



PROYECTO DE GRADO

Presentado ante la ilustre UNIVERSIDAD DE LOS ANDES como requisito parcial para
obtener el Título de INGENIERO DE SISTEMAS

IMPLEMENTAR EL PROCESO DE
CRECIMIENTO DEL CÁNCER PULMONAR
MEDIANTE UN MODELO BASADO EN
AGENTES

Por

Br. Albert Rojas

Tutor: Prof. Carlos Echeverría

Octubre 2017

©2017 Universidad de Los Andes Mérida, Venezuela

C.C. Reconocimiento

IMPLEMENTAR EL PROCESO DE CRECIMIENTO DEL CÁNCER PULMONAR MEDIANTE UN MODELO BASADO EN AGENTES

Br. Albert Rojas

Proyecto de Grado — Investigación de Operaciones, 84 páginas

Resumen: La evolución del cáncer es considerada como un sistema complejo, ya que posee todas las características inherentes a este tipo de sistemas. En otras palabras, la interacción entre las partes o subsistemas da como resultado una conducta global que no podría ser anticipada desde el comportamiento de componentes aislados. Por lo tanto, el comportamiento global del sistema depende de la naturaleza de las interacciones y de las características de las partes que lo componen, viéndose modificado cuando estas interacciones cambian. De modo que el cáncer está conformado por una sociedad de miembros individuales denominados *células*, éstas se reproducen de una manera cooperativa y organizada siguiendo un conjunto de reglas y respondiendo a las interacciones locales con otras células, ya sean normales o cancerosas [9]. Por todo esto, el presente proyecto de grado tiene como finalidad implementar en la plataforma de simulación GALATEA el trabajo de investigación realizado por Thomas S. Deisboeck y sus colaboradores titulado “Simulating non-small cell lung cancer with a multiscale agent-based model [13]”, el cual es un modelo multiescala construido para estudiar tumores altamente malignos con dinámica compleja de sistemas auto-organizados, haciendo uso de los autómatas celulares. De manera que, la dinámica de crecimiento está definida por reglas de interacción estocásticas, dictaminadas por procesos aleatorios con inclusión de probabilidades [10].

Palabras clave: autómatas celulares, galatea, sistema complejo.

Índice

Índice de Tablas	v
Índice de Figuras	vi
Agradecimientos	vii
1 INTRODUCCIÓN	1
1.1 Antecedentes	2
1.2 Algunos conceptos biológicos importantes	7
1.2.1 Escalas biológicas	13
1.3 Planteamiento del Problema	15
1.4 Justificación	15
1.5 Objetivos	15
1.5.1 Objetivo General	15
1.5.2 Objetivos Específicos	16
1.6 Metodología	16
1.7 Organización del documento	25
2 MARCO TEÓRICO Y CONCEPTUAL	27
2.1 Autómatas celulares (AC)	27
2.1.1 Definición de AC	28
2.1.2 Notación formal de un AC	28
2.2 AC en dos dimensiones	28
2.2.1 Bordes en un AC	29
2.2.2 Reglas en un AC	29

2.3	Modelos basados en Agentes (ABM)	29
2.3.1	ABM Definición Modelo Basado en Agentes	30
2.3.2	Propiedades de los ABM	30
2.4	“Simulating non-small cell lung cancer with a multiscale agent-based model ”.(Resumen).	31
2.5	Ecuaciones reacción-difusión	33
2.5.1	Aproximaciones numéricas de Ecuaciones Diferenciales Parciales	33
2.6	Proyecto GALATEA	37
2.6.1	Características de GALATEA	38
2.6.2	Estructura de la biblioteca	39
3	MODELO DE ACOPLAMIENTO	44
3.1	Supuestos de simplificación del modelo	44
3.2	Algoritmo general propuesto para el diseño del modelo	46
3.3	Resolución de la parte difusiva del modelo	46
3.3.1	Construcción del estado inicial del AC	46
3.3.2	Tratamiento de la Ecuación Diferencial Parcial	50
3.3.3	Solución a las condiciones de borde	52
3.4	Construcción de las capas del AC	54
3.4.1	¿Cómo calcular el tamaño del sistema del AC?	55
3.4.2	Funciones de consultas hechas sobre cada capa	58
3.4.3	Creación de las reglas para las celdas	60
3.5	Integración del modelo molecular Al AC	62
3.5.1	Acoplar las reglas de crecimiento celular	63
3.6	Organización de códigos fuentes	68
3.7	Aportes del presente proyecto	70
4	RESULTADOS	71
4.1	CONCLUSIONES	73
	BIBLIOGRAFÍA	79

Índice de Tablas

1.1 Ecuaciones y concentraciones iniciales.	19
1.2 Parámetros cinéticos.	20
1.3 Coeficientes de difusión de la glucosa, O_2 y EGF.	22

www.bdigital.ula.ve

Índice de Figuras

1.1	Miembros de la familia ERBB ó HER y especificidad de sus ligandos.	13
1.2	Modelo cinético de la vía de señalización.	18
1.3	Algoritmo de decisión fenotípico celular.	24
2.1	Vecinos Von Neumann en fronteras absorbentes.	29
2.2	Características básicas de las tres escalas complejas	31
2.3	Solución aproximada para ecuaciones en derivadas parciales.	34
2.4	Reemplazo del dominio continuo del problema por una malla	34
2.5	Matriz tridiagonal.	37
2.6	Estructura de la librería Galatea	43
3.1	Distribución de capas pertenecientes al Autómata Celular	48
3.2	Inicialización de la capa Glucosa	51
3.3	EDP representada en Galatea	52
3.4	Tiempo de difusión del Oxígeno sobre un rango de distancias	53
3.5	Interfaz Gráfica del AC	69
4.1	Influencias de la proliferación y migración en el artículo base	72
4.2	REGLA A: Migración	72
4.3	REGLA B: Proliferación	73
4.4	Comparación comportamiento de PLCy	74
4.5	Comparación de soluciones	75
4.6	Difusión Capa Glucosa	76
4.7	Difusión Capa Oxígeno	77
4.8	Difusión Capa Oxígeno	77

Capítulo 1

INTRODUCCIÓN

Los autómatas celulares comparten varias propiedades del mundo físico, permitiendo modelar con ellos distintos fenómenos físicos y biológicos, por consiguiente, constituyen uno de los modelos más antiguos de la computación natural, ésta se refiere al estudio de sistemas computacionales que se inspiran y usan ideas a partir de sistemas de la naturaleza, incluyendo sistemas biológicos, ecológicos y físicos [42]. John Von Neumann fue quien, inspirado en la biología, alrededor del año 1940, intentó diseñar sistemas artificiales auto-replicables, que sirvieran para otros propósitos. La idea de Von Neumann era investigar dispositivos computacionales análogos al cerebro humano, en los que la memoria y las unidades de procesamiento no estuvieran separadas entre sí, sino que trabajaran en paralelo y fueran capaces de repararse y reconstruirse con el material necesario. Así, comenzó a pensar en un universo discreto que consistía en una malla bidimensional de máquinas de estado finito, llamadas celdas, interconectadas localmente. Las celdas modificaban su estado sincrónicamente, dependiendo de algunas celdas cercanas, a partir de ciertas reglas[41].

Entre los sistemas biológicos en los que se puede inspirar la computación natural, se puede mencionar el estudio de enfermedades, como las epidemias o neoplasias malignas, entre otras más. El cáncer es una de las enfermedades más antiguas de la humanidad, actualmente constituye un problema social y biológico, aunque hoy en día se ha mejorado en su tratamiento y su detección, todavía representa un reto biológico por lo que su investigación continua, ésta se ha expandido a diferentes campos

de conocimiento incluyendo las ciencias de la computación [42].

Avanzando en este razonamiento, se describe la expresión de Autómatas Celulares basados en un modelo de agentes, que regido por reglas establecidas, simula la proliferación y la migración de células cancerígenas en el tejido del pulmón, siendo estos los comportamientos más importantes en el desarrollo inicial del tumor, a partir de un grupo de células cancerígenas iniciales. Se propone el desarrollo de una herramienta que permita modelar las condiciones iniciales en el desarrollo del cáncer de pulmón de células no pequeñas[42]. En los párrafos siguientes se explicará a mayor profundidad, todo lo relacionado al presente tema, mayormente términos biológicos que son necesarios para la comprensión de este tipo de enfermedades, con el objetivo de tener una idea clara y concisa al momento de realizar un primer prototipo de simulación en la plataforma ya mencionada.

1.1 Antecedentes

El cáncer es el estado final de un extenso y complejo proceso evolutivo originado por la mutación en una única célula que exhibe un crecimiento descontrolado [1], rompiendo la cooperación que mantiene la integridad de un organismo multicelular [2], dicho de otra manera, las células están destinadas a la colaboración y por consiguiente envían, reciben, interpretan y elaboran un conjunto de señales químicas extracelulares, pero en el momento que hay algún problema molecular con estas señales, el equilibrio que existe entre ellas se rompe ocasionando un aumento en la velocidad de crecimiento (aumentando su masa) y división (proliferación) dando origen a lo que se le es denominado con el término de tumor [3].

Por otro lado, habría que decir también que el 75-80% de los cánceres se deben a la acción de agentes externos que actúan sobre el organismo, causando alteraciones en las células. Se estima que el tabaco es el responsable de al menos 30% de los tumores, la dieta de otro 30% y los agentes infecciosos del 18% [4]. El cáncer es una de las principales causas de muertes en todo el mundo, de tal manera que en el año 2012 hubo 8,2 millones de muertes relacionadas con este padecimiento, siendo los más frecuentes de pulmón, mama, estómago e hígado. En concreto 20% de las muertes mundiales por

cáncer general, y alrededor del 70% de las muertes mundiales por cáncer de pulmón [5].

En cuanto a algunos hechos históricos relevantes del cáncer pulmonar, habría que mencionar que fue descrito a mediados del siglo XIX como una neoplasia poco frecuente [6]. En el año 1791 el científico Jonh Hill fue el primero en advertir que el padecimiento de esta enfermedad podría estar relacionada con el uso del tabaco [4]. Sin embargo, no fue hasta el año de 1819 cuando fueron publicadas algunas características relevantes del carcinoma, todas esas observaciones se dificultaban en virtud de las pocas herramientas tecnológicas existentes para esa época que tendía a confundir los diagnósticos médicos con la tuberculosis, padecimiento que aquejaba fuertemente en esos momentos. A fin de encontrar la conexión entre el consumo de tabaco y el cáncer de pulmón, continuaron los estudios en los años siguientes, por tanto, en el año 1929 el médico alemán Fritz Licknt reportó la primera prueba estadística sobre esta sospecha, lo que ocasionó una impresionante campaña antitabaco en la Alemania nazi. No obstante, la evidencia sólida sobre esta investigación fue proporcionada en el año 1950 por médicos británicos [8].

Por otra parte, la manera de estudiar la evolución del cáncer debe ser como un sistema complejo, ya que posee todas las características inherentes a este tipo de sistemas. Es necesario recalcar que en un sistema complejo la interacción entre las partes o subsistemas da como resultado una conducta global que no podría ser anticipada desde el comportamiento de los componentes aislados. Por lo tanto, el comportamiento global del sistema depende de la naturaleza de las interacciones y de las características de las partes que lo componen, viendóse modificado cuando estas interacciones cambian y a su vez pudiendo exhibir auto-organización, lo que conlleva a que sean sistemas no lineales. Así, el cáncer esta conformado por una sociedad de miembros individuales denominados *células*, éstas se reproducen de una manera cooperativa y organizada siguiendo un conjunto de reglas y respondiendo a las interacciones locales con otras células, ya sean normales o cancerosas [9].

Desde principios del siglo XX los físicos y matemáticos se han interesado en aplicaciones biológicas en este ámbito, la generación de modelos de diferente índole a partir de ecuaciones diferenciales ordinarias (EDO), ecuaciones diferenciales parciales

(EDP), modelos estocásticos discretos entre otras herramientas, con el objeto de describir la producción de tumores cancerosos también conocido con el término de tumorigénesis. Aunque la evolución de un sistema cancerígeno está limitada y condicionada a parámetros químicos y biológicos, su dispersión a través de los tejidos y órganos obedece a diversas reglas deterministas o estocásticas, las cuales se pueden modelar utilizando herramientas matemáticas. Por tanto, el uso de EDO se aplica a problemas de estabilidad en sistemas biológicos donde se tiene una población compitiendo por recursos; para el presente caso supone una población de células cancerosas compitiendo por diversos nutrientes, por ejemplo, el oxígeno y la glucosa[10]. En general, los modelos son una abstracción de algún aspecto general de la realidad, de la cual se trata de extraer aquellos puntos fundamentales para el fenómeno observado, incluyendo el cáncer.

La importancia actual de la modelización y la simulación no podría entenderse sin la existencia de medios y métodos computacionales, cada vez más potentes, que hacen posible atacar problemas que por su complejidad hasta hace cierto tiempo eran imposibles de abordar [39]. Entre la variedad de medios computacionales existentes en la actualidad, se puede encontrar la plataforma *GALATEA*, ésta es una familia de lenguajes, compiladores e interpretadores, que contiene distintas herramientas de simulación entre las que se tiene un integrador de EDO, un simulador DEVS[18], una colección de bibliotecas de apoyo y plantillas genéricas para el desarrollo de modelos de simulación de eventos discretos, sistemas continuos y sistemas multi-agentes que permite explorar alternativas para la integración con un formalismo general de modelado y simulación [17], por consiguiente, se presenta como una alternativa para evaluar su desempeño en procesos biológicos difusivos y no lineales como lo es el cáncer.

Así mismo, teniendo en cuenta la característica no lineal de los sistemas complejos hace que sea posible lograr establecer la relación con otras herramientas útiles como lo son los autómatas celulares y los modelos basados en agentes (ABM)[11], dado que ofrecen una perspectiva muy diferente al resto de las metodologías y constituyen una técnica computacional destinada fundamentalmente al estudio de sistemas dinámicos y complejos. Estos modelos utilizan múltiples agentes que, obedeciendo a ciertas reglas locales, interactúan de manera autónoma entre sí y con su entorno. De este

modo, a partir de un comportamiento local, determinado por reglas simples, es posible estudiar el comportamiento global del sistema, para el presente caso el cáncer de pulmón [14]. Esta herramienta es ideal para explorar diversidad de características en las células, condiciones del microentorno, morfología general del tumor, competencia celular, moléculas de señalización entre otros aspectos [24]. Por tanto, los modelos basados en agentes son opciones ideales a las ecuaciones diferenciales en estudiar sistemas complejos.

En relación con algunos trabajos realizados referentes al uso de los ABM y autómatas celulares dedicados a estudiar el ciclo evolutivo del cáncer, se mencionan los siguientes:

- *Modelo de crecimiento de las variantes del virus de papiloma humano tipo 16*: se estudia el progreso del cáncer del cuello uterino, mediante la construcción de un sistema complejo a través de ABM, buscando apoyar procesos de simulación en tareas de investigación del virus [11].
- *Cross-scale, cross-pathway evaluation using an agent-based non-small cell lung cancer model*: se investiga los efectos de la combinación del factor de crecimiento epidérmico con el factor de crecimiento transformante beta, es decir, ambas moléculas pueden incidir en la velocidad de la expansión del tumor [19] [20].
- *Modeling the competition between lung metastases and the immune system using agents*: se investiga la dinámica del cáncer de mama, específicamente en la etapa de metástasis, es decir, la invasión de las células malignas a través de los vasos sanguíneos a otros tejidos o membranas del cuerpo, a través de ABM [21].
- *Development of a Three-Dimensional Multiscale Agent-Based Tumor Model: Simulating Gene-Protein Interaction Profiles, Cell Phenotypes and Multicellular Patterns in Brain Cancer*: constituye un modelo multiescala para investigar tumores en el cerebro altamente malignos como dinámica compleja de sistemas auto-organizados. [22].

