suceden los eventos microscópicos

del sistema. De ahí la necesidad de contar con otro tipo de descripciones del sistema que permitan hacer comparaciones entre modelo y experimentos en escalas de tiempo más pequeñas. Además, es deseable que dichas descripciones permitan hacer una comparación más precisa entre los valores de las variables de estado y las mediciones experimentales, por ejemplo, donde las variables de estado del modelo sean continuas.

Una de las limitaciones de dicho modelo consiste en que el sistema es sincronizado, es decir, todos los nodos actualizan su estado con el mismo esquema temporal. Esta actualización sincronizada es particularmente válida en tanto que el estado del nodo es un promedio en un intervalo de tiempo cuya duración es mayor que el tiempo en el que se dan las reacciones bioquímicas de la red. Sin embargo, aunque válido en algunos casos, el describir al sistema en intervalos de actualización largos es también una desventaja, puesto que no toma en cuenta la variedad de escalas temporales que presentan los distintos tipos de procesos biológicos, [Saadatpour et al. \[48\]](#). Como consecuencia de la escala temporal artificial de los modelos sincronizados, las comparaciones de éstos con experimentos solo pueden darse a un nivel cualitativo. Finalmente, la discretización de las variables de estado es adecuada para tener una idea general –no particular– de la dinámica del sistema.

1.5 ECUACIONES DE GLASS

Una opción de modelación intermedia entre los modelos discretos y los basados en las ecuaciones diferenciales ordinarias (EDOs) son los de dinámica semicontinua. Entre este tipo de modelos se encuentran las ecuaciones lineales por pedazos que fueron esbozadas por Glass en [Glass and Kauffman \[22\]](#), consistentes en definir una derivada cuya forma específica depende del valor de un mapeo discreto en un intervalo de tiempo pequeño. En particular este mapeo puede ser una red Booleana como las presentadas en la sección 1.2.

1.5.1 Construcción de las ecuaciones de Glass

Consideremos de nuevo el mapeo discreto 1. Es posible escribirlo también como

$$\sigma_i(t + \tau_i) = F_i[\sigma_{i_1}(t), \sigma_{i_2}(t), \dots, \sigma_{i_k}(t)] \quad (2)$$

donde $\tau_i = 1$. Si $\tau_i \rightarrow 0$, haciendo un desarrollo en serie de Taylor

de $\sigma_i(t + \tau_i)$ hasta la derivada de primer orden, para posteriormente despejar esa derivada se obtiene

$$\frac{d\sigma_i(t)}{dt} = \frac{1}{\tau_i}[\sigma_i(t + \tau_i) - \sigma_i(t)] \quad (3)$$

y sustituyendo $\sigma_i(t + \tau_i)$ por

$$F_i[\hat{\sigma}_{i_1}, \hat{\sigma}_{i_2}, \dots, \hat{\sigma}_{i_k}] \quad (4)$$

donde

$$\hat{\sigma}_{i_j} = H(\sigma_{i_j}(t), \theta_{i_j}) \quad 1 \leq j \leq k \quad (5)$$

obtenemos

$$\frac{d\sigma_i(t)}{dt} = \frac{1}{\tau_i}(F_i[\hat{\sigma}_{i_1}, \hat{\sigma}_{i_2}, \dots, \hat{\sigma}_{i_k}] - \sigma_i(t)) \quad (6)$$

donde $\frac{1}{\tau_i}$ es el inverso del tamaño del intervalo donde se hace el desarrollo en serie de Taylor y $H(\sigma_{i_j}, \theta_{i_j})$ es una función escalón que discretiza los valores de estado del nodo i_j usando un valor de umbral θ_{i_j} . Esta discretización es necesaria puesto que las variables de estado σ_i son ahora variables continuas, al igual que el tiempo. Todos los F_i son los mapeos del modelo discreto a partir del cual se construyeron las ecuaciones de Glass.

Mediante este procedimiento, es posible contar con una descripción semicontinua de la dinámica del sistema. La derivada de cada nodo no es única sino que depende del valor de la función discreta en cada intervalo de tamaño τ_i . La derivada para un nodo Booleano es de la forma

$$\frac{d\sigma_i(t)}{dt} = \begin{cases} \frac{1}{\tau_i}(1 - \sigma_i(t)) & \text{si } F_i[\hat{\sigma}_{i_1}, \hat{\sigma}_{i_2}, \dots, \hat{\sigma}_{i_k}] = 1 \\ \frac{1}{\tau_i}(0 - \sigma_i(t)) & \text{si } F_i[\hat{\sigma}_{i_1}, \hat{\sigma}_{i_2}, \dots, \hat{\sigma}_{i_k}] = 0 \end{cases} \quad (7)$$

A este tipo de ecuaciones se les conoce también como *Piecewise Linear Differential Equations*, o Ecuaciones Diferenciales Lineales por Pedazos.

1.5.2 Ventajas y desventajas

El beneficio directo de usar el formalismo de las ecuaciones de Glass es contar con un refinamiento de las descripciones temporal y de estado que es capaz de producir un modelo discreto. Además, como se acaba de mostrar en su planteamiento, por construcción las ecuaciones de Glass incluyen un modelo discreto que a su vez fue planteado a partir de observaciones y conocimiento biológico. De este modo, estas ecuaciones semicontinuas no son un formalismo *ad hoc* que reproduce el comportamiento de un sistema a través de ajustar arbitrariamente trayectorias, sino que se basan en un conjunto de observaciones y conocimiento previo.

El costo por usar este formalismo radica en la estimación de los valores de umbral para cada nodo de la red. En el caso de que se quiera un modelo sincronizado, los inversos de los tiempos característicos, que denotamos como $\alpha_i = \frac{1}{\tau_i}$ pueden establecerse todos iguales. En particular, el modelo discreto en tiempo se recupera al establecer el tiempo característico en $\tau_i = 1, \forall i$.

En este sentido, el problema de construir un modelo semicontinuo usando ecuaciones de Glass se reduce a un problema de estimación de parámetros. Estos parámetros son los umbrales de activación de cada nodo, además de los inversos del tamaño de cada uno de los intervalos de tiempo. Los parámetros deben elegirse de modo que la dinámica del sistema y de los nodos en particular sea coherente desde un punto de vista biológico.

1.6 ESTIMACIÓN DE PARÁMETROS

La estimación de parámetros es un problema de optimización, consistente en encontrar parámetros tales que un modelo que dependa de ellos, minimice el error entre la solución al modelo y mediciones experimentales, es decir, que se *ajuste* o *explique* un conjunto de datos experimentales.

Para resolver un problema de optimización, es necesario definir una función, conocida como función objetivo, función de adaptación o función de costo, que permita asignar un valor a cada solución factible. Dependiendo de la manera en que esté definida, puede buscarse minimizar o maximizar la función objetivo. Los problemas de maximización y minimización son equivalentes, ya que minimizar una función f es lo mismo que maximizar $h = -f$. Además de plantear una función para minimizar o maximizar, resulta útil contar con una estrategia de exploración del espacio de soluciones.

De manera formal, sea $f(\mathbf{p})$ la función de costo a minimizar (o función de adaptación a maximizar), donde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y \mathbf{p} es un vector de números reales. El valor de la función $f(\mathbf{p})$ indica la adaptación o costo del vector \mathbf{p} . Típicamente, en muchos problemas el gradiente

de f es desconocido. El objetivo consiste en encontrar $\mathbf{m} \in \mathbb{R}^n$ tal que $f(\mathbf{m}) \leq f(\mathbf{p})$, $\forall \mathbf{p} \in \mathbb{R}^n$. En el caso de maximización, basta con definir la función $h = -f$. En este trabajo se optó por la minimización de funciones objetivo.

1.6.1 Funciones de comparación

Para el tipo de problemas de optimización discutidos en esta tesis, las funciones objetivo que se planteen, deben de permitir establecer una noción de distancia para determinar el grado de similitud entre una serie temporal proveniente de datos experimentales, y una serie temporal de la solución a un problema de condiciones iniciales de la ecuación de Glass correspondiente a dicha medición experimental.

Existen varias maneras de comparar series de tiempo, desde técnicas generales y ampliamente usadas como el *Error Cuadrático Medio* (MSE) [55], el *Coeficiente de Correlación de Pearson* [56], hasta aquellas que fueron diseñadas con el objetivo de establecer la similitud entre series de tiempo surgidas de procesos biológicos, como el *Índice de Pendiente o Slope Index* (SI) Cho et al. [9].

1.6.1.1 Error Cuadrático Medio (MSE)

El MSE [55] es el cuadrado del valor esperado de la diferencia entre las observaciones y la respuesta, trayectoria o dinámica predicha por un modelo. Al ser minimizado, es posible discriminar entre distintos modelos, estableciendo cuál de entre un conjunto de modelos ajusta mejor, es decir, explica mejor los datos. El MSE está dado por

$$E[(Y - X)^2] \quad (8)$$

donde X son los datos experimentales, Y son los datos del modelo que se quiere ajustar a dichos datos y E es el valor esperado de la diferencia entre los conjuntos de datos X y Y , es decir

$$E = \frac{1}{n}(Y_i - X_i)^2 \quad (9)$$

donde X_i y Y_i denotan las i ésimas mediciones de X y Y , respectivamente y n es el número de datos.