Los tumores son caracterizados por un complejo patrón de alteraciones moleculares. A lo largo de las últimas décadas se han identificado varios genes con un papel

fundamental en el cáncer pulmonar. Se trata de genes cuya expresión permite distinguir de manera objetiva entre un estado normal y un estado tumoral, y además pueden servir como herramienta para establecer un diagnóstico temprano de la enfermedad y así permitir optar una terapia dirigida y personalizada con inhibidores propios de los genes expresados. Son biomarcadores [31].

“El EGFR es frecuentemente sobreexpresado en muchos tipos de cánceres, incluyendo el NSCLC. La unión del EGF con el EGFR para el dominio extracelular genera un número de efectos sobre el comportamiento fenotípico de la célula en cada uno de los procesos de señalización (proliferación, invasión, metástasis, otros). En particular, el aumento de la expresión del EGFR genera la hiperactividad del proceso de proliferación celular, disminuye la apoptosis y facilita la invasión sobre los tejidos u órganos de las células malignas [13], [29]. Un modelo *in silico* (hecho por computadora) es considerado ser una herramienta cada vez más prometedora, puesto que facilita la generación de visiones útiles dentro de la dinámica del EGFR como vía de transducción de señales. Sin embargo, la mayoría de los modelos anteriores se centraron solamente en el nivel molecular o celular, despreciando las retroalimentaciones cruciales entre estas escalas, así como también la interacción con el microambiente bioquímico heterogéneo”.

El modelo presentado está constituido por la escala molecular, donde se representa mediante 20 moléculas simbolizadas a través de ecuaciones diferenciales ordinarias, así mismo también esta conformado por la escala microscópica, la cual es descrita mediante agentes, tratados como autómatas celulares. La estructura de un modelo discreto de autómatas celulares está basado en una región del espacio discreta (rejilla cuadrada), donde coexisten determinadas especies celulares. La dinámica de crecimiento está definida por algunas reglas de interacción a lo largo de los nodos y entre los vecinos de cada celda. Las reglas pueden ser determinísticas o estocásticas, estas últimas son dictaminadas por procesos aleatorios con inclusión de probabilidades [10]. En definitiva, se tienen dos modelos con escalas de tiempo y espacio distintas sobre un mismo modelo.

A continuación se presentará a mayor detalle el planteamiento del problema, donde se describirá el sistema a estudiar, sus características, componentes principales y una variedad de aspectos que son relevantes de comprender; posteriormente el modelo

planteado para el problema denominado “Vía de señalización molecular”, donde se observará el papel importante que juega el EGFR y otra variedad de proteínas sobre el desarrollo del NSCLC, el ambiente o microentorno donde se lleva a cabo estos procesos y un algoritmo de decisión fenotipo celular, que describe las alternativas que puede tomar una célula dependiendo de ciertas características que involucra la concentración de nutrientes consumidos por las células, para que se lleve a cabo las correspondientes consecuencias fenotípicas.

1.2 Algunos conceptos biológicos importantes

Existen dos tipos principales de cáncer de pulmón, dependiendo de las células que conforman el tumor: cáncer de pulmón de células no pequeñas y cáncer de pulmón de células pequeñas. El cáncer de pulmón de células no pequeñas es el más frecuente y conforma el 87% de todos los tumores de pulmón, a su vez éste se clasifica en tres clases: el adenocarcinoma, el carcinoma de células escamosas y el carcinoma de células grandes. No obstante, es importante mencionar algunas características que poseen éstos para poder diferenciarlos:

- El adenocarcinoma representa el 40% de todos los NSCLC, se desarrolla generalmente en las regiones más externas del pulmón.
- El carcinoma de células escamosas es también denominado como carcinoma epidermoide, conforma del 25% al 30% de NSCLC, es generalmente encontrado cerca de los bronquios, hacia el centro de la cavidad torácica.
- El carcinoma de células grandes constituye entre el 10% y el 15% de NSCLC, es caracterizado por su crecimiento rápido y por aparecer en cualquier parte del pulmón.

El cáncer pulmonar de células pequeñas, se caracteriza principalmente por tender a crecer y expandirse más rápido que el cáncer de pulmón de células no pequeñas, lo que conlleva a causar la aparición de síntomas con mayor prontitud [26].

Por otra parte, es indispensable describir varios aspectos biológicos a mayor profundidad para llegar a comprender en su totalidad el planteamiento del problema del

trabajo de Wang, ya que es la base para poder entender los modelos y poder describirlos sobre GALATEA. Por lo tanto, en los siguientes párrafos se describirá el proceso de señalización molecular como preámbulo al modelo, es decir, algunos aspectos sobre la comunicación celular, moléculas que intervienen en el proceso, proteínas que regulan la expresión de los genes implicados en la proliferación y supervivencia celular, entre otros puntos, que se detallarán a continuación.

Las bases de la comunicación celular juegan un rol fundamental en la regulación de procesos como proliferación, diferenciación, apoptosis, entre otros más. Estos procesos están directamente relacionados con la carcinogénesis, y se realizan mediante la comunicación entre las células a través de mecanismos complejos de transducción de señales, por consiguiente, todas las células deben ejecutar de forma precisa y regulada sus funciones, de manera que se establece una dependencia mutua entre ellas [29]. Las diversas formas que puede establecerse la comunicación incluyen:

- La comunicación paracrina: la señal actúa sobre las células vecinas;
- Endocrina: la señal viaja por el torrente sanguíneo y alcanza células lejanas;
- Autocrina: la señal llega a la misma célula de la cual salió;
- Neurotransmisión: la señal es liberada por la célula emisora al espacio sináptico, donde es captada por la célula receptora.
- Contacto célula-célula: la señal permanece anclada de la membrana de la célula emisora mientras interactúa con la célula receptora.
- Mediante uniones de hendiduras o GAP: la señal se difunde desde la célula emisora a la receptora.

Las células están sometidas a un flujo continuo de información que proviene tanto de otras células como del exterior del organismo. La correcta interpretación e integración de todas estas señales es crucial para la supervivencia y adaptación tanto de la propia célula como del organismo completo. Durante la evolución, se han desarrollado múltiples mecanismos para interpretar e integrar señales, y los receptores localizados en la membrana son una parte esencial de dichos mecanismos. Entre estos, los receptores

con actividad cinasa de tirosinas son capaces de traducir la información que proviene del exterior a un mensaje intracelular, que en última instancia controlará la expresión génica y por lo tanto la fisiología de la célula [33].

En cada organismo existen diferentes tipos de señales químicas que reciben el nombre de ligandos y forman complejos con receptores específicos. Cada tipo celular es sensible a distintas señales y cada interacción ligando-receptor está asociada a una función particular, éstos últimos pueden estar localizados en el interior de la célula (receptor intracelular) o bien anclados en la membrana plásmatica (receptor de membrana). Así, el complejo ligando-receptor transmite el mensaje al interior de la célula e inicia un camino que lleva a la ejecución de una respuesta biológica específica, esto es un proceso complejo conocido como transducción de señal. La fosforilación y desfosforilación de proteínas son parte de la transducción del mensaje. Esto significa que la mayoría de los mensajes que se transmiten al interior de las células producen la activación de cinasas que regulan su función mediante la fosforilación de diversos sustratos. Los caminos de transducción pueden ser lineales, pero también pueden existir puntos potenciales de regulación y de intersección entre distintos caminos, por consiguiente, la respuesta celular está conformada por un circuito complejo de activación que involucra más de un camino de señalización. Así, una molécula señal puede ser reconocida por diferentes receptores asociados a diferentes mecanismos de transducción, también podría ocurrir que un mismo receptor active diferentes vías. Cuando un ligando interactúa con su receptor de membrana, la señal es transmitida al interior de la célula, generando una cascada de eventos que incluye la síntesis de segundos mensajeros y la fosforilación de enzimas catalizadas por proteincinasas, por tanto, estos segundos mensajeros son moléculas pequeñas que se producen en gran cantidad y rápidamente en respuesta a la activación de un receptor, llevando la señal a otras partes de la célula, amplificandola mediante la activación de cinasas y otras enzimas [27].

En cuanto al crecimiento celular, éste comienza por la unión de un factor de crecimiento a un receptor específico; las proteínas del receptor pueden estar situadas en la superficie de la célula diana (cualquier célula en la cual una hormona se une a su receptor) o encontrarse en el citoplasma o en el núcleo, por tanto, la respuesta celular

depende de la especificidad que posee el receptor para unirse a determinados ligandos. En la superficie celular hay tres tipos de receptores importantes para el crecimiento celular, al unirse a su respectivo ligandos emiten señales hacia el núcleo por distintas vías. Los principales tipos de receptores son [28]:

- Receptores con actividad intrínseca cinasa: tiene una región extracelular para la unión al ligando, una región a cada lado de la membrana y una región que puede tener actividad tirosina cinasa. La mayoría de los factores de crecimiento tienen receptores tirosina cinasa. La unión con un factor de crecimiento hace que el receptor se dimerice y autofosforile los residuos de tirosina.
- Receptores sin actividad catalítica intrínseca: poseen una porción extracelular de unión al ligando; una sola región que atraviesa la membrana y una porción que se asocia directamente y activa la tirosina cinasas.
- Receptores ligados a proteínas G: no están directamente vinculados con la regulación del crecimiento celular, consta de siete unidades que atraviesan la membrana y pertenecen los receptores de ciertas hormonas (adrenalina y glucagón).

En general, las células de todo organismo están en permanente intercomunicación a través de diferentes tipos de señales (citocinas, hormonas, factores de crecimiento, entre otros) que regulan la proliferación, diferenciación, metabolismo, comportamiento celular e inclusive apoptosis. Estas señales interaccionan con sus respectivos receptores en las células diana produciendo la activación de diferentes rutas específicas de transducción de señales, por lo que serán transmitidas hacia el interior de la célula para generar la respuesta adecuada. Cada señal es transmitida a través de múltiples rutas, o cascadas de señalización, en la que intervienen numerosas proteínas que ganan o pierden su actividad biológica mediante diversas modificaciones como fosforilación (adición de un grupo de fosfato a una molécula), desfosforilación, entre otras [34]. Además, los genes que codifican las proteínas implicadas en las vías de transducción de señales se designan como “proto-oncogenes” [29].

Dentro de las cascadas de señalización intracelular más importantes destaca la vía PI3K/AKT/mTOR, ésta juega un papel importante en el crecimiento y supervivencia

celular. Las fosfatidilinositol-3-kinasas (PI3Ks) son una familia de cinasas lipídicas que producen una cascada de señal intracelular que regula un amplio abanico de procesos celulares y que suele estar activada en distintos tumores, interviniendo en la proliferación incontrolada, resistencia a la apoptosis y metástasis, por consiguiente, la activación anormal de ésta resulta en la alteración de los mecanismos de control de crecimiento y supervivencia celular, lo que ocasiona el crecimiento competitivo celular y frecuentemente una mayor resistencia a los tratamientos, por tanto, su estudio es parte crucial para el entendimiento de los procesos de carcinogénesis. La activación de esta vía se puede producir por distintas alteraciones en diversos puntos como mutación de PI3K, mutación de la proteína cinasa B (AKT) o activación de los receptores tirosina cinasa, además de ello, varias alteraciones genéticas y rearrreglos cromosómicos pueden comprometer esta vía, generando su activación permanente [34], [35]. Así mismo, otra cascada relevante es la RAS/RAF/MEK/ERK también conocida como la vía MAPK / ERK (proteína cinasas activadas por mitógenos, originalmente llamado ERK, cinasas reguladas por señales extracelulares), controla una gran variedad de procesos intracelulares (crecimiento, proliferación, diferenciación, migración y apoptosis) en respuesta a los cambios que ocurren en el medio extracelular; la activación de esta vía se puede producir por medio de los receptores de tirosina cinasa o por mutaciones en los componentes intracelulares que concluyen en la fosforilación de la cinasa reguladora extracelular (ERK) [34], [37].

En relación al EGFR, éste pertenece a una familia de receptores tirosina cinasa que consta de cuatro receptores [33]:

- Receptor 1 del factor de crecimiento epidérmico EGFR ó ERBB1: activa generalmente la vía MAPK/ERK, en algunas circunstancias también puede activar la vía PI3K/AKT. Participa en el desarrollo de las estructuras epiteliales del sistema respiratorio, el tracto gastrointestinal, el tejido epidérmico y los folículos pilosos.
- Receptor 2 del factor de crecimiento epidérmico ERBB2 ó HER2/neu: activa generalmente la vía MAPK/ERK. Tiene un papel fundamental en el desarrollo del corazón y estructuras neuronales.

- Receptor 3 del factor de crecimiento epidérmico ERBB3 ó HER3: activa muy eficientemente la vía PI3K/AKT, es el único de la familia que carece de actividad cinasa. Promueve el desarrollo de la cresta neuronal.
- Receptor 4 del factor de crecimiento epidérmico ERBB4 ó HER4: tiene un papel fundamental en el desarrollo del corazón y estructuras neuronales.

Éstos receptores regulan la expresión de genes implicados en la proliferación y supervivencia celular a través de la vía de las MAPK/ERK y la vía PI3K/AKT, la actividad en que ellos se centran es en la mediación y regulación de las señales del medio externo hacia el interior de las células. Además, son proteínas ancladas en la membrana plasmática, que comparten una estructura similar. Constan de una región extracelular que se divide en cuatro dominios denominados I, II, III Y IV; donde los dominios I y III participan en la unión al ligando y los dominios I y II están involucrados en la dimerización entre receptores, a su vez constan de un dominio transmembrana (anclado a la membrana) y un dominio intracelular con actividad tirosina cinasa (transducción de señales). Estos receptores forman heterodímeros con ERBB2, existiendo como monómeros (una sola molécula) inactivos, que se dimerizan después de la activación por su ligandos, por tanto, la unión de ambos produce un cambio conformacional que permite formar una molécula e iniciar una cascada de eventos intracelulares a través de las vías ya mencionadas. Los ligandos se clasifican de acuerdo a los receptores que se unen específicamente, el EGF, epirregulina, anfirregulina y TGF α son ligandos del EGFR, mientras que las neuregulinas (NRG) son los ligandos de ERBB3 y ERBB4. Únicamente los receptores EGFR y ERBB4 son completamente funcionales en términos de unión al ligando. ERBB3 presenta una incompleta actividad cinasa y requiere heterodimerización para su total activación. ERBB2 es incapaz de unirse al ligando pero potencia la actividad cinasa de los heterodímeros.

En resumen, la activación del EGFR se produce tanto por sobreexpresión del receptor como por mecanismos dependientes e independientes del ligando, es decir, se activa al interactuar con sus ligandos o en algunos casos con otro miembro de la familia ERBB, lo que implica la transición desde una forma monomérica inactiva a una forma homodimérica activa, esto genera el inicio del proceso de transducción de señales. Además, la expresión anómala (sobreexpresión, activación constitutiva o

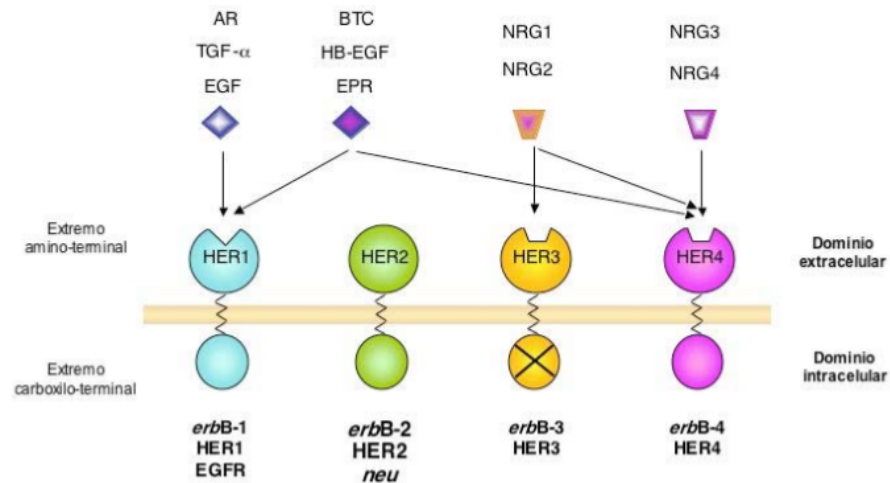


Figura 1.1: Miembros de la familia ERBB ó HER y especificidad de sus ligandos.

mutaciones) de estos receptores, especialmente EGFR y ERBB2, está implicada en el desarrollo y malignidad de numerosos tipos de cáncer [30], [36], [29]. Por lo tanto, el receptor del factor de crecimiento epidérmico y los otros miembros de su familia son receptores que juegan papeles esenciales tanto en condiciones fisiológicas normales como en condiciones cancerosas [33].

1.2.1 Escalas biológicas

Las escalas biológicas donde puede desarrollarse el proceso evolutivo del carcinoma de pulmón son: molecular, microscópica y macroscópica correspondiente a átomos, células y órganos respectivamente. Puesto que el cáncer es un sistema complejo, su simulación no debe limitarse a una escala biológica específica, ya que se omitirían aspectos importantes y esenciales para el estudio del mismo, de modo que son seleccionadas tanto la *escala molecular* como la *escala microscópica* por los aspectos que se destacarán a continuación:

En la escala molecular se puede modelar algunos de los siguientes aspectos:

- Interacción receptor-ligando (concentración de oxígeno y otros nutrientes en las células).

- Procesos celulares de señalización ya mencionados.
- Concentración de las moléculas.

Del mismo modo, el modelo propuesto denominado vía de señalización molecular se sitúa a este nivel, más adelante se observará a mayor detalle la implementación de una vía de transducción de señales, específicamente la correspondiente a las proteínas EGFR -ERK, fosfolipasa C (PLC), proteína cinasa C (PKC) entre otras más, constando exactamente de 20 proteínas. La forma de modelar que se realiza a nivel molecular es a través del uso de EDO, éstas sirven para describir cambios de concentración y de velocidad en las moléculas, por tanto, es necesario contar con 20 EDO para representar cada una de las proteínas. Habría que decir también que, tanto la proteína PLC como la ERK, son de gran importancia para la determinación de cambios fenotípicos de las células, mediante la evaluación de su dinámica se decide el proceso que va a seguir la célula, es decir, si permanece en el mismo sitio, prolifera o migra. Por consiguiente, el modelo contempla cuatro fenotipos de células tumorales: la proliferación, la migración, la inactividad y la muerte. Además, es importante mencionar que es necesario el uso de algunos parámetros cinéticos para observar los resultados proporcionados por la simulación, es decir, las concentraciones y velocidades iniciales de las proteínas que conforman el camino de señalización.

En cuanto a la escala microscópica es importante seleccionarla, en vista de que las condiciones del microentorno donde están dispuestas las células pueden inducir a la transformación de una de éstas en una célula maligna. Algunos de los siguientes aspectos que dan lugar en este nivel son los siguientes:

- Interacción entre las células (proximidad con sus vecinos).
- Permitir ver una mutación de células sanas en fenotipos cancerosos.

El modelado de esta escala es basado en agentes, donde a partir de algunas reglas biológicas definidas se puede representar cada célula de forma individual, centrándose primordialmente en las interacciones con su entorno [24]. La dinámica del crecimiento del tumor será estudiada en una rejilla discreta de dos dimensiones, donde las células responden a señales bioquímicas del entorno y de esta manera determinan su fenotipo

en cada paso del tiempo. Es decir, las señales químicas como la glucosa, la tensión de oxígeno y el EGFR son distribuidos a lo largo del entorno, de acuerdo a métodos probabilísticos.

1.3 Planteamiento del Problema

En el trabajo de Wang y sus colaboradores como base del presente proyecto, consta de las escalas: molecular y microscópica. La primera es descrita a través de EDOs, mientras que la segunda hace uso de autómatas celulares y un algoritmo de decisión fenotipo celular, en consecuencia, constituyen dos modelos en escalas espacio-tiempo distintas que se quieren implementar usando GALATEA, buscando la manera de integrar ambas escalas y haciendo de ellas una sola simulación. Por lo tanto, se quiere construir un primer prototipo, de modo que permita evaluar el modelo y además, ver el desempeño de la plataforma de simulación aplicada a procesos biológicos moleculares.

1.4 Justificación

En primer lugar, la comprensión de la dinámica del cáncer de pulmón a través de la modelización y simulación. En segundo lugar, la construcción de una herramienta genérica en el software GALATEA, cuya finalidad estudiar y entender problemas biológicos moleculares, usando ABM y autómatas celulares como herramientas de desarrollo.

1.5 Objetivos

Acorde con lo mencionado hasta este momento, los objetivos que se persiguen con la realización de este proyecto son los siguientes:

1.5.1 Objetivo General

Implementar en el software GALATEA un modelo de vía de señalización molecular, a través de las técnicas que proporcionan los modelos basados en agentes, para describir el

crecimiento del cáncer pulmonar tanto en la escala molecular como en la microscópica.

1.5.2 Objetivos Específicos

- Revisar bibliografía relacionada con el uso del modelado basado en agentes y autómatas celulares, específicamente en el cáncer pulmonar.
- Entender los procesos de señalización molecular: proliferación, invasión, metástasis, migración e inhibición de apoptosis, transducción de señales y su relación con el factor de crecimiento epidérmico como activador de células claves.
- Comprender los modelos de ecuaciones diferenciales ordinarias establecidas en el artículo, herramienta para generar el modelo molecular.
- Desarrollar un conjunto de reglas que caracterice el comportamiento de las células como autómatas celulares.
- Desarrollar un primer prototipo de simulación en la plataforma GALATEA.

1.6 Metodología

En cuanto a la metodología a llevar a cabo en el presente proyecto, es necesario recalcar que, para estudiar los aspectos concernientes al cáncer pulmonar a través de autómatas celulares, se va a implementar el modelado empleado en el artículo “Simulating non-small cell lung cancer with a multiscale agent-based model”, [13]. Avanzando en nuestro razonamiento, se va a describir detalladamente cada una de las etapas que constituyen al mismo en los siguientes párrafos:

Modelo de vía de señalización molecular.

El modelo de la implementación de la vía de señalización molecular para el cáncer de pulmón de células no pequeñas, está conformado por 20 moléculas incluyendo el EGFR, las cinasas no receptoras (PLC_γ y la proteína Cinasa C (PKC)), Raf, la vía de proteínas cinasas activadas por mitógenos MEK, y ERK. La selección de estas proteínas se justifica en que juegan un papel importante en el proceso evolutivo del cáncer

pulmonar, hecho que ha sido comprobado con diversos estudios a nivel experimental y además, dado que sus parámetros cinéticos pueden ser encontrados en la literatura con mayor facilidad a diferencia de otras moléculas. Entre estas proteínas, tanto la *PLC_y* como la ERK son de particular interés para la determinación de cambios fenotípicos de las células como se observará más adelante. Cada proteína es representada a través de ecuaciones diferenciales ordinarias, denominadas para el presente estudio como ecuaciones cinéticas, siendo escritas en términos de concentraciones y éstas en función de las velocidades de reacción. Las ecuaciones para diferentes moléculas se calcularon a través de la siguiente ecuación:

$$\frac{dX_i}{dt} = \sum Velocidad(producto) - \sum Velocidad(consumo) \quad (1.1)$$

Donde X_i representa cada uno de los 20 componentes moleculares que conforman el camino de señalización. La ecuación describe el cambio en la concentración de la molécula X_i , siendo el resultado de la diferencia entre la velocidad de reacción producida y la velocidad de reacción consumida. Cada reacción bioquímica se caracteriza por el diagrama que se presentará a continuación (figura 1), donde se aprecia constantes de velocidad tanto de avance como de retroceso. Además, las flechas representan las reacciones especificadas en las tablas 1 y 2, éstas resumen los parámetros cinéticos y las EDO utilizadas en el modelo.

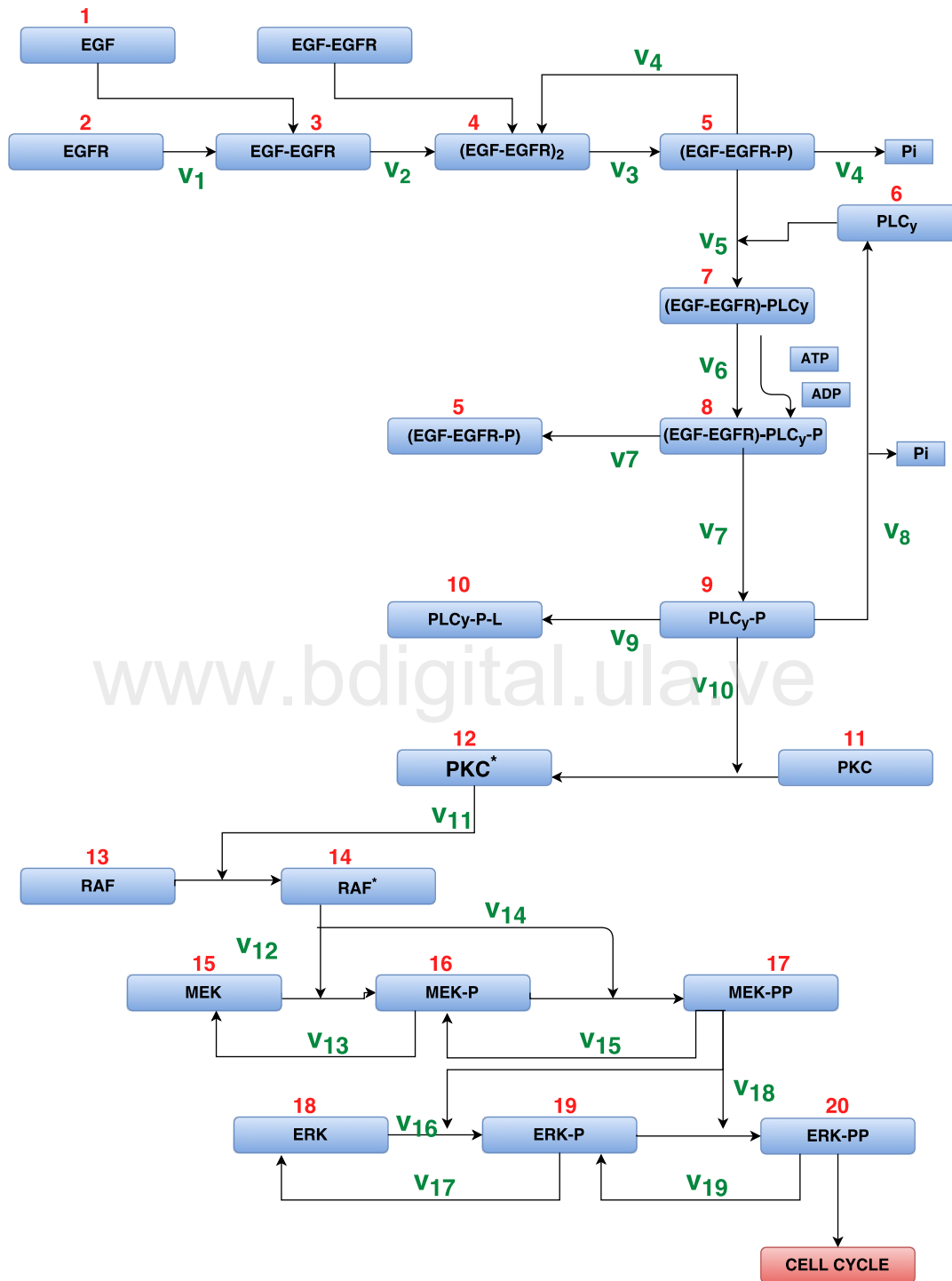


Figura 1.2: Modelo cinético de la vía de señalización.

Reactante	Ecuaciones Cinéticas		
	Var. Molecular	[.] ₀	EDO
X_1	EGF	Variada	$\frac{dX_1}{dt} = -V_1$
X_2	EGFR	80	$\frac{dX_2}{dt} = -V_1$
X_3	EGF-EGFR	0	$\frac{dX_3}{dt} = V_1 - 2V_2$
X_4	$(EGF - EGFR)_2$	0	$\frac{dX_4}{dt} = V_2 + V_4 - V_3$
X_5	$EGF - EGFR - P$	0	$\frac{dX_5}{dt} = V_3 + V_7 - V_4 - V_5$
X_6	PLC_y	10	$\frac{dX_6}{dt} = V_8 - V_5$
X_7	$EGF - EGFR - PLC_y$	0	$\frac{dX_7}{dt} = V_5 - V_6$
X_8	$EGF - EGFR - PLC_y - P$	0	$\frac{dX_8}{dt} = V_6 - V_7$
X_9	$PLC_y - P$	0	$\frac{dX_9}{dt} = V_7 - V_8 - V_9 - V_{10}$
X_{10}	$PLC_y - P - L$	0	$\frac{dX_{10}}{dt} = V_9$
X_{11}	PKC	10	$\frac{dX_{11}}{dt} = -V_{10}$
X_{12}	PKS^*	0	$\frac{dX_{12}}{dt} = V_{10} - V_{11}$
X_{13}	Raf	100	$\frac{dX_{13}}{dt} = -V_{11}$
X_{14}	Raf^*	0	$\frac{dX_{14}}{dt} = V_{11} - V_{12} - V_{14}$
X_{15}	MEK	120	$\frac{dX_{15}}{dt} = V_{13} - V_{12}$
X_{16}	$MEK - P$	120	$\frac{dX_{16}}{dt} = V_{12} + V_{15} - V_{13} - V_{14}$
X_{17}	$MEK - PP$	0	$\frac{dX_{17}}{dt} = V_{14} - V_{15} - V_{16} + V_{18}$
X_{18}	ERK	100	$\frac{dX_{18}}{dt} = V_{17} - V_{16}$
X_{19}	MEK	0	$\frac{dX_{19}}{dt} = V_{16} + V_{19} - V_{17} + V_{18}$
X_{20}	MEK	0	$\frac{dX_{20}}{dt} = V_{18} - V_{19}$

Tabla 1.1: Ecuaciones y concentraciones iniciales.