1.6.1.2 Coeficiente de correlación de Pearson

El Coeficiente de Correlación de Pearson [56] es una medida de la correlación (dependencia lineal) entre dos variables aleatorias X y Y , denotado por $r \in [-1, 1]$, está dado por

$$r_{X,Y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (10)$$

donde \bar{X} y \bar{Y} son los promedios de X y Y , respectivamente, X_i y Y_i denotan las i ésimas mediciones de X y Y , respectivamente y n es el número de datos. En este trabajo se identifica a los datos experimentales como X y a Y como los datos provenientes de la solución de un modelo (es decir, el modelo que se desea hacer coincidir con esos datos experimentales).

1.6.1.3 Índice de Pendiente (SI)

Propuesto por [Cho et al. \[9\]](#), el índice de pendiente entre dos series de tiempo X y Y consiste en ponderar la suma de los signos de las derivadas entre los puntos de ambas series. Está dado por

$$SI(X, Y) = \frac{1}{n-1} \sum_{i=1}^{n-1} \text{signo}\left(\frac{X_{i+1} - X_i}{Y_{i+1} - Y_i}\right) \quad (11)$$

donde X_i y Y_i denotan las i ésimas mediciones experimental de X y Y , respectivamente, n es el total de mediciones experimentales y signo es una función definida como $\text{signo} : \mathbb{R} \rightarrow \{-1, 0, 1\}$ dada por

$$\text{signo}(a) = \begin{cases} -1 & \text{si } a < 0, \\ 0 & \text{si } a = 0, \\ 1 & \text{si } a > 0. \end{cases}$$

Se excluyen aquellos puntos en los cuales $Y_{i+1} - Y_i = 0$. De manera similar al coeficiente de correlación de Pearson, $SI(X, Y) \in [-1, 1]$, X representa a los datos experimentales y Y los datos provenientes de la solución de un modelo (es decir, el modelo que se desea hacer coincidir con esos datos experimentales).

1.6.2 Estrategias de exploración

Una vez que se tiene definida una función objetivo o función de costo que permita evaluar las distintas soluciones factibles al problema, es necesario contar con una estrategia de búsqueda o exploración que permita explorar el espacio de soluciones factibles. Además, es deseable que la estrategia de búsqueda explore de manera inteligente el espacio de soluciones, es decir, que evite que la búsqueda quede atrapada en puntos extremos locales y que por ende sea imposible encontrar los puntos extremos globales.

1.6.2.1 Tipos de exploración

Además de la búsqueda al azar, existen dos estrategias generales para resolver problemas de optimización: aquellas basadas en la diferenciabilidad de la función de costo, y aquellas basadas en el uso de

criterios heurísticos. Un ejemplo clásico de un método basado en diferenciabilidad es el de Gradiente Conjugado, [Press et al. \[46\]](#). Entre los métodos populares basados en criterios heurísticos se encuentran el RECOCIDO SIMULADO, [Kirkpatrick et al. \[37\]](#), y los ALGORITMOS GENÉTICOS, [Goldberg \[23\]](#).

Los métodos basados en diferenciabilidad suelen usarse cuando se conoce con un buen nivel de detalle el espacio de soluciones, y en ese caso permiten encontrar los puntos extremos globales fácilmente. Sin embargo, una mala elección del punto inicial de búsqueda puede resultar en que estos algoritmos de optimización queden atrapados en un punto extremo local. Hacer malas elecciones del punto inicial de búsqueda suele ser común cuando no se conoce muy a fondo el espacio de soluciones. Por ello, estos métodos son útiles para refinar soluciones, es decir, para aumentar la precisión de una solución una vez que se cree que se está cerca de un punto extremo global, o al menos, cerca de un punto cuyo valor es cercano al punto extremo global.

Por otro lado, los métodos de búsqueda basados en criterios heurísticos explotan propiedades del problema que no necesariamente están relacionadas con la diferenciabilidad de la función objetivo o función de costo. Además, se ha observado que son más resistentes a quedar atrapados en puntos extremos locales, debido a que incluyen un subprocedimiento que permite variar de una u otra forma las soluciones propuestas, independientemente de si esta solución mejora o no el costo de la función objetivo. Por lo tanto, estos métodos no son tan sensibles a una mala elección del punto o puntos iniciales de búsqueda, y no es necesario conocer la forma específica del espacio de soluciones, [Moles et al. \[41\]](#), [Storn and Price \[49\]](#).

1.7 JUSTIFICACIÓN Y OBJETIVO

Tal y como se mostró en el experimento de bloqueo de [HVA](#), el tipo de comparaciones que se pueden hacer entre el modelo discreto y los experimentos son a un nivel cualitativo. A pesar de que aún a este nivel es posible ganar entendimiento de la vía de señalización, el contar con un modelo más detallado en tiempo y estado sería útil para hacer comparaciones a nivel cuantitativo. Adicionalmente, hay que considerar que al hacer más pequeño el tiempo en el que se actualiza el estado de un componente de la red, la asincronía, es decir, que cada componente o nodo de la red actualice su estado de acuerdo a su propio tiempo característico de reacción, se vuelve más relevante, [Saadatpour et al. \[48\]](#).

Una opción para refinar el modelo es plantear ecuaciones diferenciales ordinarias. Sin embargo, para esto se requiere conocer un conjunto de parámetros relacionados directamente con magnitudes físicas y bioquímicas, a saber: las concentraciones de cada componente de la vía de señalización; las tasas de asociación y disociación de cada

complejo bioquímico, y finalmente, la cooperatividad de las reacciones bioquímicas, es decir, la cantidad de moléculas necesarias de un compuesto para que este reaccione con otro. Por lo general, estos parámetros no son conocidos en su totalidad y su estimación experimental es costosa, complicada desde el punto de vista de diseño y ejecución del experimento, o ambas. En el caso particular de esta red de señalización, no todos estos parámetros e información son conocidos, por lo que como primera aproximación se desarrolló el modelo discreto antes mencionado.

Un modelo basado en las ecuaciones de Glass es un tipo de descripción del sistema a un nivel de detalle que se encuentra entre una aproximación discreta y una continua como la proporcionada por un modelo basado en ecuaciones diferenciales ordinarias.

Es importante recalcar que las ecuaciones de Glass no son en modo alguno una alternativa excluyente del modelado que se pueda hacer con ecuaciones diferenciales ordinarias. Por el contrario, si bien es cierto que los parámetros temporales y de umbral de las ecuaciones de Glass no son los mismos parámetros cinéticos y de reacción que aparecen en los modelos basados en EDOs, conocer los parámetros de las ecuaciones de Glass puede dar una idea cualitativa de la magnitud de los parámetros cinéticos. Este conocimiento cualitativo resultaría de gran utilidad al momento de querer plantear un sistema de ecuaciones diferenciales acopladas para la vía de señalización de speract.

El objetivo de este trabajo es mejorar, mediante el uso de ecuaciones de Glass, la descripción temporal y de estado de la vía de señalización que la proporcionada por el modelo discreto de Espinal et al. [14].

Algunas de las funciones de regulación del modelo de Espinal et al. [14], en particular las correspondientes a los nodos de Ca^{2+} y potencial de membrana, son bastante más complejas que las que aparecen en otros modelos discretos, por ejemplo el de diferenciación de la flor de la planta modelo *Arabidopsis thaliana* que presentan Espinosa-Soto et al. [15] y Alvarez-Buylla et al. [3], o el de diferenciación en el nicho de células madre de la misma planta desarrollado en Azpeitia et al. [4]. En estos modelos, todos los nodos y por ende las funciones de regulación son estrictamente binarios, en contraste con los valores ternarios de los nodos de Ca^{2+} y potencial de membrana y algunos de los reguladores de éstos en el modelo de Espinal et al.. Por ello, se determinó que sería conveniente ganar entendimiento del proceso de traducción de un modelo discreto en uno basado en ecuaciones de Glass, ayudándose de un modelo discreto con una dinámica más sencilla. En particular, dentro de Saadatpour et al. [48] se presenta una red discreta pequeña de tres nodos, donde los valores de estado son estrictamente Booleanos. La elección de esta red se basa solo en su simplicidad, a la par de que no es una red artificial sino una subred de una vía de señalización en plantas. Los modelos que se presentan en Espinosa-Soto et al. [15], Alvarez-Buylla et al. [3] y Azpeitia et al. [4] son bastante más

complejos que esa red de tres nodos. La red de tres nodos se describe brevemente en el apéndice [A](#).

1.8 RESUMEN

Este capítulo introdujo la vía de señalización de speract, que se ha relacionado con la motilidad del espermatozoide de erizo de mar. Mostró también un modelo discreto para esta vía de señalización, resaltando sus alcances y limitaciones. Introdujo el formalismo semicontinuo de las ecuaciones de Glass y mostró algunos aspectos relevantes para la construcción de un modelo de este tipo a partir de uno discreto. Además, se mencionó que el objetivo de esta tesis es desarrollar un modelo de ecuaciones de Glass para la vía de señalización de speract, y que para ganar entendimiento de este proceso, se desarrolló también un modelo semicontinuo para una red simple de tres nodos.

El siguiente capítulo aborda detalles necesarios para la construcción de los programas de ajuste de parámetros para ambos modelos de ecuaciones de Glass.

MATERIALES Y MÉTODOS

Este capítulo describe los materiales y métodos empleados para el ajuste de parámetros de los modelos de ecuaciones de Glass de la vía de señalización de speract presentada en el capítulo anterior y la red de tres nodos del apéndice A. Define los datos experimentales que se usaron para la comparación, las funciones objetivo y estrategias de búsqueda, así como mencionar detalles de la implementación computacional de los programas.

2.1 DATOS EXPERIMENTALES

Los datos experimentales que se usaron en todas las comparaciones para discriminar modelos de ecuaciones de Glass corresponden a mediciones de fluorescencia de Ca^{2+} obtenidos en el laboratorio del Dr. Alberto Darszon. Un ejemplo de estas mediciones se muestra en la figura 1. Como se mencionó en la sección 1.1 del capítulo 1, los detalles experimentales de la obtención de estas mediciones se describen a profundidad en Darszon et al. [12] y Wood et al. [59].

Es importante recalcar que la fluorescencia de calcio es el único tipo de dato experimental que cuenta con mediciones largas (del orden de 10 segundos, aproximadamente).

2.2 FUNCIONES OBJETIVO

Como se mencionó en 1.6.1, es necesario establecer una noción de distancia para determinar qué tan bueno es el ajuste a las mediciones experimentales de un modelo de ecuaciones de Glass con un conjunto de parámetros determinado. Para este efecto, anteriormente se mencionó que los criterios que pueden ayudar a esta determinación son el Error Cuadrático Medio 1.6.1.1, el Índice de Correlación de Pearson 1.6.1.2 y el Índice de Pendiente 1.6.1.3.

Se definieron funciones de costo para los anteriores criterios. En cada función, \mathbf{X} representa a las mediciones experimentales, \mathbf{Y} representa a los datos del modelo de ecuaciones de Glass. Finalmente, X_i y Y_i representan a cada uno de los puntos de la medición y de la solución de las ecuaciones de Glass, respectivamente. Cabe recalcar que las únicas mediciones experimentales con las que se cuenta son las

de fluorescencia de Ca^{2+} , y por lo tanto, los datos del modelo que se compararon con esas mediciones fueron los correspondientes a la dinámica del nodo de $[Ca^{2+}]_i$.

Las funciones son

$$f_{MSE}(\mathbf{X}, \mathbf{Y}) = \frac{1}{n} (Y_i - X_i)^2 \quad (12)$$

$$f_{Pearson}(\mathbf{X}, \mathbf{Y}) = 1 - r_{\mathbf{X}, \mathbf{Y}} \quad (13)$$

$$f_{SI}(\mathbf{X}, \mathbf{Y}) = 1 - SI(X, Y) \quad (14)$$

donde $r_{\mathbf{X}, \mathbf{Y}}$ es el Coeficiente de Correlación de Pearson, definido en 10, SI es el Índice de Pendiente, definido en 11. De la definición de el Coeficiente de Correlación de Pearson y del Índice de Pendiente, se puede ver que $1 - r_{\mathbf{X}, \mathbf{Y}} = 2$, si $r_{\mathbf{X}, \mathbf{Y}} = -1$, es decir, cuando \mathbf{X} y \mathbf{Y} están correlacionados negativamente, mientras que $1 - r_{\mathbf{X}, \mathbf{Y}} = 0$, si $r_{\mathbf{X}, \mathbf{Y}} = 1$, es decir, cuando la correlación entre los datos experimentales y el modelo es perfecta. Entonces, minimizar $f_{Pearson}(\mathbf{X}, \mathbf{Y})$ permite elegir un modelo que produzca los datos Y tal que la similitud con los experimentos sea mayor. Un argumento similar muestra que minimizar $f_{SI}(\mathbf{X}, \mathbf{Y})$ permite seleccionar un modelo que produzca datos Y con una similitud mayor a los datos experimentales.

2.3 ESTRATEGIAS DE BÚSQUEDA

En este trabajo, las estrategias de exploración utilizadas fueron BÚSQUEDA ALEATORIA, ALGORITMOS GENÉTICOS y EVOLUCIÓN DIFERENCIAL, cada uno de los cuales se describe a continuación.

2.3.0.2 Búsqueda Aleatoria

La búsqueda aleatoria es la más simple de las estrategias de búsqueda. Consiste en elegir al azar una solución, evaluarla y, si y solo si su costo es mejor que el mejor costo hasta el momento, guardar la solución como la mejor hasta el momento; en caso contrario no se guarda la solución ni se actualiza el valor del mejor costo. En ambos casos, si no se ha llegado a una precisión determinada para la función de costo o bien se ha alcanzado un número máximo de iteraciones, la búsqueda finaliza; en caso contrario, se elige otra solución al azar y se repite el proceso.

2.3.0.3 Algoritmos Genéticos

Inspirados en el trabajo de Holland [30] y tratados de manera un poco más rigurosa por Goldberg [23], los algoritmos genéticos son una estrategia de búsqueda que simula el proceso de evolución natural a través de los procesos de cruce, mutación y selección.

El algoritmo inicia con un conjunto de agentes denominados población, donde cada agente tiene un *genoma*, es decir, una representación codificada de una solución al problema de optimización. Posteriormente se selecciona a algunos individuos de la población para procrear a la siguiente generación. La selección se realiza mediante diferentes esquemas, aunque en términos generales se suele favorecer a aquellos individuos que tengan un mejor costo. Los individuos no seleccionados no sobreviven. De entre aquellos que sí han sido seleccionados, se eligen pares de agentes y estos combinan su genoma de acuerdo a algún esquema para crear un tercer agente. Este proceso de cruce se realiza hasta completar una cantidad preestablecida de miembros de la nueva generación. Con cierta probabilidad, por lo general baja, algunos individuos de la nueva población sufrirán una mutación en su genoma. El proceso de selección, cruce y mutación continúa hasta que se ha alcanzado cierta precisión en el costo de la solución o bien se ha alcanzado un número determinado de iteraciones.

Tradicionalmente el genoma y los operadores de los algoritmos genéticos son discretos. Dado que los parámetros que se requirió estimar en este trabajo toman valores continuos, se usó una variante de los algoritmos genéticos tradicionales, adaptados a trabajar con variables continuas encontrado en Haupt and Haupt [29], en donde el genoma de un agente se codifica como un vector de valores continuos y los operadores de cruce y mutación pueden ser aplicados a este tipo de variables.

2.3.0.4 Evolución Diferencial

Storn and Price [49] observaron que algunos métodos heurísticos existentes no eran tan robustos y no convergían tan rápidamente al ser aplicados a problemas de optimización que involucraran variables continuas. A raíz de esto, desarrollaron un método específicamente pensado para este tipo de problemas.

El método consiste en una población de soluciones candidato, llamadas agentes. Cada agente se mueve a través del espacio de búsqueda combinando las posiciones de los agentes existentes en la población. Si la nueva posición del agente representa una mejora, esta nueva posición se acepta y pasa a formar parte de la población; en caso contrario, la nueva posición se rechaza. El proceso se repite y se espera que una solución satisfactoria se descubrirá eventualmente.

Sea $x \in \mathbb{R}^n$ un agente en la población de tamaño $NP > 3$; sea $F \in [0, 2]$, conocido como peso diferencial; sea también $CR \in [0, 1]$, la probabilidad de cruce. El algoritmo de evolución diferencial consiste en:

1. Inicializar cada agente x con una posición aleatoria en el espacio de búsqueda.
2. Hasta que un criterio de búsqueda sea satisfecho, repetir:
 - a) Escoger al azar tres agentes distintos entre sí a, b, c .

- b) Escoger un índice al azar $R \in 1, \dots, n$ donde n es la dimensionalidad del problema
 - c) Calcular la posible nueva posición del agente, dada por $y = [y_1, \dots, y_n]$ iterando sobre cada $i \in 1, \dots, n$ como sigue:
 - 1) Elegir al azar de manera uniforme $r \in (0, 1)$
 - 2) Si $i = R$ o $r_i < CR$, hacer $y_i = a_i + F(b_i - c_i)$. En caso contrario $y_i = x_i$
 - d) Si $f(y) < f(x)$, entonces $x = y$
3. Elegir el agente con el menor costo o máxima adaptación y regresarlo como la mejor solución candidato encontrada.

La elección de F , CR y NP puede tener un impacto significativo en el desempeño de la optimización. [Storn and Price \[49\]](#), y [Liu and Lampinen \[40\]](#) proporcionan valores iniciales para estos parámetros que parecen ser un buen punto de partida en general.

2.4 BÚSQUEDA DE PARÁMETROS

La búsqueda de parámetros para los modelos de ecuaciones de Glass se realizó usando diferentes combinaciones de funciones objetivo y estrategias de búsqueda. La implementación de las rutinas se realizó en lenguaje C [Kernighan \[36\]](#) bajo el estándar ISO C99 publicado por ISO [32] usando el compilador CLANG de [The clang Developers and Contributors \[53\]](#). Adicionalmente, se usaron los módulos de estadística, métodos de solución de ecuaciones diferenciales ordinarias, vectores y matrices, y generación de números aleatorios de la biblioteca de funciones para cómputo científico GNU SCIENTIFIC LIBRARY (GSL) v.1.15 [Galassi et al. \[16\]](#).

Para la solución de ecuaciones diferenciales se utilizó el método de Euler [Press et al. \[46\]](#). También se utilizaron el método de Runge-Kutta 4,5 [Press et al. \[46\]](#), [Galassi et al. \[16\]](#); y el método de paso adaptativo de Adams-Bashforth [16]. Estos dos últimos métodos son menos sensibles a errores que el método de Euler original.