Ecuaciones Cinéticas			
Nro. de Reacción	Ecuación	Parámetros cinéticos	
V_1	$K_1 \cdot X_1 \cdot X_2 - K_{-1} \cdot X_3$	$K_1 = 0.003$	$K_{-1} = 0.06$
V_2	$K_2 \cdot X_3 \cdot X_3 - K_{-2} \cdot X_4$	$K_2 = 0.01$	$K_{-2} = 0.1$
V_3	$K_3 \cdot X_4 - K_{-3} \cdot X_5$	$K_3 = 1$	$K_{-3} = 0.01$
V_4	$\frac{V_4 \cdot X_5}{(K_4 + X_5)}$	$V_4 = 450$	$K_4 = 50$
V_5	$K_5 \cdot X_5 \cdot X_3 \cdot X_6 - K_{-5} \cdot X_7$	$K_5 = 0.06$	$K_{-5} = 0.2$
V_6	$K_6 \cdot X_7 - K_{-6} \cdot X_8$	$K_6 = 1$	$K_{-6} = 0.05$
V_7	$K_7 \cdot X_8 - K_{-7} \cdot X_5 \cdot X_9$	$K_7 = 0.3$	$K_{-7} = 0.006$
V_8	$\frac{V_8 \cdot X_9}{(K_8 + X_9)}$	$V_8 = 1$	$K_8 = 100$
V_9	$K_9 \cdot X_9 - K_{-9} \cdot X_{10}$	$K_9 = 1$	$K_{-9} = 100$
V_{10}	$K_{10} \cdot X_9 \cdot K_{11} - K_{-10} \cdot X_{12}$	$K_{10} = 0.214$	$K_{-10} = 5.25$
V_{11}	$\frac{V_{11} \cdot X_{12} \cdot X_{13}}{(K_{11} + X_{13})}$	$V_{11} = 4$	$K_{11} = 64$
V_{12}	$\frac{V_{12} \cdot X_{14} \cdot X_{15}}{[K_{12} \cdot (1 + \frac{X_{16}}{K_{14}}) + X_{15}]}$	$V_{12} = 3.5$	$K_{12} = 317$
V_{13}	$\frac{V_{13} \cdot X_{16}}{[K_{13} \cdot (1 + \frac{X_{17}}{K_{15}}) + X_{16}]}$	$V_{13} = 0.058$	$K_{13} = 2200$
V_{14}	$\frac{V_{14} \cdot X_{14} \cdot X_{16}}{[K_{14} \cdot (1 + \frac{X_{15}}{K_{12}}) + X_{16}]}$	$V_{14} = 2.9$	$K_{14} = 317$
V_{15}	$\frac{V_{15} \cdot X_{17}}{[K_{15} \cdot (1 + \frac{X_{16}}{K_{13}}) + X_{17}]}$	$V_{15} = 0.058$	$K_{13} = 60$
V_{16}	$\frac{V_{16} \cdot X_{17} \cdot X_{18}}{[K_{16} \cdot (1 + \frac{X_{19}}{K_{18}}) + X_{18}]}$	$V_{16} = 9.5$	$K_{16} = 1.46 \cdot 10^5$
V_{17}	$\frac{V_{17} \cdot X_{19}}{[K_{17} \cdot (1 + \frac{X_{20}}{K_{19}}) + X_{19}]}$	$V_{17} = 0.3$	$K_{17} = 160$
V_{18}	$\frac{V_{18} \cdot X_{17} \cdot X_{19}}{[K_{18} \cdot (1 + \frac{X_{18}}{K_{16}}) + X_{19}]}$	$V_{18} = 16$	$K_{18} = 1.46 \cdot 10^5$
V_{19}	$\frac{V_{19} \cdot X_{20}}{[K_{19} \cdot (1 + \frac{X_{19}}{K_{17}}) + X_{20}]}$	$V_{19} = 0.27$	$K_{19} = 60$

Tabla 1.2: Parámetros cinéticos.

Microentorno.

Como se ha mencionado en párrafos anteriores, los agentes son una generalización de los autómatas celulares. Por lo tanto, la estructura de un modelo discreto de autómatas celulares está basado en una región del espacio discreta (rejilla, cuadrada), donde coexisten determinadas especies celulares. La dinámica de crecimiento está definida por algunas reglas de interacción a lo largo de los nodos y entre los vecinos de cada celda. Cada punto de la rejilla puede representar una célula individual o un grupo de células. En una red de autómatas el estado de una célula dentro de la región depende de las reglas definidas, del estado de los vecinos y de su propio estado en la generación anterior. Un ejemplo básico de autómatas celulares es una rejilla en dos dimensiones, con puntos nodales X_{ij} donde i y j son las dimensiones de la cuadrícula. Se empieza con una población inicial, que se actualiza en cada paso de tiempo bajo unas determinadas reglas de muerte y reproducción [10]. En definitiva, el entorno consiste en una rejilla cuadrada de dos dimensiones, usándose la notación $p(i, j)$ para expresar cada punto en la red en términos Euclidianos. Habría que decir también que, el inicio de la simulación se da en el centro de la red por un conjunto de células.

Por otra parte, es necesario recordar que las células cancerosas compiten en el organismo por los nutrientes como el oxígeno, glucosa y otros presentes en éste, por tanto, donde exista mayor valor de éstos será mas atractivo para las células tumorales. Por consiguiente, son seleccionadas tres señales químicas ha emplearse en el modelo: EGF, la glucosa y la tensión de oxígeno. Haciendo uso de la distribución de probabilidades normal, cada punto del plano se le asigna un perfil de concentración de las señales mencionadas, donde los niveles de estas distribuciones son ponderados por la distancia, d_{ij} , de un punto de la celda dada a la fuente de nutrientes. Así, la distribución de estas tres señales se describe a través de las siguientes ecuaciones:

$$EGF^{ij} = T_m \cdot \exp \frac{-2d_{ij}^2}{\sigma_t^2} \quad (1.2)$$

$$Glucose^{ij} = G_a + (G_m - G_a) \cdot \exp \frac{-2d_{ij}^2}{\sigma_g^2} \quad (1.3)$$

$$Oxygen^{ij} = O_a + (O_m - O_a) \cdot \exp \frac{-2d_{ij}^2}{\sigma_o^2} \quad (1.4)$$

Por otra parte, las tres señales quimiotácticas continúan la difusión sobre la red a través de todo el proceso de simulación con una tasa fija, utilizando la siguiente ecuación:

$$\frac{\partial M^{ij}}{\partial t} = D_M \cdot \nabla^2 \cdot M^{ij}, t = 1, 2, 3 \quad (1.5)$$

Donde M representa una de las 3 señales externas, y t representa un paso de tiempo. Por otra parte, los coeficientes de las ecuaciones mostradas (2-5) son descritos en la tabla 3. Por lo tanto, mientras más cerca se encuentre la fuente de nutrientes de un lugar que se desee estudiar, más altos son los niveles de estas 3 señales en ese punto, por consiguiente, las células toman continuamente de glucosa para apoyar su metabolismo.

Coeficientes de distribución (EGF, glucosa y O_2)			
Coefficiente	Valor	Unidades	Descripción
T_m	2.56	Nm	Máxima concentración de EGF
G_a	17.0	mM	Concentración normal de la glucosa
G_m	57.0	mM	Concentración máxima de la glucosa
O_a	0.0017	DC	Concentración normal del oxígeno
O_m	0.0025	DC	Concentración máxima del oxígeno
D_{EGF}	6.7×10^{-11}	$m^2 \cdot s^{-1}$	Coefficiente de difusión de EGF
$D_{Glucose}$	5.18×10^{-11}	$m^2 \cdot s^{-1}$	Coefficiente de difusión de la glucosa
D_{Oxygen}	8.0×10^{-9}	$m^2 \cdot s^{-1}$	Coefficiente de difusión del Oxígeno

Tabla 1.3: Coeficientes de difusión de la glucosa, O_2 y EGF.

Algoritmo de decisión fenotipo celular

Acorde con lo mencionado hasta este momento, 4 fenotipos de células tumorales son consideradas para el modelo: la proliferación, la migración, la inactividad y la muerte. Por tanto, la muerte celular se produce cuando la concentración de la glucosa sobre el sitio cae por debajo de 8 mM, a su vez una célula se vuelve inactiva cuando la concentración de esta misma señal se encuentra en un rango de 8-16 mM, cuando no cumple las condiciones para migrar o dividirse (proliferar) ó cuando no se puede encontrar una ubicación vacía para que ésta pueda moverse o proliferar. Así, los dos rasgos fenotípicos más importantes para la expansión espacio-temporal (migración, proliferación) se deciden mediante la evaluación dinámica de las siguientes moléculas intracelulares:

1) La proteína PLC_y por estar involucrada en la dirección de movimiento de las células en respuesta al factor de crecimiento epidérmico (EGF), uno de los rasgos más relevantes, la dinámica acelerada que presenta durante la migración de las células malignas. Por lo tanto, en este modelo la velocidad de cambio (rate of change) ROC_{PLC_y} decide si una célula procede a migrar o no. Es decir, si ROC_{PLC_y} supera un cierto límite (T_{PLC_y}), la célula tiene potencial para migrar.

2) Análogamente a la condición anterior, la velocidad de cambio de la molécula ERK (ROC_{ERK}) decide si una célula procede con la proliferación, se ha encontrado experimentalmente que esta molécula tiene una fuerte influencia sobre la proliferación celular y además, la activación transitoria de ERK con EGF conduce a la replicación celular.

De manera que el algoritmo describe la consecuencia que tendría cualquier célula en el proceso de cambiar de fenotipo al caer en una de estas 4 condiciones:

- Si una célula experimenta la condición (1), el resultado es que no hay cambio fenotípico y tanto el ROC_{PLC_y} , como el ROC_{ERK} permanecen iguales.
- Si una célula experimenta la condición (2), conlleva a que la célula migra debido a que ROC_{PLC_y} es superior a su límite.
- Si una célula experimenta la condición (3), conlleva a que la célula prolifera

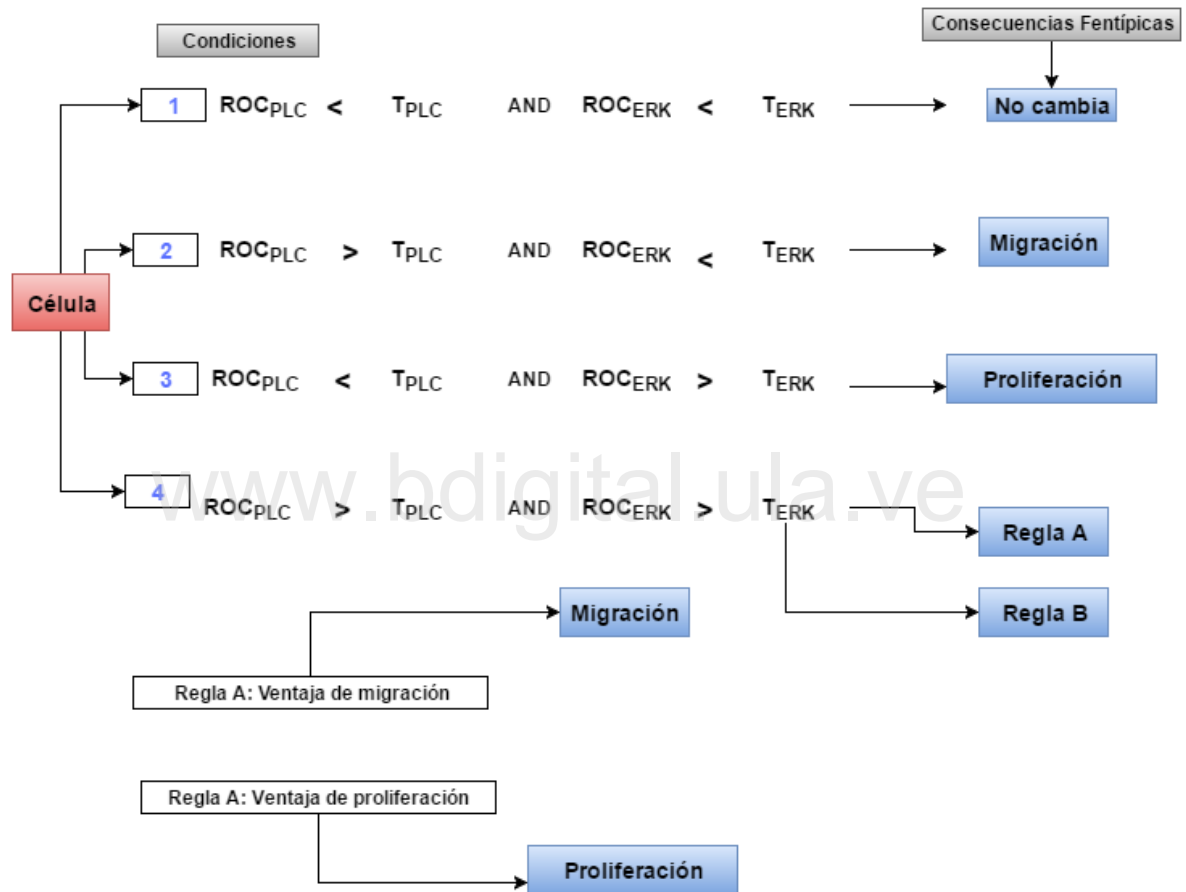


Figura 1.3: Algoritmo de decisión fenotípico celular.

debido a que ROC_{ERK} es superior a su límite .

- Para a condición (4), ambas moléculas sobrepasan sus límites y por consiguiente, se estudian dos hipótesis a lo que se le denomina normas especiales.

Los cambios en el número de células, localización e inclusive el entorno son factores importantes a tomar en cuenta para la modelización, éstos influirán en las fuentes de componentes moleculares en el camino de señalización. Al mismo tiempo la distribución espacial o concentración intracelular de las moléculas o especies de proteínas influirán en el destino de la célula. Razón por la cual, es fundamental tener una secuencia de pasos para poder implementar la relación entre los cambios moleculares y el fenotipo de las células.

1.7 Organización del documento

En cuanto al contenido del presente proyecto, éste sera distribuido en 4 capítulos:

- **Capítulo I:** constituido por la parte introductoria del trabajo, dicho de otra manera: historia, planteamiento del problema, objetivos de la investigación, justificación, metodología a seguir y la presente subsección.
- **Capítulo II:** en cualquier trabajo de investigación es fundamental tener ciertas bases teóricas claras, para posteriormente poder encontrar la respectiva solución al problema planteado. Dicho lo anterior, en esta parte del trabajo se hablará en relación a los siguientes aspectos: autómatas celulares como herramienta de simulación, un resumen de los aspectos fundamentales del artículo base del proyecto (Simulating non- small cell lung cancer with a multiscale agent-based model), la plataforma de simulación GALATEA; entre otros puntos que serán explicados a mayor detalle en los párrafos siguientes.
- **Capítulo III:** luego de comprender gran parte de las bases teóricas descritas en el capítulo anterior, se procede a la construcción del modelo sobre el software GALATEA. Por tanto, en el presente apartado se describirá detalladamente el proceso de solución al problema visto en la sección dos, es decir, resolución del

problema de bordes, selección de los pasos de tiempo, supuestos de simplificación, integración del modelo molecular dentro de las celdas, algoritmos propuestos, entre otros puntos. Es necesario recalcar que, en el proceso de realizar una simulación es necesario realizar supuestos de simplificación para disminuir la complejidad del problema.

- **Capítulo IV:** comprende de los resultados obtenidos en el proceso de simulación, comparandolos con los mostrados en el artículo científico base del presente proyecto. Además, se expondrá el aporte que genera la terminación del presente proyecto de grado.

www.bdigital.ula.ve

C.C. Reconocimiento

Capítulo 2

MARCO TEÓRICO Y CONCEPTUAL

2.1 Autómatas celulares (AC)

En el presente capítulo, se aborda el concepto de Autómatas Celulares y Modelo Basados en Agentes, esta compuesto por secciones en donde se describen los campos en donde pueden influir, como están definidos formalmente, se presentan sus antecedentes históricos, se detallan ejemplos clásicos de autómatas, se mencionan sus propiedades, también las reglas de transición utilizadas, entre otros puntos que se observarán a mayor detalle a continuación [42].

Los Autómatas Celulares son una clase de sistemas dinámicos discretos en espacio y tiempo, cuyas características los hacen un candidato idóneo para el estudio de sistemas con un alto grado de complejidad, es así que, han sido empleados en una gran variedad de campos como lo son: física, biología, sistemas de cómputo y la medicina [40]. Por ejemplo, se pueden mencionar los modelos de crecimiento, interacciones genéticas, interacciones sociales, patrones de reconocimiento, procesos de enfermedades, siendo tan solo una parte de sus aplicaciones [11].

2.1.1 Definición de AC

Un AC es un objeto matemático que consiste en un conjunto de celdas, un conjunto finito de estados y un conjunto de reglas de transición [41].

2.1.2 Notación formal de un AC

Un AC se puede definir formalmente según los elementos que lo componen como una cuatro tupla (Γ, S, V, δ) donde:

- Γ : Esta representado por un espacio celular regular, para $\Gamma = \{c \in C^d\}$, en donde $d \in Z^+$ corresponde a la dimensión, cubriendo un espacio de n-dimensiones, cuenta con patrones idénticos y conexiones locales con otras células.
- S : Es el conjunto finito de los posibles estados de cada una de las celdas en Γ .
- V : Una colección local de células que definen la vecindad para cada $c \in \Gamma$ tal que $V_c = \{K_1, K_2, \dots, K_n \mid K_j \in \Gamma \text{ para } j=0, 1, 2, \dots, n\}$. El conjunto de células relacionado directamente a una célula.
- δ Sea la función de transición aplicada simultáneamente a las células que conforman la retícula $\delta = S^d \rightarrow S$, determina el proceso de evolución en un tiempo t para cada una de las células en la lattice [42].

2.2 AC en dos dimensiones

Un AC de dos dimensiones, está representado por un arreglo definido por las parejas de ordenadas $Z \times Z$, de manera que una célula está representada por un elemento del conjunto de estados en las posiciones (i, j) .

Von Neumann: Esta vecindad consiste en una célula central, sobre la cual se realizara la actualización, acompañada de las cuatro células horizontales y verticales adyacentes a la célula central, geográficamente podemos decir que se encuentran localizadas al Norte, Sur, Este y Oeste es decir ($\aleph = 4$) [11].

$$V_{C_{i,j}^{(t+1)}} = \{C_{i,j}^{(t)}, C_{i-1,j}^{(t)}, C_{i+1,j}^{(t)}, C_{i,j-1}^{(t)}, C_{i,j+1}^{(t)}\} \quad (2.1)$$

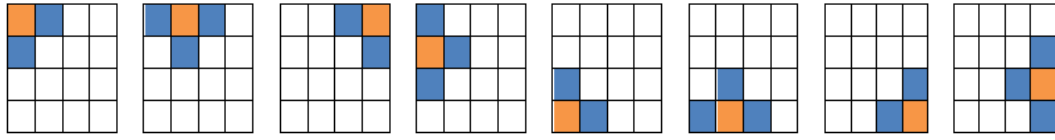


Figura 2.1: Vecinos Von Neumann en fronteras absorbentes.

2.2.1 Bordes en un AC

A continuación se determinan las reglas que definen las fronteras, y se muestran gráficamente siguiendo el mismo código de colores de la figura anterior [41].

Fronteras absorbentes

- 1. No hay vecino más cercano izquierdo de la primera celda de una fila.
- 2. No hay vecino más cercano derecho de la última celda de una fila.
- 3. No hay vecino más cercano de arriba de la primera celda de una columna.
- 4. No hay vecino más cercano de abajo de la última celda de una columna.

2.2.2 Reglas en un AC

Las reglas de un autómata celular se definen de acuerdo a lo que se desea modelar. El estado de una celda en el próximo paso de tiempo, depende del estado de ésta en el paso de tiempo actual, así como también de la vecindad considerada[41].

Diseñar y producir las reglas de los AC para que exhiban un comportamiento requiere conocer muy bien el problema objetivo, lo que implica dominar el conocimiento del problema a tratar [11].

2.3 Modelos basados en Agentes (ABM)

Los modelos basados en agentes son muy útiles en el estudio de sistemas complejos conformados por un gran número de entidades que interactúan con varios grados de libertad. También son referidos como Modelos Basados en Individuos (IBM), Sistemas

Multi-Agentes (MAS), Sistemas Multi Agentes Masivos (MMAS). En el presente documento se utilizará el acrónimo ABM, en general [11].

El concepto de ABM consiste en construir la contraparte computacional de un modelo conceptual de un sistema bajo estudio sobre la base de entidades discretas, denominados agentes, con algunas propiedades y reglas de comportamiento, para luego simularlos en un computador que represente el fenómeno real [11].

Los métodos basados en agentes facilitan el estudio y modelado de sistemas complejos a partir de las unidades que los componen, permitiendo construir modelos experimentales de la realidad desde un punto de vista diferente al tradicional: desde lo más simple hacia lo más complejo [45]. De esta manera, los ABM son herramientas útiles para estudiar el comportamiento que emerge de un sistema complejo como resultado de las características de los componentes del sistema y el funcionamiento de los mismos [44].

2.3.1 ABM Definición Modelo Basado en Agentes

Los modelos basados en agentes son una generalización de los autómatas celulares. Constituyen una técnica computacional destinada principalmente al estudio de los sistemas dinámicos, y son utilizados en campos tan diversos como la economía, la ingeniería, la química, la biología y la sociología. Estos modelos utilizan múltiples agentes que, obedeciendo ciertas reglas locales, interactúan de manera autónoma entre sí y con su entorno. De este modo, a partir de un comportamiento local, determinado por reglas simples, es posible estudiar el comportamiento global del sistema, estructuras emergentes, complejidad y otros. Los agentes son entidades computacionales que representan los actores reales, el entorno donde se mueven e interactúan los agentes y las reglas determinan el comportamiento de los mismos [14].

2.3.2 Propiedades de los ABM

Entre las propiedades que más se destacan, se encuentran la heterogeneidad, el espacio explícito y la interacción local. En la heterogeneidad, se selecciona un parámetro relevante por cada distribución, lo cual implica que unos pocos parámetros son

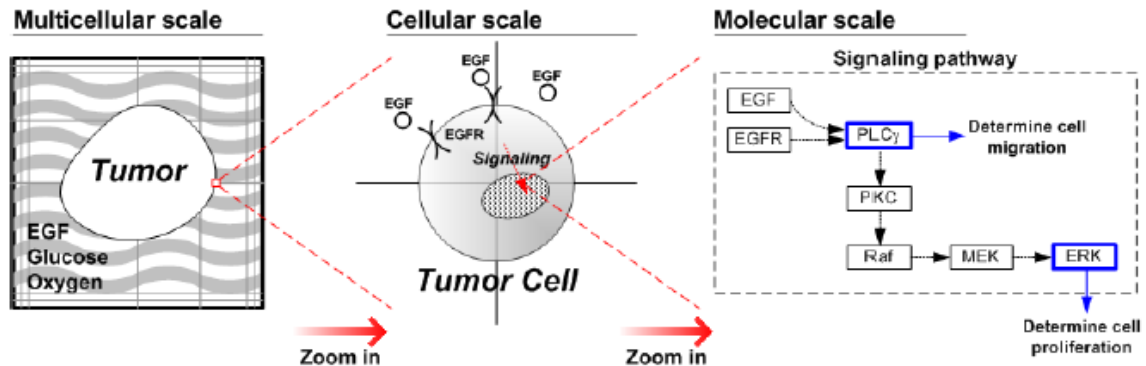


Figura 2.2: Características básicas de las tres escalas complejas

adicionados al modelo y son los que gobiernan la distribución. En el espacio explícito, con frecuencia los individuos difieren del lugar físico en donde se encuentran y/o de los vecinos con quienes pueden o tienen que interactuar. En la interacción local, los modelos asumen una interacción global o una interacción local muy simple [11].

2.4 “Simulating non-small cell lung cancer with a multiscale agent-based model ”.(Resumen).

Modelo multiescala que contempla por lo menos dos escalas para representar de una manera simplificada una realidad, fenómeno o como en el presente caso, una enfermedad. El estudio del cáncer pulmonar de células no pequeñas realizado por Zhihui Wang, Le Zhang, Jonathan Sagotsky y Thomas Deisboeck, contempla dos escalas: escala molecular y escala microscópica. Es importante acotar que también existe una tercera escala, denominada escala multicelular, como se puede apreciar en la figura 2.5 donde además se señalan las características más relevantes de estos niveles de biocomplejidad. Sin embargo, sólo seleccionaron dos escalas ya que la inserción de una tercera, ocasionaría mayor incertidumbre alrededor del mecanismo biológico. Fue elaborado en base a una variedad de modelos multiescala que se habían hecho para la zona del cerebro. No obstante, se hicieron las modificaciones necesarias,

ya que al ser tejidos diferentes implicaría una conformación molecular distinta, dicho en otras palabras, a nivel molecular significó la implementación de un camino de transducción de señales específico para el pulmón. Es necesario recalcar que al hablar de transducción de señales, se está refiriendo a una serie de eventos moleculares de transferencia de información que ocurren dentro de las células o entre ellas, lo que implica que relacionan los cambios externos con las respuestas internas que éstas generan. Así mismo, dicho camino consta de 20 especies moleculares representadas mediante Ecuaciones Diferenciales Ordinarias, de tal manera que conforman un sistema acoplado, escritas en términos de concentraciones y las velocidades de reacción, las cuales pueden ser apreciadas en la Tabla 1.1, a su vez en la Figura 1.2 pueden ser observadas cada una de las moléculas que intervienen en este proceso, teniendo principal interés sobre el Receptor del factor de crecimiento(EGFR), la proteína quinasa C (PKC)y las quinasas reguladoras de señales extra-celulares (ERK), ya que se ha comprobado experimentalmente que juegan un papel importante en la tumorigénesis del cáncer de pulmón.

La otra parte del modelo, corresponde a la parte microscópica o el estudio de las células, a diferencia del anterior que sólo se enfocaba en la interacción entre las moléculas. La dinámica del crecimiento del tumor puede ser investigada en un entorno de dos dimensiones con una rejilla discreta, donde cada punto de la cuadrícula puede ser ocupado por una sola celda o estar vacío. El oxígeno, la glucosa y el factor de crecimiento se difunden en cada paso del tiempo a través de las ecuaciones:

$$EGF^{ij} = T_m \cdot \exp \frac{-2d_{ij}^2}{\sigma_t^2} \quad (2.2)$$

$$Glucosa^{ij} = G_a + (G_m - G_a) \cdot \exp \frac{-2d_{ij}^2}{\sigma_g^2} \quad (2.3)$$

$$Oxigeno^{ij} = O_a + (O_m - O_a) \cdot \exp \frac{-2d_{ij}^2}{\sigma_o^2} \quad (2.4)$$

Se empieza con una matriz 7×7 de células malas en un espacio de 200×200 . Las concentraciones iniciales van a depender de la distancia de la fuente a la célula. En cuanto a la última parte del modelo, correspondiente al algoritmo de decisión fenotipo celular, se debe agregar que dependiendo de ciertos límites en los cuales deben permanecer la molécula **ERK** y **PLCy**, se decide si la célula procede a proliferar, permanecer inactivo o migrar. Identificar los elementos que constituyen el cáncer pulmonar de células no pequeñas ha hecho posible establecer los factores claves que deben ser considerados, en la construcción de reglas de comportamiento y de tal forma, adaptarlas a la lógica de herramientas computacionales como son los AC y los ABM, traduciendo de esta manera un comportamiento biológico a un lenguaje computacional capaz de interpretar reglas.

2.5 Ecuaciones reacción-difusión

La distribución espacial de ciertas propiedades inhomogéneas en un sistema, como lo son: concentraciones, parámetros de control, poblaciones de individuos, entre otras. Las propiedades espacio-temporales de *reacción-difusión*, aquellas que tienen en cuenta simultáneamente la variación temporal como consecuencia de las relaciones cinéticas y los procesos de difusión. En consecuencia, muchos de los fenómenos de auto-organización espacio-temporal observados en biología son modelados mediante sistemas de ecuaciones de reacción difusión [52].

2.5.1 Aproximaciones numéricas de Ecuaciones Diferenciales Parciales

Las ecuaciones de reacción-difusión no se pueden resolver de forma analítica, por lo que se hace imprescindible el uso de algún método numérico, para realizar aproximaciones a la solución real [50].

El objetivo que se persigue con la utilización de cualquiera de los métodos discretos, es obtener la solución en un conjunto reducido de puntos (P) de un dominio continuo (D). En la mayoría de los casos, el conjunto de puntos elegidos responde a un patrón

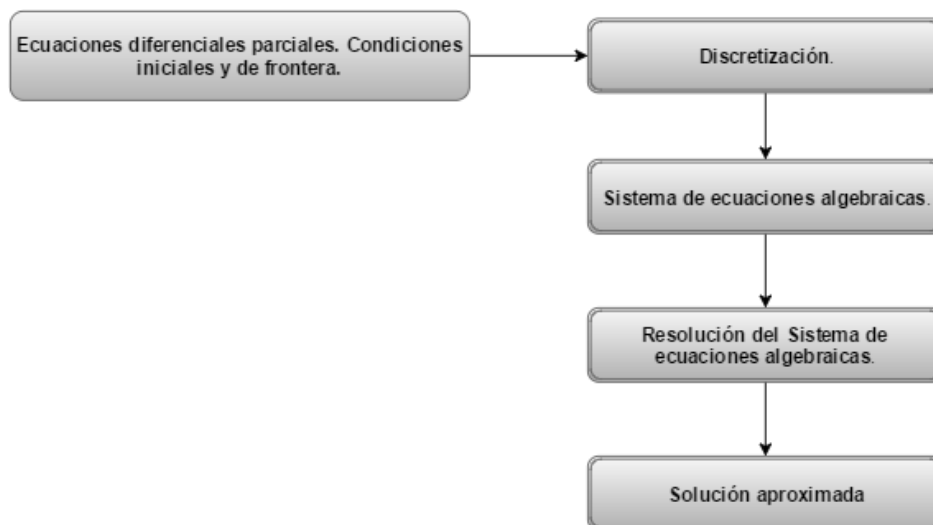


Figura 2.3: Solución aproximada para ecuaciones en derivadas parciales.

determinado, como puede ser una malla rectangular homogénea [50]. Las intersecciones de las líneas que constituyen ésta son denominados nodos y es en donde se calcula la solución numérica de la ecuación diferencial parcial [51].

La resolución numérica de una ecuación diferencial parcial utilizando el método conocido como Diferencias Finitas, es el reemplazo de las derivadas continuas de la ecuación diferencial por las expresiones equivalentes en diferencias finitas [51].