El método de Euler se implementó directamente, mientras que para el caso del método de Runge-Kutta 4,5 y del método de Adams-Bashforth se utilizó la implementación de la GNU SCIENTIFIC LIBRARY (GSL) v.1.15. La generación de números aleatorios se realizó usando la implementación de la GNU SCIENTIFIC LIBRARY (GSL) v.1.15 del algoritmo RANLXS2 [Galassi et al. \[16\]](#).

Como esquema general, los programas de inferencia de parámetros consistían en proponer soluciones candidato y evaluarlas, repitiendo el proceso hasta cumplir con un criterio de paro. Una solución candidato consiste en un conjunto de parámetros propuestos de acuerdo a una estrategia de exploración del espacio de búsqueda y las trayectorias solución del problema de condiciones iniciales del sistema de

ecuaciones de Glass correspondiente a dichos parámetros. La evaluación de soluciones candidato consistió en calcular su costo de acuerdo a las funciones objetivo de la sección 2.2, para las cuales X son las mediciones experimentales de fluorescencia de calcio y Y la dinámica del nodo de $[Ca^{2+}]_i$ del modelo de Glass dado por un conjunto de parámetros particulares.

2.5 RESUMEN

En este capítulo se definieron las herramientas necesarias para implementar los programas de búsqueda de parámetros de los modelos de ecuaciones de Glass de la vía de señalización de speract presentada en el capítulo anterior y la red de tres nodos del apéndice A. El siguiente capítulo muestra y discute los resultados de la ejecución de las estrategias de búsqueda.

RESULTADOS

Este capítulo muestra los resultados de la estimación de parámetros para la construcción de modelos de ecuaciones de Glass tanto para la red de tres nodos del apéndice A como para la vía de señalización de speract presentada en el capítulo 1. Posteriormente se presenta una breve discusión de los resultados.

3.1 MODELO DE TRES NODOS

Como caso de ejemplo de estimación de parámetros en un problema de transformación de un modelo discreto a un modelo semicontinuo, se resolvió mediante el método de Euler el sistema de ecuaciones de Glass de la red de 3 nodos con parámetros $\alpha_{\sigma_x} = \alpha_{\sigma_y} = \alpha_{\sigma_z} = 1$, umbrales de activación $\theta_{\sigma_x} = 0.3$, $\theta_{\sigma_y} = 0.7$ y $\theta_{\sigma_z} = 0.8$, y $\sigma_x = 1$, $\sigma_y = 0$, $\sigma_z = 0$ como condiciones iniciales. El objetivo consistió en encontrar mediante un procedimiento de búsqueda el conjunto de parámetros tales que la dinámica de Glass correspondiente al σ_x fuera lo más similar posible a la solución conocida para los parámetros antes mencionados. Esta solución conocida se tomó como una señal artificial, que se muestra en la figura 6.

Se eligió búsqueda aleatoria como estrategia de exploración del espacio de búsqueda, mientras que la función objetivo consistió en minimizar $f_{SI}(X, Y) = 1 - SI(X, Y)$ con X la señal original y Y los datos de la dinámica de Glass del nodo σ_z de cada uno de los candidatos propuestos por el algoritmo de búsqueda.

En este caso, fue posible recuperar el valor de los tres parámetros de umbral $\theta_{\sigma_x}, \theta_{\sigma_y}, \theta_{\sigma_z}$ en varias ejecuciones del programa de búsqueda. La figura 7 muestra la dinámica original y el resultado obtenido con dos ejecuciones distintas del programa de búsqueda.

Se probó realizar la estimación de parámetros para el mismo sistema usando una versión de algoritmos genéticos adaptada a problemas con variables continuas con los parámetros recomendados por omission, tal como se refiere en Haupt and Haupt [29]. En este caso fue posible recuperar el valor del parámetro de umbral para σ_z , si bien con una precisión menor.

Cabe mencionar que a pesar de que este caso de ejemplo no está relacionado en un sentido bioquímico con la red de señalización de

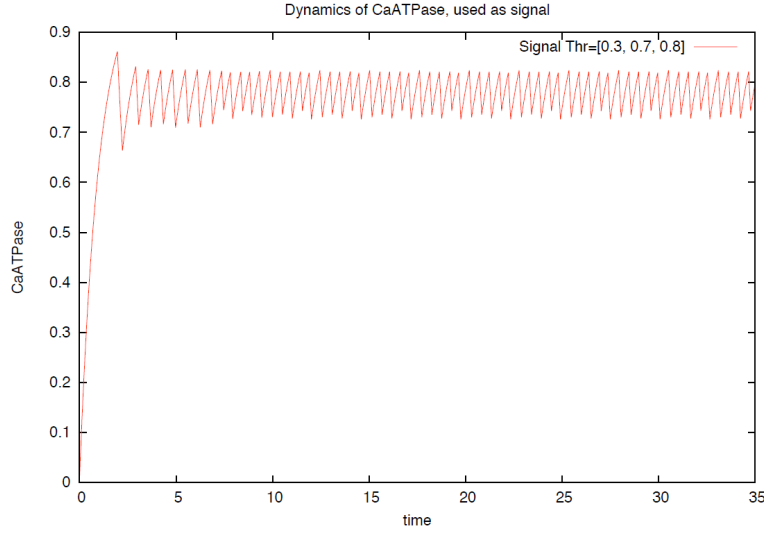


Figura 6: Solución del problema de condiciones iniciales $\sigma_x = 1$, $\sigma_y = 0$, $\sigma_z = 0$ de las ecuaciones 25, 26 y 27, correspondientes al modelo de Glass de la red de 3 nodos del apéndice A, con parámetros $\alpha_{\sigma_x} = \alpha_{\sigma_y} = \alpha_{\sigma_z} = 1$, umbrales de activación $\theta_{\sigma_x} = 0.3$, $\theta_{\sigma_y} = 0.7$ y $\theta_{\sigma_z} = 0.8$. Se muestra la solución con respecto a tiempo de σ_z .

speract, sí lo está en términos de ser un modelo discreto que se quiere reescribir como semicontinuo. Además, para la estimación de parámetros se usó un solo tipo de señal o “medición experimental” contra la cual comparar la dinámica de cada uno de los candidatos, situación que se presentó también en la red de señalización, donde solo se cuenta con un único tipo de mediciones para tiempos largos. La experiencia ganada con este caso de ejemplo sirvió como base para abordar la estimación de parámetros del modelo semicontinuo de la vía de señalización de speract.

3.2 MODELO DE LA RED DE SEÑALIZACIÓN

El modelo semicontinuo de la vía de señalización de speract se implementó planteando ecuaciones de Glass cuyo valor depende de un conjunto de funciones discretas, estas últimas basadas en el modelo de Espinal et al. [14]. A modo de ejemplo, puede considerarse la ecuación de Glass para guanilato ciclasa (GC). La función Booleana para este nodo de la red es

$$GC = F_{GC}(SR \wedge \neg pH_i) \quad (15)$$

y la ecuación de Glass correspondiente es

$$\frac{dGC}{dt} = \alpha_{GC}[F_{GC}(\widehat{SR}, \widehat{pH_i}) - GC] \quad (16)$$

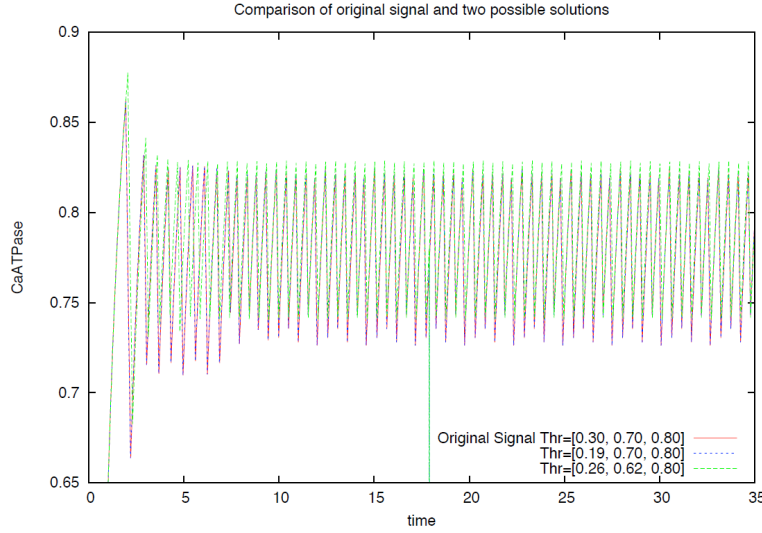


Figura 7: Dinámicas original y estimada de σ_z de la red de 3 nodos. En rojo la señal original, verde y azul representan dos distintas ejecuciones del programa de búsqueda.

donde

$$\widehat{SR} = H(SR - \theta_{SR}) \quad (17)$$

$$\widehat{pH_i} = H(pH_i - \theta_{pH_i}) \quad (18)$$

son los valores discretizados de las variables continuas receptor de speract (SR) y pH intracelular (pH_i), respectivamente.

El problema de estimación de parámetros en este caso consiste en ajustar el valor de 26 parámetros de umbral θ_i (22 para los nodos binarios y 4 extras para los nodos que tienen un valor terciario), en el caso de un sistema sincronizado donde todos los parámetros $\alpha_i = x$, $x \in [0, T_{max}]$, donde T_{max} es el tiempo característico más grande del sistema.

Un problema similar, si bien más complicado, es permitir que cada α_i tome un valor distinto de los demás α_i , es decir, un problema asíncrono. El problema asíncrono añade la necesidad de estimar 22 parámetros extra, uno para cada componente del sistema.