Una vez obtenida la solución discreta y mediante técnicas de interpolación, se

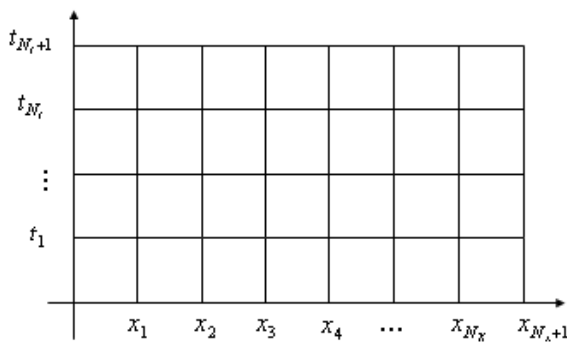


Figura 2.4: Reemplazo del dominio continuo del problema por una malla

pueden obtener los valores de las variables dependientes, sus derivadas y sus integrales en cualquier otro punto. Utilizando diferencias finitas, la aproximación a una derivada parcial se puede calcular con una sencilla expresión algebraica si simplemente se despeja el correspondiente término del desarrollo en serie de Taylor, truncado adecuadamente. Por ejemplo,

$$u(x + \Delta x) = u(x) + \Delta x \frac{\partial u}{\partial x}(x) + O(\Delta x^2) \quad (2.5)$$

de donde se obtiene

$$\frac{\partial u}{\partial x}(x) = \frac{u(x + \Delta x) - u(x)}{\Delta x} + O(\Delta x^2) \quad (2.6)$$

El *error de truncamiento local* se define como el error que se comete en cada punto (x_i, t_n) debido a la sustitución de un operador diferencial por el correspondiente operador de diferencias. La diferencia entre los desarrollos de Taylor para $u(x + \Delta x)$ y $u(x - \Delta x)$ permite obtener la siguiente aproximación con menor error de truncado:

$$\frac{\partial u}{\partial x}(x) = \frac{u(x + \Delta x) - u(x - \Delta x)}{2 \Delta x} + O(\Delta x^2) \quad (2.7)$$

Para introducir algunos conceptos relacionados con la resolución de las ecuaciones de difusión para el cáncer pulmonar, se considera la ecuación de difusión en su forma más simple:

$$\frac{\partial u}{\partial t} = D \cdot \nabla^2 = D \cdot \frac{\partial^2 u}{\partial x^2} \quad (2.8)$$

De tal manera que, utilizando una malla de N_x puntos para la discretización. En esta ecuación se considera la siguiente discretización del operador de la derecha (espacial), evaluada en un punto x_i y en el instante t_n :

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u_{i-1}^n - 2u_i^n + u_{i+1}^n}{\Delta x^2} \quad (2.9)$$

donde $u_i^n = u(x_i, t_n)$.

Para la derivada de u en t se puede optar por utilizar diferencias progresivas, regresivas o centrales, dependiendo de la evaluación temporal de los puntos elegidos para la discretización:

$$\frac{\partial u}{\partial t} \approx \frac{u_i^{n+1} - u_i^n}{\Delta t} \quad (2.10)$$

$$\frac{\partial u}{\partial t} \approx \frac{u_i^n - u_i^{n-1}}{\Delta t} \quad (2.11)$$

$$\frac{\partial u}{\partial t} \approx \frac{u_i^{n+1} - u_i^{n-1}}{2 \Delta t} \quad (2.12)$$

La combinación de *diferencias espaciales centradas* y *progresivas* se conoce como método de las diferencias progresivas de Euler o, simplemente, método de Euler.

Si se utiliza un método de diferencias progresivas, es fácil observar que la solución u_i^{n+1} se puede calcular en cualquier punto x_i simplemente despejando la ecuación de diferencias en dicho punto:

$$u_i^{n+1} = u_i^n + \frac{D \Delta t}{\Delta x^2} (u_{i-1}^n - 2u_i^n + u_{i+1}^n) \quad (2.13)$$

Como esto es válido para cualquier punto x_i , sólo será necesario realizar operaciones triviales en cada punto. Por lo tanto, se trata de una matriz de $N_x \times N_x$ elementos, donde cada fila \mathbf{i} de la matriz está asociada a la ecuación de diferencias evaluada en un punto x_i de la malla. En la discretización de la ecuación de difusión, los únicos valores de \mathbf{j} que se requieren realmente son $\mathbf{j} = (\mathbf{i}-1, \mathbf{i}, \mathbf{i}+1)$. El resto de los coeficientes de la matriz son nulos. Este tipo de discretización, en el que la ecuación de diferencias en cada punto se puede resolver de forma independiente, se denomina *método explícito*.

Si se utiliza un método de diferencias regresivas, la discretización de la ecuación de difusión que se obtiene es:

$$u_i^n - \frac{D \Delta t}{\Delta x^2} (u_{i-1}^n - 2u_i^n + u_{i+1}^n) = u_i^{n-1} \quad (2.14)$$

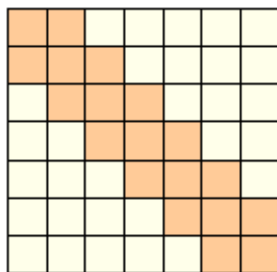


Figura 2.5: Matriz tridiagonal.

Aunque la ecuación anterior es muy parecida a la obtenida por diferencias progresivas, es muy diferente computacionalmente. En este caso, la ecuación en cada punto contiene tres incógnitas, por lo que no se puede resolver por separado, sino de forma conjunta para todos los puntos. Este tipo de discretización, en el que la ecuación de diferencias en cada punto no se puede resolver de forma independiente, se denomina *método implícito* [50].

Al aplicar la ecuación discretizada en cada punto de la malla se obtiene un sistema de ecuaciones denominado **sistema de ecuaciones de diferencias finitas**. El proceso de aproximación requiere de la selección de un método adecuado para obtener la solución del sistema de ecuaciones algebraicas planteado. Una vez resuelto el sistema se obtiene el valor de la función en los nodos de la malla, es decir, que al emplear el método de diferencias finitas se obtiene una solución aproximada discreta [51].

2.6 Proyecto GALATEA

La plataforma de simulación **GALATEA** pretende aprovechar el esfuerzo de las investigaciones realizadas en la Universidad de Los Andes (CESIMO -IEAC - SUMA) en las áreas de simulación, modelado y desarrollo de software. GALATEA surge como un proyecto para incorporar o extender la plataforma GLIDER con los conceptos y herramientas que permitan simular sistemas bajo los enfoques distribuido, interactivo, continuo, discreto y combinado. Igualmente, esta plataforma incorpora el soporte para modelado y simulación de sistemas multiagentes, así como también una serie de

módulos que permiten la interacción ente los modelos desarrollados con GALATEA y un conjunto de sistemas externos como por ejemplo, Base de Datos, Sistemas de Información Geográficos, Herramientas de Visualización entre otras.

GLIDER, lenguaje de simulación, desarrollado en la Universidad de Los Andes, incluye herramientas para el modelado y simulación de sistemas continuos, además presenta facilidades, en su diseño, para la ejecución simultánea y cooperativa de ambos tipos de sistemas.

Este lenguaje representa el sistema modelado, en subsistemas o nodos, que intercambian información, a través de mensajes. Individualmente un subsistema puede almacenar, transformar, transmitir, crear y eliminar información del estado del sistema, mediante un código enlazado, que precisa el cambio de los estados del sistema, cuando ocurre un evento en un subsistema.

2.6.1 Características de GALATEA

La plataforma de simulación GALATEA, está conformada por:

- Una familia de lenguajes que permiten simular sistemas multi-agentes. GALATEA no es un solo lenguaje para crear modelos monolíticos. Es una familia de lenguajes que incluye a una versión extendida de GLIDER, y a su vez varios lenguajes de programación lógica para modelar los agentes.
- Un compilador de GLIDER extendido, desarrollado en Java.
- Un ambiente para la construcción de modelos.
- Un motor de inferencia que funciona como interpretador independiente de los programas de cada agente.
- Un conjunto de módulos para la comunicación con sistemas externos. Para mayor información, se podría consultar en [15],[16] y [17].

2.6.2 Estructura de la biblioteca

Acorde con las características enunciadas, en la figura 2.9 se puede apreciar algunos de los componentes del software galatea. Por tanto, a continuación se explica muy brevemente la funcionalidad de algunos de estos:

- **gca**: Es la librería de Autómatas Celulares que posee el software.
- **ggui**: Su aplicabilidad es para el desarrollo de interfaces gráficas, con sus respectivos cuadros de dialogo de información.
- **glider**: Es el lenguaje para la simulación de eventos continuos y discretos.
- **glorias**: Enfocada para el desarrollo de modelos lógicos, por tanto, puede hacer uso del lenguaje de programación lógica mejor denominado como Prolog.
- **gode**: Es el módulo diseñado para la resolución de Ecuaciones Diferenciales Ordinarias. De tal forma que cuenta con algunos de los métodos más comunes e importantes, como lo son: Euler, Euler Modificado y Runge Kutta de cuarto orden.
- **gspaces**: Se centra en la simulación de agentes, limitados por espacios, así por ejemplo, se podría representar el desalojo de un edificio a través de este módulo.

Dicho lo anterior, la librería que tiene mayor relevancia en el presente proyecto es la “**gca**”, ya que representa una herramienta que posee el software para lograr la construcción del modelo. Por tanto, es fundamental explicar a mayor detalle su conformación, especialmente aquellos archivos con código fuente en el lenguaje de programación java, que constan de métodos que serán utilizados a lo largo del enlace de los los dos modelos, con la finalidad de darle al lector mayor comprensión al momento de revisar el presente trabajo. Así mismo, al igual que otras librerías que posee dicho lenguaje, posee constructores, métodos de consulta y métodos de asignación.

En definitiva, los archivos por los cuales esta constituido el módulo de AC son los siguientes:

1. CA.java: Integrado principalmente por los siguientes métodos:

- `CA(String name, Node ArrayList layers)`: Primer constructor. Crea una instancia del AC.
 - `CA(String name, Layer layer)`: Segundo constructor. Crea una instancia del AC con una capa.
 - `public NodeArrayList getLayerLst()`: Método de consulta. Lista de capas.
 - `void addLayer(Layer l)`: Su utilidad es agregar una capa a la lista de capas del AC.
 - `public Layer getLayer(String Name)`: Con este método se recorre la lista de capas, comparándola con el nombre que se pasa como parámetro. Al coincidir, devuelve la capa solicitada.
 - `public void act(int a)`: Indica la próxima activación del AC.
2. `CAOutputEvent.java`: Conformado fundamentalmente por los siguientes métodos:
- `Public String getLayerId(int i)`: Método de consulta. Retorna el id de la capa i.
 - `public int[] getLayerData(int i)`: Método de consulta. Retorna los datos de la capa i.
3. `Cell.java`. Es el archivo contenedor de la clase designada para representar las celdas o células. Los atributos de los cuales esta compuesta son:
- (a) Capa a la que pertenece (`Layer`).
 - (b) Identificador (`id`). Define la localización en el espacio.
 - (c) Un valor tomado a partir de un conjunto finito de estados definidos por la capa a la que pertenece la celda.
 - (d) Conjunto de vecinos: `ArrayList <cell>neighbordlist`.

Así mismo, también está constituida por los siguientes métodos:

- `public Cell(Layer L, int memSize)`: Constructor de la clase celda.

- `public int getActualState()`: Método de consulta. Retorna el estado actual de las celdas.
 - `public int getNeighborState(int i)`: Método de consulta. Retorna el estado del vecino `i` de la celda.
 - `public void setActualState(int s)`: Método de asignación. Por tanto, configura la celda `i` con el valor `s` en el estado actual.
 - `public void setNextState(int s)`: Método de asignación. Por consiguiente, configura la celda `i` con el valor de `s`, en el próximo paso del tiempo.
 - `public int getId()`: Método de consulta. Devuelve el identificador de la celda.
 - `rule()`: Regla que se le aplica a la celda.
4. `Layer.java`: Denota el archivo contenedor de la capa del AC, dicho de otra manera, corresponde al tablero o conjunto de tableros conformados por las celdas. Los atributos que posee dicha clase son los siguientes:
- (a) Un identificador de la capa (`id`).
 - (b) Dimensión de la capa (`1d`, `2d`, `3d`).
 - (c) Tamaño de la memoria de los estados.
 - (d) Un índice del estado actual.
 - (e) Conjunto de celdas de la capa.

A su vez, algunos de los métodos que integran esta clase son los siguientes:

- `public Layer(String id, int dimension, int memSize, double ts, boolean enableDisplay)`: Constructor de la capa, crea una instancia de `Layer`.
- `public int getSize()`: Método de consulta. Retorna el número de celdas del AC.
- `void setDimension(int d)`: Asigna una dimensión a la capa.
- `public int getNextStateIndex()`: Retorna el índice del estado siguiente de las celdas.

- `public void update()`: Actualizador de la capa.
 - `public Cell getCell(int i)`: Retorna la celda `i` perteneciente a la capa.
5. `Layer1D.java`: Clase capa para autómatas celulares de una dimensión. Posee métodos análogos a la clase `Layer`, ya que hereda de ésta.
 6. `Layer2D.java`: Clase capa para AC de dos dimensiones, hereda de la clase `Layer` también. Sin embargo, entre sus atributos se agregan:
 - (a) Número de filas.
 - (b) Ancho de las filas.
 - (c) Número de columnas.
 - (d) Ancho de las columnas.
 7. `MooreLayer.java`: Clase diseñada para capas, cuyas vecindades sea la de Moore.
 8. `NearestNeighbor1DLayer.java`: Clase para vecindades cuyo el número de vecinos, sea igual a dos.
 9. `VonNeumannLayer.java`: Clase diseñada para capas, cuyas vecindades sea la de Von Neumann.

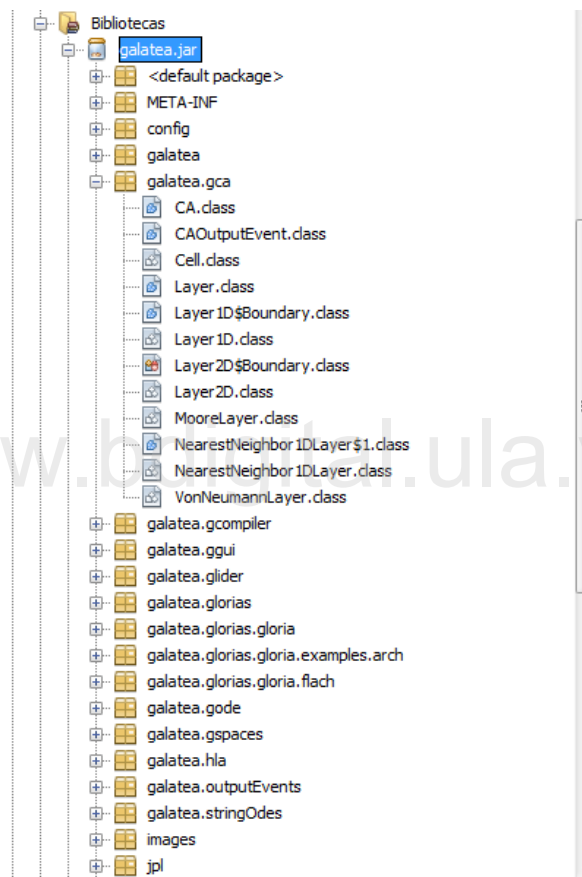


Figura 2.6: Estructura de la librería Galatea

Capítulo 3

MODELO DE ACOPLAMIENTO

Un modelo puede ser considerado como una representación simplificada de la realidad, en donde distintos componentes y procesos que forman parte del sistema en estudio se desean exhibir. El proceso de construcción de un modelo contribuye a identificar, seleccionar y ordenar la información disponible en relación al funcionamiento del sistema de estudio[44]. En el presente caso una enfermedad que se ha hecho muy cotidiana hoy en día, el cáncer pulmonar.