Como una primera aproximación al problema, se consideró el problema sincronizado, estableciendo $\alpha_i = 1.0 \forall i$, de modo que solo se debió estimar el valor de los 26 parámetros de umbral. Las condiciones iniciales se fijaron tales que al discretizar los valores continuos de la condición inicial, se tuviera una situación muy similar a la del organismo silvestre o *wild type*, es decir, aquél al que no se le ha hecho modificación alguna. Tras la discretización de los valores continuos, la condición inicial del modelo semicontinuo es similar a la condición

inicial del modelo discreto presentado en la sección 1.3, tal y como fue presentada en la figura 4.

En particular la condición inicial del modelo de ecuaciones de Glass para la vía de señalización se estableció con los valores

$$\begin{aligned} V &= 1.0 \\ HVA &= 1.0 \\ LVA &= 1.0 \\ [Ca^{2+}]_i &= 0.9 \\ \text{Todos los demás} &= 0.2 \end{aligned}$$

Estos valores indican, como se mencionó en el pie de la figura 4 de la sección 1.3, estados en los que el potencial de membrana (V) está en el estado que representa la hiperpolarización de la membrana del flagelo, los canales canal de Calcio dependiente de alto voltaje (HVA) y canal de Calcio dependiente de bajo voltaje (LVA) se encuentran activos (pero no abiertos) y el calcio intracelular ($[Ca^{2+}]_i$) está en un valor muy cercano al inicio de las oscilaciones tónicas.

En vista de que aun para el problema sincronizado el espacio de parámetros es mayor que el del problema de 3 nodos presentado en la sección anterior, y de que los algoritmos genéticos no presentaron una mejora con respecto a la búsqueda aleatoria en el mismo problema, se buscó otro método de exploración del espacio de búsqueda. A este respecto, Moles et al. [41] muestran que para problemas de estimación de parámetros en redes de señalización bioquímica los métodos con mejores resultados son aquellos basados en estrategias evolutivas. En especial algunos como Evolución Diferencial resultan aún mejores que los algoritmos genéticos.

Para la búsqueda de parámetros de la vía de señalización de speract se usó entonces Evolución Diferencial como estrategia de búsqueda (y evaluación de candidatos), usando los parámetros recomendados en la literatura Storn and Price [49]. Para comparar la dinámica del nodo de $[Ca^{2+}]_i$, (denotado en el modelo como dCa) de cada modelo planteado con las mediciones experimentales de fluorescencia de $[Ca^{2+}]_i$, se buscó minimizar $f_{Pearson}(X, Y) = 1 - r_{X,Y}$, donde X son los datos experimentales y Y son los datos de la dinámica del nodo de calcio del modelo.

Los sistemas de ecuaciones de Glass se resolvieron mediante el método solución de ecuaciones diferenciales de Euler. Para algunas dinámicas elegidas por el algoritmo de Evolución Diferencial presentado en 2.3.0.4, se verificó la solución usando el método para solución de ecuaciones diferenciales ordinarias de Adams-Bashforth, Galassi et al. [16]. No se encontraron diferencias significativas entre las soluciones de uno y otro métodos.

Se logró ajustar la dinámica del nodo de Ca^{2+} de acuerdo a mediciones experimentales, como se muestra en la figura 8.

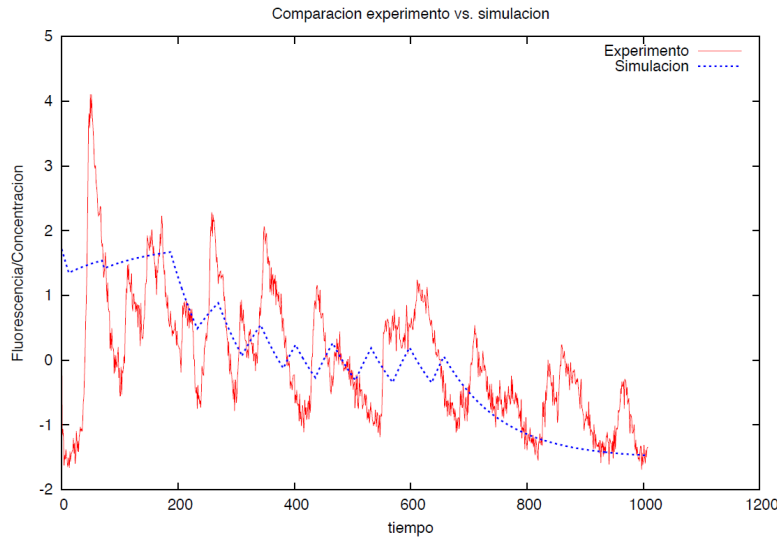


Figura 8: Dinámica del nodo de $[Ca^{2+}]_i$ del modelo y medición experimental de Ca^{2+} . En rojo la medición experimental, en azul el ajuste usando la función objetivo basada en la correlación de Pearson. Para efectos de comparación, las series se normalizaron de manera que tuvieran promedio 0 y varianza 1.

A pesar de lo alentador de este resultado, el comportamiento dinámico de otros nodos aún necesita mayor revisión. Por ejemplo, el nodo correspondiente al potencial de membrana (V) debería de disminuir su valor al inicio de la dinámica para luego aumentarlo. Sin embargo, el potencial aumenta al principio y disminuye posteriormente. Este comportamiento es imposible en la realidad, ya que al haber un aumento de $[Ca^{2+}]_i$ el potencial de membrana (V) debería disminuir. Véase la figura 9.

Varias ejecuciones del algoritmo de evolución diferencial, diferentes parámetros para el mismo algoritmo y funciones objetivo distintas, es decir, usando error cuadrático medio MSE o el índice de pendiente (SI) arrojaron resultados similares.

3.3 DISCUSIÓN DE LOS RESULTADOS

En cuanto al modelo Booleano de tres nodos, es interesante resaltar que la relación *parámetros a estimar/tipos de mediciones* es alta (1/3). Contar con un tipo de medición para un sistema de 3 componentes parece permitir la estimación de parámetros, aún con una estrategia de exploración bastante pobre como la búsqueda aleatoria.

El hecho de que la búsqueda aleatoria se haya comportado mejor que los algoritmos genéticos bien podría deberse a que algunos de los parámetros sugeridos para el algoritmo genético resultaran inadecuados. De manera muy particular, el parámetro de mutación de 0.2 puede evitar que algunas soluciones buenas no logren mantenerse el tiem-

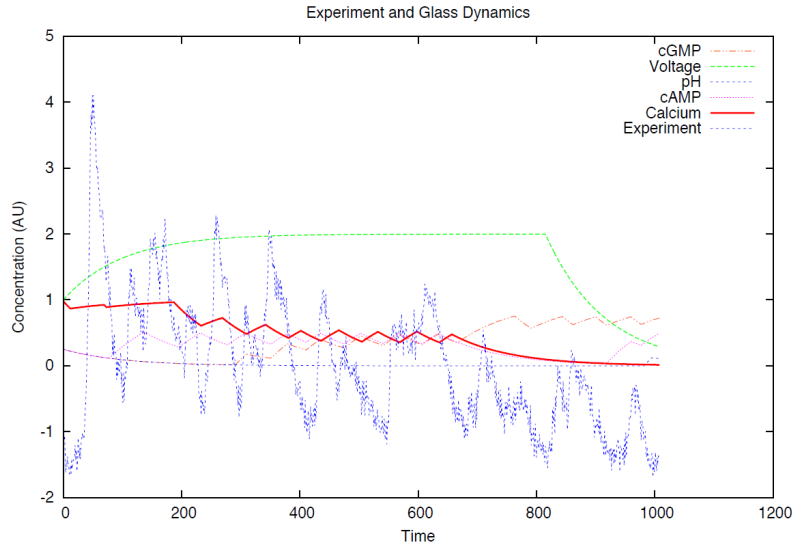


Figura 9: Dinámica del nodo de $[Ca^{2+}]_i$, v y otros componentes del modelo, y medición experimental de Ca^{2+} . En este caso en azul punteado se muestra la medición experimental, en rojo el nodo de $[Ca^{2+}]_i$ del modelo y en verde el nodo que representa al potencial de membrana (v) en el modelo. Nótese que se trata de una ejecución del algoritmo de búsqueda distinta a la de la figura 8. Para efectos de comparación, las series se normalizaron de manera que tuvieran promedio 0 y varianza 1.

po suficiente para propagarse y dominar la población. En este sentido, puede ser que los parámetros por defecto que se utilizaron en esta búsqueda no resulten adecuados para este problema en particular.

Además, dado que los algoritmos genéticos fueron formulados originalmente para problemas discretos, el usar adaptaciones a problemas continuos puede implicar que algunos parámetros para el algoritmo genético deban de elegirse con mayor cuidado, Haupt and Haupt [29]. Fue para evitar este problema que se optó por realizar la búsqueda mediante el algoritmo de evolución diferencial, presentado en la sección 2.3.0.4.