Avanzando en este razonamiento, en el presente capítulo se apreciará el proceso de construcción e implementación de un modelo de integración o acoplamiento, es decir, la conformación del nivel molecular y microscópico en una sola simulación, basados en la dinámica expuesta por Wang, ya mencionada en los apartados de los capítulos anteriores.

3.1 Supuestos de simplificación del modelo

En el proceso de construcción de un modelo es recomendable comenzar con pocos detalles, enfocarse en obtener un modelo sencillo que permita generar y analizar los resultados, de tal manera que se puedan introducir más detalles en las áreas que impacten más a los mismos [49].

Por otra parte, dado que el objetivo fundamental es generar un primer prototipo de simulación sobre la plataforma Galatea, fue necesario realizar ciertos supuestos sobre

el modelo acoplado, los cuales son los siguientes:

1. En un medio biológico cada célula individual tiene sus propias características y responde de cierta manera a los agentes externos que la rodean, siendo de distintos tamaños de acuerdo al tejido que éstas conforman [9]. Ahora bien, para el presente trabajo se considerará que todas poseen volúmenes homogéneos de 8 micras.
2. Se pueden tener constantes de tiempo distintas para la parte reactiva y la difusiva. Dicho de otra manera, la escala de tiempo a la cual se produce el modelo molecular o parte reactiva (involucra la resolución de las 20 EDOs), podría ser distinta al proceso difusivo de la glucosa, oxígeno e inclusive del receptor del factor de crecimiento. Por consiguiente, se asume que el proceso difusivo ocurre más rápido, que el ciclo celular y todo lo que conlleva éste. Dicho de otra manera, los procesos de proliferación, migración, muerte e inactividad celular ocurren posterior a realizarse el proceso de difusión.
3. Para la producción del **ADENOSINTRIFOSFATO (ATP)** compuesto indispensable en muchas de las reacciones que se llevan a cabo en la célula, se necesitan varios sustratos energéticos, entre los cuales la glucosa es el de mayor importancia, la cual debe ser transportada al interior de la célula[53]. De donde se infiere que la difusión que se da dentro de cada celda no es considerada en la presente simulación, ya que esta ocurriendo a una escala infinitamente más pequeña que la escala microscópica.
4. No se toman en cuenta los desechos durante el proceso difusivo, dicho en otras palabras, no se considera el dióxido de carbono.
5. Actualmente se está tratando de simular la parte básica del crecimiento de la célula en un tejido. De tal manera que la parte de vascularización no es considerada en el modelo. Por lo tanto, la activación de células cancerosas sólo ocurre en las cercanías o vecindario de la célula que se está estudiando.
6. Las células van naciendo y muriendo. Por consiguiente, la coordenada donde está dispuesta una célula que muere, es reemplazada por una célula sana.

7. No hay reglas para células sanas.
8. Dado que en el artículo base de este proyecto, no enuncia los niveles de oxígeno se hace la hipótesis que esta presente en todos lados y únicamente se tienen gradientes espaciales de glucosa. El oxígeno se difunde mucho más rápido que la glucosa, por tanto, es más trascendental ésta última. Sin embargo, debe estar presente para que la glucosa haga su proceso.

3.2 Algoritmo general propuesto para el diseño del modelo

Un algoritmo bien estructurado combina sentencias de alguna de las formas siguientes:

1. Sentencias sencillas
2. Decisión (if).
3. Bucles
4. Llamadas a procedimientos [54].

En este orden de ideas, se presenta un algoritmo general que involucra el enlace de los modelos molecular y microscópico. De tal manera que antes de proceder a la aplicación de las reglas de crecimiento celular, se da solución al problema de difusión. Además, la red está inicializada en el centro con una submatriz de 5×5 de celdas malignas. Por tanto, dentro de ésta se hallan los perfiles de concentración para cada una de las especies químicas ya mencionadas, por consiguiente, fuera de la misma permanecen los valores arrojados por las funciones de distribución radial con respecto a la fuente.

3.3 Resolución de la parte difusiva del modelo

3.3.1 Construcción del estado inicial del AC

Es el primer paso para la construcción del modelo que represente al AC, como se puede detallar al inicio del algoritmo presentado. Cada capa perteneciente al autómata es

Algoritmo 1 Algoritmo de integración del modelo molecular y microscópico

```

1: while tiempoDif < TasaNacimientoCeldas do
2:   Para tiempo = 0, Inicializar malla para cada capa con la distribución radial:
   Capa 0:  $Glucose^{ij} = G_a + (G_m - G_a) \cdot \exp \frac{-2d_{ij}^2}{\sigma_g^2}$ 
   Capa 1:  $EGF^{ij} = T_m \cdot \exp \frac{-2d_{ij}^2}{\sigma_t^2}$ 
   Capa 2:  $Oxygen^{ij} = O_a + (O_m - O_a) \cdot \exp \frac{-2d_{ij}^2}{\sigma_o^2}$ 
3:   if tiempo > 0 then
4:     Colocar matriz  $n \times n$  de células cancerígenas en el centro del Lattice
     Resolver el problema de bordes para dicha matriz
5:     if ConcentracionGlucosa < 8mM then
6:       La célula MUERE
7:     end if
8:     if TasaNacimiento == 8.064 minutos then
9:       Nace una célula Cancerígena
       Resolver las 20 EDOs
10:    if ((velocidadPLC < LimPLC AND velocidadERK < LimiteERK) OR
        (8mM ≤ ConcentracionGlucosa AND ConcentracionGlucosa ≤ 16mM))
        then
11:      La célula permanece INACTIVA
12:    end if
13:    if velocidadPLC > LimPLC AND velocidadERK < LimiteERK then
14:      La célula MIGRA
15:    end if
16:    if velocidadPLC < LimPLC AND velocidadERK > LimiteERK then
17:      La célula PROLIFERA
18:    end if
19:    if velocidadPLC > LimPLC AND velocidadERK > LimiteERK then
20:      La célula DECIDE: PROLIFERAR O MIGRAR
21:    end if
22:  end if
23: end while

```

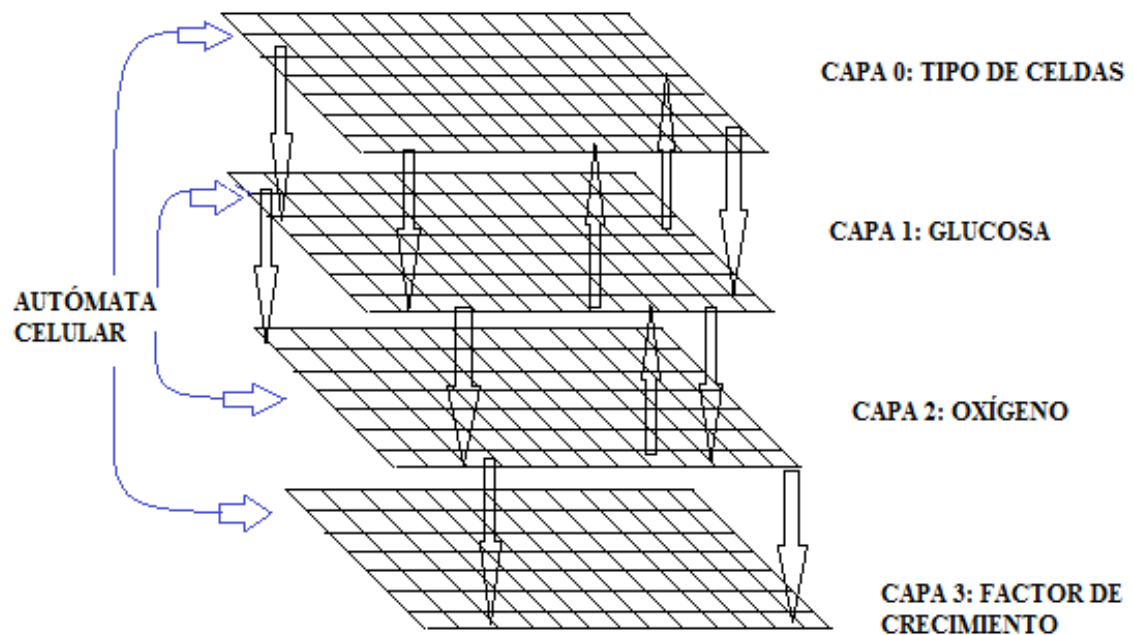


Figura 3.1: Distribución de capas pertenecientes al Autómata Celular

inicializada con una función de distribución distinta, éstas tienen como rasgo en común que son radiales de acuerdo a la fuente, a la distancia de ésta con respecto a la posición de las demás células, dicho en otras palabras, a su vecindario. Por tanto, en la figura 3.1 se puede apreciar a mayor precisión como se encuentran distribuidas las mallas, teniendo en consideración la forma en que éstas son agregadas al AC. Es necesario hacer énfasis en que los métodos para asignar condiciones iniciales dependen únicamente de la capa, ya que las celdas, como elementos individuales, no pueden hacerse asignaciones a sí mismas.

Por otra parte, la difusión se produce independiente para cada componente. Lo que quiere decir que en dicho proceso, los vecinos de la capa glucosa sólo dependen de los vecinos de esta capa, no de los encontrados de manera perpendicular en la capa correspondiente al factor de crecimiento, análogamente sucede con este último.

A continuación, se presenta un algoritmo con el cual se da solución a la inicialización de las capas mencionadas, de acuerdo a sus distribuciones como se ha mencionado en párrafos anteriores:

Algoritmo 2 Algoritmo para obtener el estado inicial del AC

1: ENTRADA: Coordenadas de la fuente de nutrientes.

SALIDA: Matriz con valores de la distribución radial.

2: **for** $i = 0$ hasta *filas* **do**

3: **for** $j = 0$ hasta *columnas* **do**

4: Obtener el módulo entre la fuente de nutrientes situada en el punto con coordenada “X, Y”, con respecto a las células restantes situadas en las coordenadas “i, j”. Es decir, calcular la distancia de la fuente a cada una de las celdas del vecindario.

5: Obtener la distancia ponderada.

6: Para cada capa respectivamente. Sustituir la distancia ponderada en cada distribución radial correspondiente:

$$\text{Capa 0: } Glucose^{ij} = G_a + (G_m - G_a) \cdot \exp \frac{-2d_{ij}^2}{\sigma_g^2}$$

$$\text{Capa 1: } EGF^{ij} = T_m \cdot \exp \frac{-2d_{ij}^2}{\sigma_t^2}$$

$$\text{Capa 2: } Oxygen^{ij} = O_a + (O_m - O_a) \cdot \exp \frac{-2d_{ij}^2}{\sigma_o^2}$$

7: Asignar valores arrojados por la distribución radial a la capa correspondiente del AC.

8: **end for**

9: **end for**

En la figura 3.2 se observa la codificación de dicho algoritmo en el lenguaje de programación java, para la malla de la glucosa. Hay que mencionar además, una breve explicación del mismo para omitir cualquier duda que se pueda presentar a futuros lectores de este proyecto.

En primer lugar, se puede apreciar que la función esta constituida por dos estructuras de repetición anidadas, por el hecho de tener que usar matrices. En la estructura interna de los mismos, se procede a calcular la distancia de la fuente al vecindario y luego la distancia ponderada, el objetivo sin duda alguna es registrar estos valores, ya que son usados en la función radial. Esta última es almacenada en la matriz signada para la glucosa. Al finalizar el método se observa el uso del procedimiento de asignación de la biblioteca Galatea `setActualStateNSCLC()`, quien le proporciona a todo el vecindario los valores que han sido registrados en la matriz.

En segundo lugar, es importante señalar que de forma similar se programó el módulo para las capas restantes, cuya única diferencia es la función de distribución con sus respectivos parámetros.

3.3.2 Tratamiento de la Ecuación Diferencial Parcial

Galatea posee una variedad de módulos o herramientas útiles para la simulación, como se pudo detallar en el capítulo anterior. Sin embargo, a pesar de poseer la librería gode, ésta únicamente da solución a EDOs, no cubriendo el requerimiento de solventar las tres ecuaciones diferenciales parciales de las señales químicas. Por tal razón, fue necesario encontrar un método de linealización del operador de difusión, con la finalidad de transformarlo a un operador algebraico y que su comportamiento fuese similar. Como se ha dicho, las tres señales químicas continúan la difusión sobre la red utilizando la EDP 3.1 siguiente:

$$\frac{\partial M^{ij}}{\partial t} = D_M \cdot \nabla^2 \cdot M^{ij}, t = 1, 2, 3 \quad (3.1)$$

Donde M representa una de las 3 señales externas, y t representa un paso de tiempo.

Avanzando en este razonamiento, luego de comprender algunos de los métodos discretos para aproximar EDP descritos en el capítulo anterior, dicho en otras

```

public double[][] DistribucionRadial_inicializacionMallaGlu() {

    double Distancias[][] = new double[getNRows()][getNColumns()];
    double DistPond[][] = new double[getNRows()][getNColumns()];
    for (int i = 0; i < getNRows(); i++) {
        for (int j = 0; j < getNColumns(); j++) {
            double X = i - source_x;
            double Y = j - source_y;

            Distancias[i][j] = Math.sqrt(Math.pow(X, 2.0) + Math.pow(Y, 2.0));
            DistPond[i][j] = Math.pow(Distancias[i][j], 2.0);
            MatrizGlucosa[i][j] = (Ga + (Gm - Ga) * Math.exp(-2 * DistPond[i][j] / (Sigma * Sigma)));
            // Asignación a la capa.
            CancerPulmonCell celdaGlucosa;
            celdaGlucosa = (CancerPulmonCell) getCell(i, j);
            celdaGlucosa.setActualStateNSCLC(MatrizGlucosa[i][j]);
        }
    }
    return MatrizGlucosa;
}

```

Figura 3.2: Inicialización de la capa Glucosa

palabras, la conversión de un dominio continuo a uno discreto a través del proceso de discretización que conforma al método de diferencias finitas, la EDP anterior se puede representar de la siguiente manera, para cada una de las especies químicas:

$$U_{i,j}^{(t+1)} = U_{i,j}^{(t)} + r \times (U_{i+1,j}^{(t)} + U_{i,j+1}^{(t)} + U_{i-1,j}^{(t)} + U_{i,j-1}^{(t)} - 4 \times (U_{i,j}^{(t)})) \quad (3.2)$$

En Galatea se podría tener una generalización de la ecuación anterior de la siguiente manera:

$$U_{i,j}^{(t+1)} = (1 - 2 \times dim \times r) \times U_{i,j}^{(t)} + r \times \sum U_k^{(t)} \quad (3.3)$$

En la figura 3.3, se observa la representación de la forma generalizada de dicha ecuación en Galatea. Se detalla la variable *dim*, quien representa la dimensión de la capa, *r* es la variable que denota el producto del coeficiente de difusión por el paso del tiempo, entre el paso de la distancia al cuadrado.

```

double nuevoEstado = this.getActualStateNSCLC();
for (int i = 0; i < neighborhoodSize(); i++){
nuevoEstado=(int) ((1-2*DimensionCapa*r2)*this.getActualStateNSCLC()+r2*this.getNeighborState(i));
this.setNextStateNSCLC(nuevoEstado);
}

```

Figura 3.3: EDP representada en Galatea

3.3.3 Solución a las condiciones de borde

Para los bordes de un AC, se debe tener en cuenta el tipo, estacionario y los no estacionarios. Con esto se quiere decir que, en los del primer tipo existe un gran margen de diferencia entre el tiempo de difusión del nutriente y el tiempo de los procesos del ciclo celular, ya sea la proliferación o la migración. En los del segundo tipo, los tiempos son comparables o similares, lo que se refleja en encontrar soluciones mucho más complejas.