Con lo que respecta al modelo de la vía de señalización de speract, la relación *parámetros a estimar/tipos de mediciones* es, al contrario del modelo de 3 nodos, muy baja (1/26). En principio se quiere encontrar un conjunto de parámetros que produzcan una dinámica coherente desde el punto de vista biológico, es decir, que el comportamiento de todos y cada uno de los nodos tenga sentido. Por ejemplo, que el potencial de membrana (v) se hiperpolarice cuando el calcio intracelular ($[Ca^{2+}]_i$) disminuya, o viceversa. Al contar solamente con un tipo de medición experimental, es posible ajustar con cierto grado de precisión el comportamiento del nodo de la red relacionado directamente con esa medición, como es el caso del relativo buen ajuste que se logró para el $[Ca^{2+}]_i$. Por el contrario, al no comparar el comportamiento de

los demás nodos de la red con mediciones experimentales que estén directamente relacionadas con ellos, resulta difícil encontrar un conjunto de parámetros que produzca un modelo biológicamente válido. En este sentido, la metodología presentada no permite distinguir los modelos biológicamente coherentes de aquellos que cuya interpretación fisiológica no es correcta.

Una opción para resolver este problema es utilizar funciones objetivo que califiquen el comportamiento de varios nodos e imponer restricciones a priori a las soluciones posibles.

En este trabajo no se incorporaron otros tipos de mediciones experimentales para tiempos largos por carecer de ellas. Algunos avances en las técnicas experimentales necesarias para realizar otras mediciones podrían en un futuro arrojar datos que sea posible incluir en el diseño de otras funciones objetivo.

3.4 RESUMEN

Este capítulo mostró los resultados de los programas de ajuste de parámetros para dos modelos de ecuaciones de Glass y luego presentó una breve discusión de estos resultados. El siguiente capítulo presenta algunas conclusiones y perspectivas de trabajo a futuro.

CONCLUSIONES Y PERSPECTIVAS

Este capítulo presenta algunas conclusiones, y al final menciona ideas que podrían ser de utilidad para extender y refinar la investigación iniciada en este trabajo.

4.1 CONCLUSIONES

LA TRANSFORMACIÓN DE UN MODELO DISCRETO EN SEMICONTINUO PRODUCE BUENOS RESULTADOS EN SISTEMAS DONDE LA RELACIÓN *parámetros a estimar/tipos de mediciones* ES ALTA. El modelo Booleano de 3 nodos muestra que el paso de un formalismo de modelación a otro puede ser exitoso cuando se cuenta con datos suficientes contra los cuales comparar el modelo, en relación a la cantidad de parámetros a estimar. Por el contrario, en modelos como el de la vía señalización, en los que la cantidad de parámetros es mucho mayor que los criterios contra los cuales comparar, es difícil encontrar resultados biológicamente relevantes.

LA ELECCIÓN DE PARÁMETROS PARA MODELOS DE GLASS NO ES SIMPLE. A pesar del éxito que han tenido algunos modelos ([Alvarez-Buylla et al. \[3\]](#), [Azpeitia et al. \[4\]](#)) en reproducir el comportamiento observado experimentalmente, este trabajo muestra que no siempre es directo encontrar un conjunto de parámetros tales que la dinámica del modelo tenga una interpretación biológica correcta. Establecer los inversos de tiempo característicos en un valor de 1 y los umbrales de cada nodo en 0.5 parece ser una buena primera elección ante falta de información detallada, pero no necesariamente es correcta en todos los casos.

ES NECESARIA UNA METODOLOGÍA GENERAL PARA LA TRANSFORMACIÓN DE MODELOS DISCRETOS EN MODELOS DE ECUACIONES DE GLASS. La metodología empleada en este trabajo no presenta una forma generalizada para la transformación de modelos discretos en semicontinuos basados en ecuaciones de Glass. Este tipo de modelos resultan atractivos puesto que constituyen un paso intermedio entre los modelos discretos, que requieren poca información, y los basados en ecuaciones diferenciales, que requieren más información detalla-

da. Contar con una metodología generalizada permitiría construir de manera semiautomática una representación más detallada del sistema que la de un modelo discreto, a partir de apenas un poco más de información que la usada para construir a este último pero sin llegar al nivel de detalle requerido por un modelo de ecuaciones diferenciales ordinarias.

4.2 PERSPECTIVAS Y TRABAJO FUTURO

Este trabajo se desarrolló bajo el supuesto de que la transformación de modelos discretos en semicontinuos podía realizarse mediante el empleo de técnicas de estimación de parámetros, particularmente aquellas basadas en estrategias evolutivas. Las funciones objetivo utilizadas se basan en criterios simples de correlación o minimización de diferencias.

Existe un marco teórico más desarrollado para la solución de problemas inversos, que cuenta con técnicas elaboradas para el problema de discriminación de modelos. Una técnica que puede resultar prometedora para la solución de este tipo de problemas puede ser la REGULARIZACIÓN PROMOTORA DE DISPERSIÓN O SPARSITY ENFORCING REGULARIZATION, Engl et al. [13], en donde se busca hacer uso de criterios estadísticos para la comparación entre mediciones experimentales y modelos.

Otra posibilidad que valdría la pena explorar consiste en modificar el modelo discreto de modo que contenga menos nodos y por lo tanto haya menos parámetros que estimar. Hasta este momento cuál es un buen criterio general de reducción de modelos discretos sigue siendo una pregunta abierta. Sin embargo, se ha explorado la posibilidad de utilizar criterios de robustez de las redes para separar nodos esenciales para mantener la dinámica y función de la red de aquellos nodos que podrían parecer redundantes.

Usando este último criterio, se ha logrado plantear un modelo discreto de la vía de señalización de speract con 11 nodos. Queda abierto el determinar si a partir de este modelo reducido y los datos experimentales con que se cuenta hasta el momento, es posible ajustar el comportamiento de un modelo semicontinuo de manera que la interpretación biológica del modelo sea más completa que lo alcanzado hasta el momento.

MODELO AUXILIAR DE 3 NODOS

Con el objetivo de explorar y ganar entendimiento acerca de cómo transformar modelos discretos en modelos de Glass, se utilizó una red Booleana de tres nodos que había sido estudiada previamente por Saadatpour et al. [48]. Esta pequeña red de tres nodos pertenece a un módulo más grande relacionado con la vía de señalización de ácido abscísico en plantas.

Se decidió incluir en este trabajo a la red de tres nodos puesto que modela el comportamiento de un fragmento de una vía de señalización bioquímica y está compuesta de pocos nodos, cuyos estados cambian en el tiempo de acuerdo a funciones Booleanas. Coincidentemente, esta red pequeña parece estar relacionada con oscilaciones de calcio en plantas Saadatpour et al. [48]. Finalmente, dentro de la documentación de BooleanNet (<http://code.google.com/p/booleannet/>), un software especializado en la simulación de redes Booleanas y de Glass, Albert et al. [1], los autores usan esta misma red como un ejemplo que muestra la manera de transformar un modelo Booleano en uno de ecuaciones de Glass, basados en el trabajo de Li et al. [39].

La red de tres nodos consta de CIS , Ca_c^{2+} y $Ca^{2+}ATPase$. Como se mencionó anteriormente, estos nodos corresponden a elementos de la vía de señalización del ácido abscísico en plantas. CIS representa el flujo de Ca^{2+} hacia adentro del citosol desde almacenes intracelulares, Ca_c^{2+} es el calcio en citosol y finalmente $Ca^{2+}ATPase$ representa a enzimas e intercambiadores calcio-protones relacionados con la salida de calcio del citosol, Saadatpour et al. [48]. La red completa mostrada en Saadatpour et al. [48] está basada en el trabajo de Li et al. [39]. La figura 10 muestra el modelo de red Booleana de tres nodos.

Las reglas lógicas para la red de CIS , Ca_c^{2+} y $Ca^{2+}ATPase$ son

$$CIS(t+1) = Ca_c^{2+}(t) \quad (19)$$

$$Ca_c^{2+}(t+1) = CIS(t) \wedge \neg Ca^{2+}ATPase(t) \quad (20)$$

$$Ca^{2+}ATPase(t+1) = Ca_c^{2+}(t) \quad (21)$$

Por simplicidad, y debido a que el interés de trabajar con esta red es usar un modelo pequeño independientemente de las interpretaciones que se puedan hacer de la vía de señalización del ácido abscísico, deno-

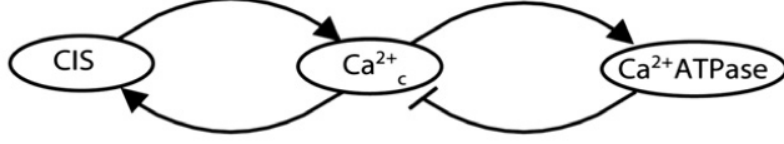


Figura 10: Modelo Booleano de 3 nodos de [Saadatpour et al. \[48\]](#). Las flechas punteadas representan regulación positiva, mientras que las flechas planas representan regulación negativa.

temos a CIS , Ca^{2+}_c y $Ca^{2+}ATPase$ como σ_x , σ_y y σ_z , respectivamente. Las reglas lógicas quedan entonces como

$$\sigma_x(t+1) = \sigma_y(t) \quad (22)$$

$$\sigma_y(t+1) = \sigma_x(t) \wedge \neg \sigma_z(t) \quad (23)$$

$$\sigma_z(t+1) = \sigma_y(t) \quad (24)$$

El modelo de ecuaciones de Glass para la red de tres nodos queda dado por

$$\frac{d\sigma_x}{dt} = \alpha_{\sigma_x} [F_{\sigma_x}(\widehat{\sigma}_y) - \sigma_x] \quad (25)$$

$$\frac{d\sigma_y}{dt} = \alpha_{\sigma_y} [F_{\sigma_y}(\widehat{\sigma}_x, \widehat{\sigma}_z) - \sigma_y] \quad (26)$$

$$\frac{d\sigma_z}{dt} = \alpha_{\sigma_z} [F_{\sigma_z}(\widehat{\sigma}_y) - \sigma_z] \quad (27)$$

donde

$$\widehat{\sigma}_x = H(\sigma_x - \theta_{\sigma_x}) \quad (28)$$

$$\widehat{\sigma}_y = H(\sigma_y - \theta_{\sigma_y}) \quad (29)$$

$$\widehat{\sigma}_z = H(\sigma_z - \theta_{\sigma_z}) \quad (30)$$

corresponden a los valores discretizados de las variables continuas σ_x , σ_y y σ_z , respectivamente, H es la función de Heavyside y los θ_{σ_i} son valores de umbral. Es necesario discretizar el valor de estas variables pues hay que recordar que F_{σ_x} , F_{σ_y} y F_{σ_z} son funciones cuyos argumentos son valores discretos.

BIBLIOGRAFÍA

- [1] István Albert, Juilee Thakar, Song Li, Ranran Zhang, and Réka Albert. Boolean network simulations for life scientists. *Source code for biology and medicine*, 3:16, 2008. (Citado en la página 35.)
- [2] Réka Albert and Hans G Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *Journal of Theoretical Biology*, 223(1):1–18, July 2003. (Citado en la página 5.)
- [3] Elena R Alvarez-Buylla, Alvaro Chaos, Maximino Aldana, Mariana Benítez, Yuriria Cortes-Poza, Carlos Espinosa-Soto, Diego A Hartasánchez, R Beau Lotto, David Malkin, Gerardo J Escalera Santos, and Pablo Padilla-Longoria. Floral morphogenesis: stochastic explorations of a gene network epigenetic landscape. *PLoS ONE*, 3(11):e3626, 2008. (Citado en las páginas 16 y 33.)
- [4] Eugenio Azpeitia, Mariana Benítez, Iliusi Vega, Carlos Villarreal, and Elena R Alvarez-Buylla. Single-cell and coupled GRN models of cell patterning in the *Arabidopsis thaliana* root stem cell niche. *BMC Systems Biology*, 4:134, 2010. (Citado en las páginas 16 y 33.)
- [5] D F DF Babcock, M M MM Bosma, D E DE Battaglia, and A A Darszon. Early persistent activation of sperm K⁺ channels by the egg peptide speract. *Proceedings of the National Academy of Sciences of the United States of America*, 89(13):6001–6005, July 1992. (Citado en la página 8.)
- [6] C Beltrán, O Zapata, and A Darszon. Membrane potential regulates sea urchin sperm adenylylcyclase. *Biochemistry*, 35(23):7591–7598, June 1996. (Citado en la página 3.)
- [7] Martin M Böhmer, Qui Q Van, Ingo I Weyand, Volker V Hagen, Michael M Beyermann, Midori M Matsumoto, Motonori M Hoshi, Eilo E Hildebrand, and Ulrich Benjamin UB Kaupp. Ca²⁺ spikes in the flagellum control chemotactic behavior of sperm. *The EMBO Journal*, 24(15):2741–2752, August 2005. (Citado en la página 4.)
- [8] Stefan Bornholdt. Systems biology. Less is more in modeling large genetic networks. *Science (New York, NY)*, 310(5747):449–451, October 2005. (Citado en la página 5.)
- [9] Kwang-Hyun Cho, Jeong-Rae Kim, Songjoon Baek, Hyung-Seok Choi, and Sang-Mok Choo. Inferring biomolecular regulatory

- networks from phase portraits of time-series expression profiles. *FEBS Lett*, 580(14):3511–3518, Jun 2006. doi: 10.1016/j.febslet.2006.05.035. URL <http://dx.doi.org/10.1016/j.febslet.2006.05.035>. (Citado en las páginas 13 y 14.)
- [10] S P Cook and D F Babcock. Activation of Ca^{2+} permeability by cAMP is coordinated through the pHi increase induced by speract. *The Journal of biological chemistry*, 268(30):22408–22413, October 1993. (Citado en la página 3.)
- [11] A Darszon, T Nishigaki, C. Wood, and CL Treviño. Calcium Channels and Ca^{2+} Fluctuations in Sperm Physiology. *International review of ...*, 2005. (Citado en la página 3.)
- [12] Alberto Darszon, Adan Guerrero, Blanca E. Galindo, Takuya Nishigaki, and Christopher D. Wood. Sperm-activating peptides in the regulation of ion fluxes, signal transduction and motility. *Int J Dev Biol*, 52(5-6):595–606, 2008. doi: 10.1387/ijdb.07255oad. (Citado en las páginas 1, 2 y 19.)
- [13] Heinz W Engl, Christoph Flamm, Philipp Kuegler, James Lu, Stefan Mueller, and Peter Schuster. Inverse problems in systems biology. *Inverse Problems*, 25(12):123014, 2009. URL <http://stacks.iop.org/0266-5611/25/i=12/a=123014>. (Citado en la página 34.)
- [14] Jesus Espinal, Maximino Aldana, Adan Guerrero, Christopher Wood, Alberto Darszon, and Gustavo Martinez-Mekler. Discrete dynamics model for the speract-activated Ca^{2+} signaling network relevant to sperm motility. *PLoS ONE*, 6(8):e22619, 08 2011. doi: 10.1371/journal.pone.0022619. URL <http://dx.doi.org/10.1371/journal.pone.0022619>. (Citado en las páginas VII, 3, 5, 6, 7, 8, 9, 16 y 26.)
- [15] Carlos Espinosa-Soto, Pablo Padilla-Longoria, and Elena R Alvarez-Buylla. A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. *THE PLANT CELL*, 16(11):2923–2939, November 2004. (Citado en las páginas 5 y 16.)
- [16] Mark Galassi, James Theiler, Jim Davies, Brian Gough, Reid Priedhorsky, Gerard Jungman, Mike Booth, Jorma Olavi Tahtinen, Thomas Walter, Fabrice Rossi, Carlo Perassi, Szymon Jaroszewicz, Nicolas Darnis, Jason H. Stover, Ivo Alxneit, Tuomo Kesitalo, Lowell Johnson, Patrick Alken, Rhys Ulerich, Pavel Holoborodko, and Pedro Gonnet. *GNU Scientific Library Reference Manual*, 3rd edition edition. URL <http://www.gnu.org/software/gsl/>. (Citado en las páginas 22 y 28.)

- [17] Blanca E BE Galindo, Anna T AT Neill, and Victor D VD Vacquier. A new hyperpolarization-activated, cyclic nucleotide-gated channel from sea urchin sperm flagella. *Biochemical and Biophysical Research Communications*, 334(1):96–101, August 2005. (Citado en la página 3.)
- [18] Blanca Estela Galindo, José Luis de la Vega-Beltrán, Pedro Labarca, Victor D Vacquier, and Alberto Darszon. Sp-tetraKCNG: A novel cyclic nucleotide gated K(+) channel. *Biochemical and Biophysical Research Communications*, 354(3):668–675, March 2007. (Citado en la página 3.)
- [19] D L Garbers. Sea urchin sperm guanylate cyclase. Purification and loss of cooperativity. *The Journal of biological chemistry*, 251(13):4071–4077, July 1976. (Citado en la página 2.)
- [20] R Gauss, R Seifert, and U B Kaupp. Molecular identification of a hyperpolarization-activated channel in sea urchin sperm. *Nature*, 393(6685):583–587, June 1998. (Citado en las páginas 3 y 4.)
- [21] Carlos Gershenson. Introduction to Random Boolean Networks. *arXiv.org*, nlin.AO, August 2004. (Citado en las páginas 4 y 5.)
- [22] L. Glass and S. A. Kauffman. The logical analysis of continuous, non-linear biochemical control networks. *J Theor Biol*, 39(1):103–129, Apr 1973. (Citado en la página 10.)
- [23] David E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1989. ISBN 0201157675. (Citado en las páginas 15 y 20.)
- [24] G Granados-Gonzalez, I Mendoza-Lujambio, E Rodriguez, B E Galindo, C Beltrán, and A Darszon. Identification of voltage-dependent Ca²⁺ channels in sea urchin sperm. *FEBS Letters*, 579(29):6667–6672, December 2005. (Citado en las páginas 1, 3 y 4.)
- [25] Iain A Greenwood and Normand Leblanc. Overlapping pharmacology of Ca²⁺-activated Cl⁻ and K⁺ channels. *Trends in Pharmacological Sciences*, 28(1):1–5, January 2007. (Citado en la página 4.)
- [26] Adán Guerrero, Takuya Nishigaki, Jorge Carneiro, Yoshiro Tatsu, Christopher D Wood, and Alberto Darszon. Tuning sperm chemotaxis by calcium burst timing. *Developmental biology*, 344(1):52–65, August 2010. (Citado en la página 4.)
- [27] Adán Guerrero, Christopher D Wood, Takuya Nishigaki, Jorge Carneiro, and Alberto Darszon. Tuning sperm chemotaxis. *Biochemical Society Transactions*, 38(5):1270–1274, October 2010. (Citado en la página 4.)

- [28] J R JR Hansbrough and D L DL Garbers. Speract. Purification and characterization of a peptide associated with eggs that activates spermatozoa. *The Journal of biological chemistry*, 256(3): 1447–1452, February 1981. (Citado en la página 1.)
- [29] Randy L. Haupt and Sue Ellen Haupt. *Practical genetic algorithms*. John Wiley & Sons, Inc., New York, NY, USA, 1998. ISBN 047-1188735. (Citado en las páginas 21, 25 y 30.)
- [30] John H. Holland. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, 1975. (Citado en la página 20.)
- [31] Sui Huang, Gabriel Eichler, Yaneer Bar-Yam, and Donald E. Ingber. Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Phys. Rev. Lett.*, 94:128701, Apr 2005. doi: 10.1103/PhysRevLett.94.128701. URL <http://link.aps.org/doi/10.1103/PhysRevLett.94.128701>. (Citado en las páginas 5 y 8.)
- [32] ISO. The ANSI C standard (C99). Technical Report WG14 N1124, ISO/IEC, 1999. URL <http://www.open-std.org/JTC1/SC22/WG14/www/docs/n1124.pdf>. (Citado en la página 22.)
- [33] S A Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 22(3):437–467, March 1969. (Citado en la página 4.)
- [34] U Benjamin Kaupp, Johannes Solzin, Eilo Hildebrand, Joel E Brown, Annika Helbig, Volker Hagen, Michael Beyermann, Francesco Pampaloni, and Ingo Weyand. The signal flow and motor response controlling chemotaxis of sea urchin sperm. *Nature cell biology*, 5(2):109–117, February 2003. (Citado en la página 4.)
- [35] U Benjamin Kaupp, Nachiket D Kashikar, and Ingo Weyand. Mechanisms of sperm chemotaxis. *Annual review of physiology*, 70: 93–117, 2008. (Citado en la página 1.)
- [36] Brian W. Kernighan. *The C Programming Language*. Prentice Hall Professional Technical Reference, 2nd edition, 1988. ISBN 0131103709. (Citado en la página 22.)
- [37] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983. doi: 10.1126/science.220.4598.671. URL <http://www.sciencemag.org/content/220/4598/671.abstract>. (Citado en la página 15.)
- [38] H C Lee and D L Garbers. Modulation of the voltage-sensitive Na⁺/H⁺ exchange in sea urchin spermatozoa through membrane potential changes induced by the egg peptide speract. *The Journal of biological chemistry*, 261(34):16026–16032, December 1986. (Citado en la página 3.)

- [39] Song Li, Sarah M Assmann, and Réka Albert. Predicting Essential Components of Signal Transduction Networks: A Dynamic Model of Guard Cell Abscissic Acid Signaling. *PLoS Biology*, 4(10):e312, September 2006. (Citado en la página 35.)
- [40] J. Liu and J. Lampinen. On setting the control parameter of the differential evolution method. In *Proceedings of the 8th International Conference on Soft Computing (MENDEL)*, pages 62–67, 2002. (Citado en la página 22.)
- [41] Carmen G. Moles, Pedro Mendes, and Julio R. Banga. Parameter estimation in biochemical pathways: A comparison of global optimization methods. *Genome Research*, 13(11):2467–2474, 2003. doi: 10.1101/gr.1262503. URL <http://genome.cshlp.org/content/13/11/2467.abstract>. (Citado en las páginas 15 y 28.)
- [42] Melody K Morris, Julio Saez-Rodriguez, Peter K Sorger, and Douglas A Lauffenburger. Logic-based models for the analysis of cell signaling networks. *Biochemistry*, 49(15):3216–3224, April 2010. (Citado en la página 5.)
- [43] T Nishigaki, CD Wood, Y Tatsu, N Yumoto, T Furuta, D Elias, K Shiba, SA Baba, and A Darszon. A sea urchin egg jelly peptide induces a cGMP-mediated decrease in sperm intracellular Ca^{2+} before its increase. *Developmental biology*, 272(2):376–388, 2004. (Citado en las páginas 3 y 4.)
- [44] Gbolahan W Okunade, Marian L Miller, Gail J Pyne, Roy L Sutliff, Kyle T O'Connor, Jonathan C Neumann, Anastasia Andringa, Daniel A Miller, Vikram Prasad, Thomas Doetschman, Richard J Paul, and Gary E Shull. Targeted ablation of plasma membrane Ca^{2+} -ATPase (PMCA) 1 and 4 indicates a major housekeeping function for PMCA1 and a critical role in hyperactivated sperm motility and male fertility for PMCA4. *The Journal of biological chemistry*, 279(32):33742–33750, August 2004. (Citado en la página 4.)
- [45] Edward Perez-Reyes. Molecular physiology of low-voltage-activated t-type calcium channels. *Physiological Reviews*, 83(1): 117–161, January 2003. (Citado en la página 3.)
- [46] William Press, Saul Teukolsky, William Vetterling, and Brian Flannery. *Numerical Recipes in C*. Cambridge University Press, Cambridge, UK, 2nd edition, 1992. (Citado en las páginas 15 y 22.)
- [47] Esmeralda Rodríguez and Alberto Darszon. Intracellular sodium changes during the speract response and the acrosome reaction

- in sea urchin sperm. *The Journal of physiology*, 546(Pt 1):89–100, January 2003. (Citado en las páginas 3 y 4.)
- [48] Assieh Saadatpour, Istvan Albert, and Reka Albert. Attractor analysis of asynchronous boolean models of signal transduction networks. *Journal of Theoretical Biology*, 266(4):641 – 656, 2010. ISSN 0022-5193. doi: 10.1016/j.jtbi.2010.07.022. URL <http://www.sciencedirect.com/science/article/pii/S0022519310003796>. (Citado en las páginas 10, 15, 16, 35 y 36.)
- [49] Rainer Storn and Kenneth Price. Differential evolution. a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341–359, 1997. ISSN 0925-5001. URL <http://dx.doi.org/10.1023/A:1008202821328>. (Citado en las páginas 15, 21, 22 y 28.)
- [50] Timo T Strünker, Ingo I Weyand, Wolfgang W Bönigk, Qui Q Van, Astrid A Loogen, Joel E JE Brown, Nachiket N Kashikar, Volker V Hagen, Eberhard E Krause, and U Benjamin UB Kaupp. A K+-selective cGMP-gated ion channel controls chemosensation of sperm. *Nature cell biology*, 8(10):1149–1154, October 2006. (Citado en las páginas 1, 3 y 4.)
- [51] Yi-Hsien YH Su and Victor D VD Vacquier. A flagellar K(+)-dependent Na(+)/Ca(2+) exchanger keeps Ca(2+) low in sea urchin spermatozoa. *Proceedings of the National Academy of Sciences of the United States of America*, 99(10):6743–6748, May 2002. (Citado en las páginas 3 y 4.)
- [52] N Suzuki. Structure, Function and Biosynthesis of Sperm-Activating Peptides and Fucose Sulfate Glycoconjugate in the Extracellular Coat of Sea-Urchin Eggs. *Zoological Science*, 12(1):13–27, 1995. (Citado en la página 1.)
- [53] The clang Developers and Contributors. clang: a C language family frontend for LLVM. <http://clang.llvm.org/> ; accessed 21-Sept-2011. (Citado en la página 22.)
- [54] Dan Wang, Shelby M King, Timothy A Quill, Lynda K Doolittle, and David L Garbers. A new sperm-specific Na⁺/H⁺ exchanger required for sperm motility and fertility. *Nature cell biology*, 5(12):1117–1122, December 2003. (Citado en la página 3.)
- [55] Wikipedia. Mean squared error — wikipedia, the free encyclopedia, 2012. URL http://en.wikipedia.org/w/index.php?title=Mean_squared_error&oldid=47302847. [Online; accessed 31-January-2012]. (Citado en la página 13.)
- [56] Wikipedia. Pearson product-moment correlation coefficient — wikipedia, the free encyclopedia, 2012. URL

http://en.wikipedia.org/w/index.php?title=Pearson_product-moment_correlation_coefficient&oldid=472133432. [Online; accessed 31-January-2012]. (Citado en la página 13.)

- [57] Chris D Wood, Alberto Darszon, and Michael Whitaker. Speract induces calcium oscillations in the sperm tail. *Journal of Cell Biology*, 161(1):89–101, April 2003. (Citado en las páginas 2, 3 y 4.)
- [58] Christopher D Wood, Takuya Nishigaki, Toshiaki Furuta, Shoji A Baba, and Alberto Darszon. Real-time analysis of the role of Ca(2+) in flagellar movement and motility in single sea urchin sperm. *Journal of Cell Biology*, 169(5):725–731, June 2005. (Citado en las páginas 3 y 4.)
- [59] Christopher D. Wood, Takuya Nishigaki, Yoshiro Tatsu, Noboru Yumoto, Shoji A. Baba, Michael Whitaker, and Alberto Darszon. Altering the speract-induced ion permeability changes that generate flagellar ca2+ spikes regulates their kinetics and sea urchin sperm motility. *Dev Biol*, 306(2):525–537, Jun 2007. doi: 10.1016/j.ydbio.2007.03.036. (Citado en las páginas 2, 4 y 19.)

COLOFÓN

Este documento se produjo con las fuentes Palatino y Euler de Herman Zapf y Bera Mono, una variante de Bitsream Vera modificada por Malte Rosenau y Ulrich Dirr, usando el estilo `classicthesis` de André Miede. El estilo está inspirado en el libro de Robert Bringhurst “*The Elements of Typographic Style*”. `classicthesis` puede obtenerse para \LaTeX y \LyX en:

<http://code.google.com/p/classicthesis/>

Final Version as of 14 de mayo de 2012 (`classicthesis`).