Como se puede apreciar en la figura 3.4, el tiempo requerido para la difusión del oxígeno para una celda que ocupe 10 micras es de 23.8 ms aproximadamente. En cuanto a la tasa de nacimiento de una célula, corresponde a períodos de 2 a 3 semanas aproximadamente, es decir, se tiene fracciones de segundos en la difusión contra tiempos de dos semanas en multiplicación celular, siendo tiempos totalmente distintos para cada proceso. En consecuencia, los nutrientes van a estar en estado estacionario cuando la célula se reproduzca, lo que se traduce en encontrar una solución para un problema estacionario, donde las condiciones de los bordes van a llevar siempre a la misma solución, así se integre varias veces, sin la limitación de tener que avanzar el tiempo.

Los bordes son obtenidos luego de tener inicializado el AC con las distribuciones radiales, ya que toman los valores de éstas. De acuerdo con ello, se asume que todas las celdas que conforman el contorno no poseen vecinos, incluyendo también aquellas localizadas en las esquinas. Habría que decir también que a este tipo de bordes, se le es conocido como *bordes absorbentes o membranas permeables*, puesto que los nutrientes no son retenidos en ningún momento.

Dicho lo anterior, para proporcionar una solución al problema de bordes del AC, fue

Table 2. Time required for diffusion of O₂ over a range of distances.

Distance of Diffusion	Approximate Time Required
10 nm	23.8 ns
50 nm	595 ns
100 nm	2.38 μs
1 μm	238 μs
10 μm	23.8 ms
100 μm	2.38 s
1 mm	3.97 min
1 cm	6.61 hours
10 cm	27.56 days

Figura 3.4: Tiempo de difusión del Oxígeno sobre un rango de distancias

necesario el diseño de tres algoritmos, de manera que pudiese facilitar el entendimiento de dicha solución planteada. A continuación, se explica cada uno de los mismos:

- **Algoritmo 3:** describe de manera general la secuencia de pasos necesarios para encontrar el perfil del compuesto químico a estudiar. Partiendo de la red inicializada con los valores arrojados por la función de distribución para cada capa respectivamente, se procede a seleccionar el tamaño y la ubicación en la malla del AC de la submatriz correspondiente a las celdas cancerígenas, cuya finalidad es extraer ésta y calcular el perfil quimiotáctico de las celdas internas del compuesto.
- **Algoritmo 4:** explica a mayor profundidad el proceso de obtener el perfil molecular. Esta conformado de estructuras de repetición anidadas, ya que son necesarias para poder recorrer cada celda perteneciente a la malla, ya que ésta es tratada como una matriz. Además, la utilización de estructuras de decisión para poder delimitar el área donde se extrae la submatriz de celdas malignas y donde se reemplaza por la del perfil calculado. Por otro lado, hace uso de dos variables auxiliares del tipo entero llamadas contadores, la submatriz de celdas malignas y otra matriz auxiliar del mismo tamaño, ambas del tipo doble.

El objetivo es extraer los bordes únicamente y asignarlos en la matriz auxiliar, por ello, los contadores trabajan simultáneamente haciendo los perfiles internos de la submatriz de células malas igual a cero, por tanto, se extrae primeramente el borde superior y el inferior, se realiza la asignación, para luego realizar el mismo procedimiento sobre el borde izquierdo y derecho. Ya teniendo la submatriz con los contornos constantes, se procede a insertar en el mismo lugar donde fue extraída, se llama a la función que calcula el perfil químico, es decir, la descrita en el Algoritmo 5 y por último, a través de las funciones que posee Galatea en la biblioteca de AC, se realiza la asignación a las celdas de la malla del nuevo estado.

- **Algoritmo 5:** explica los pasos necesarios para construir la función que devuelva el calculo aproximado de los puntos internos de una malla, dándole como entrada ésta con contornos constantes. Se encuentra conformado por estructuras de repetición anidadas, así como también de estructuras de decisión. Así mismo, se puede observar el ingreso de la ecuación diferencial parcial discretizada, fragmentada en tres partes, es decir, el nodo actual y el desplazamiento en los dos ejes cartesianos, la suma de éstos corresponde al nuevo estado. En el cálculo de los puntos internos, se usa el valor absoluto de la resta entre el nuevo estado y el estado actual, denotado por la matriz que se le da como entrada a la función.

Posterior a tener el problema de condiciones de borde resuelto, se procedió a aplicar la parte dinámica del modelo, es decir, el algoritmo de decisión de fenotipos partiendo de las concentraciones iniciales que suministra el artículo.

3.4 Construcción de las capas del AC

Un AC construido en el software Galatea puede estar constituido por lo menos con una capa. No existiendo límites al momento de agregarselas a éste. Sin embargo, es relevante acotar, entre mayor es el número de capas, implica un mayor grado de complejidad para lograr el objetivo final; la ejecución de una simulación que se asemeje en gran parte al modelo planteado inicialmente. En el presente caso, fue diseñado con

Algoritmo 3 Algoritmo básico para calcular el perfil de concentración de la señal quimiotáctica

1: ENTRADA: Malla inicializada para cada una de las capas.

$$\text{Capa 0: } Glucose^{ij} = G_a + (G_m - G_a) \cdot \exp \frac{-2d_{ij}^2}{\sigma_g^2}$$

$$\text{Capa 1: } EGF^{ij} = T_m \cdot \exp \frac{-2d_{ij}^2}{\sigma_t^2}$$

$$\text{Capa 2: } Oxygen^{ij} = O_a + (O_m - O_a) \cdot \exp \frac{-2d_{ij}^2}{\sigma_o^2}$$

SALIDA: Malla con la matriz del cáncer en la mitad.

2: Delimitar el tamaño de la matriz del cáncer que se desea en la malla.

3: Extraer matriz del cáncer

4: Mantener unicamente el contorno constante

5: Hallar valores internos. Perfil de concentración.

6: Sustituir dicha matriz en el centro del lattice.

7: Continuar la simulación con la EDP discretizada, según las reglas de cada capa.

cuatro capas, en virtud de que tres son para las señales quimiotácticas de las que se ha hablado desde un comienzo y la cuarta capa, necesaria para realizar interacción con éstas tres. De tal modo que, el estado de la celda no depende solamente de las celdas de su capa y de sus vecinos, depende también de lo que esta pasando en la celda equivalente de la otra capa. Además, cada capa posee su dinámica particular y métodos de consulta, para poder reportar el estado del perfil químico de una determinada celda que se desea conocer. En la figura 3.1 apreciada anteriormente se pudo observar como están organizadas las capas dentro del AC, según el orden como fueron agregadas en el programa principal. Por otra parte, cada vez que se solicita a la capa una célula en particular, se debe realizar un casting a ésta, convirtiendo la celda al tipo que interesa. Por tanto, las celdas de cada capa tienen sus propias funciones para realizar asignaciones y consultas.

3.4.1 ¿Cómo calcular el tamaño del sistema del AC?

El principio de difusión está basado en el movimiento molecular o movimiento Browniano. De tal manera que, dependiendo del tipo de dimensiones donde se este llevando a cabo la investigación, se va a seleccionar la expresión correspondiente a ésta.

Algoritmo 4 Algoritmo Completo para calcular el perfil del compuesto químico

```

1: Procedimiento GenerandoPerfilCancer()
   Doble Submatriz 5×5, CeldasCancerosas 5×5 Entero contador1, contador2
2: for  $i = 0$  hasta  $filas$  do
3:   if  $!(i \geq 0 \text{ AND } i \leq 21) \text{ AND } !(i \geq 27 \text{ AND } i \leq 49)$  then
4:     Inicializar contador2 en cero y Aumentar Contador1
5:     for  $j = 0$  hasta  $columns$  do
6:       if  $!(j \geq 0 \text{ AND } j \leq 21) \text{ AND } !(j \geq 27 \text{ AND } j \leq 49)$  then
7:         Obtener el estado actual de las celdas  $i,j$  del tablero y asignarlos a la
           submatriz. Aumentar el contador2
8:         for  $i = 0$  hasta  $filasMC - 1$  do
9:           Extraer la primera y última fila de subMatriz y asignarla a
           CeldasCancerosas
10:          for  $j = 0$  hasta  $ColumnsMC - 1$  do
11:            Extraer la primera y última columna de subMatriz y asignarla a
            CeldasCancerosas
12:            for  $i = 0$  hasta  $filas$  do
13:              for  $j = 0$  hasta  $Columns$  do
14:                if  $(i \geq 0 \text{ AND } i \leq 21) \text{ OR } (i \geq 27 \text{ AND } i \leq 49) \text{ OR } (j \geq 0$ 
                    $\text{ AND } j \leq 21) \text{ OR } (j \geq 27 \text{ AND } j \leq 49)$  then
15:                  Volver a insertar la matriz calculada en la misma posición
16:                  LLamar al Método Calcular Perfil del compuesto
17:                  Asignar el nuevo estado actual a las celdas  $i,j$  del tablero
                   galatea
18:                end if
19:              end for
20:            end for
21:          end for
22:        end for
23:      end if
24:    end for
25:  end if
26: end for

```

Algoritmo 5 Algoritmo para la función calcular perfil químico

1: ENTRADA: Matriz con bordes constantes y valores internos cero

SALIDA: Matriz con el perfil interno calculado doble aux, diferencia

2: **while** diferencia > aux **do**

3: **for** $i = 0$ hasta *filas* **do**

4: **for** $j = 0$ hasta *columns* **do**

5: Estado Actual = Matriz bordes constantes, Luego:

$$\text{Desplazo en } x = r \times (U_{i-1,j}^{(t)} + U_{i+1,j}^{(t)})$$

$$\text{Nodo Actual} = (1 - (4 \times r)) \times (U_{i,j}^{(t)})$$

$$\text{Desplazo en } y = r \times (U_{i,j+1}^{(t)} + U_{i,j-1}^{(t)})$$

6: Nuevo Estado = Desplazo en x + Nodo Actual + Desplazo en y Matriz bordes constantes = Nuevo Estado

7: **if** Obtener valor absoluto(Nuevo Estado - Estado Actual) > diferencia **then**

8: diferencia = Obtener valor absoluto de Nuevo Estado - Estado Actual

9: **end if**

10: **end for**

11: **end for**

12: **end while**

13: Devolver Matriz con puntos internos calculados

- En 2D: $\langle r \rangle^2 = \langle x \rangle^2 + \langle y \rangle^2 = 4 \cdot D \cdot t$

En el presente caso, el estudio ha sido desarrollado en un entorno de dos dimensiones. Posterior a tener seleccionada la dimensión, se puede obtener el tiempo para recorrer todo el sistema de la siguiente manera:

- Despejar el tiempo de la ecuación. Obteniéndose la siguiente expresión:

$$t = \frac{\langle r \rangle^2}{4}$$

- $TiempoRecorrido = t \cdot filas \cdot columnas$

- Por último, se puede obtener el tamaño del sistema, a través del producto del número de filas por el número de columnas y el tamaño de la celda:

$$Tam = filas \cdot columnas \cdot celda$$

Es importante recalcar que en estudios donde se involucran varios coeficientes de difusión para las especies químicas, se debe seleccionar el más lento.

3.4.2 Funciones de consultas hechas sobre cada capa

Las celdas están distribuidas sobre capas distintas para cada compuesto químico respectivamente. No obstante, las capas se encuentran separadas unas de otras, con una localización de celdas equivalente. Es decir, que celdas de diferentes capas representan el mismo espacio.

Por otro lado, dado que las capas pertenecen al AC, tienen la posibilidad de preguntar quienes son las otras capas, lo cual es necesario para consultar los niveles de las señales químicas y de esta manera mejorar la comprensión del modelo. A continuación se presentan dos funciones de consulta realizadas sobre la capa Glucosa; similarmente son desarrolladas sobre las otras capas restantes.

Algoritmo 6 Algoritmo para la función Obtener perfil Glucosa

1: ENTRADA: Coordenada i , Coordenada j de la celda a consultar
SALIDA: Perfil de Glucosa para la celda i,j
Doble LevelGlucose inicializado en cero
Declaración del tipo celda Glucosa: CancerPulmonCell c

2: **for** $i = 0$ hasta $filas$ **do**

3: **for** $j = 0$ hasta $columns$ **do**

4: **if** Obtener celda i,j del tablero == Obtener celda Coordenada i , Coordenada j **then**

5: Realizar Casting al tipo celda de la posición requerida.

6: LevelGlucose = Estado actual de la celda c

7: **end if**

8: **end for**

9: **end for**

10: Devolver levelGlucose

Algoritmo 7 Algoritmo para la función aportar perfil Glucosa a otra capa

1: ENTRADA: Coordenada i , Coordenada j de la celda a consultar
SALIDA: Perfil de Glucosa para la celda i,j
Doble LevelGlucose inicializado en cero
Declaración del tipo celda Glucosa: CancerPulmonCell c

2: **for** $i = 0$ hasta $filas$ **do**

3: **for** $j = 0$ hasta $columns$ **do**

4: **if** Obtener celda i,j del tablero == Obtener celda Coordenada i , Coordenada j de la capa Tipo Celdas **then**

5: Realizar Casting al tipo celda de la posición requerida.

6: LevelGlucose = Estado actual de la celda c

7: Asignar Estado Actual de la celda c el valor de LevelGlucose

8: **end if**

9: **end for**

10: **end for**

3.4.3 Creación de las reglas para las celdas

Lo dicho hasta aquí supone que, cada compuesto químico estará distribuido en cada una de las capas del AC, existiendo una capa extra para realizar funciones de consulta con las anteriores. No obstante, cada celda debe saber cual es su regla y dependiendo de ésta, se actualizará, es decir, se realizan los cálculos a futuro haciendo uso de la función `NextState()` de la biblioteca de AC de galatea. Además, la regla de transición en común para las tres capas, corresponde a la EDP discretizada. Sin embargo, es importante recalcar que del artículo base del presente proyecto no existe una regla para las señales químicas del oxígeno y el EGF, a diferencia del azúcar, la cual debe estar sobre ciertos niveles para que la célula pueda vivir o simplemente permanecer inactiva. A continuación, mediante los algoritmos ocho, nueve y diez, se podrá apreciar a mayor detalle cada una de estas reglas.

www.bdigital.ula.ve

Algoritmo 8 Algoritmo para el procedimiento Regla de la capa Glucosa

```

1: Doble nuevoEstado = Obtener Actual Estado
2: for  $i = 0$  hasta  $tamañoVecindario$  do
3:   if Estado actual del index de la capa Glucosa == Estado actual del index de
     la capa TipoCeldas then
4:     if Estado Actual de Glucosa < Mínimo Establecido then
5:       nuevoEstado = 0
6:       Asignar estado siguiente = nuevoEstado
7:       if Estado actual  $\geq$  Mínimo Establecido AND Estado actual  $\leq$  Valor
         Estandar de la Glucosa then
8:         La celda permanece Inactiva
9:         nuevoEstado = 4
           Asignar estado Siguiente = nuevoEstado
10:      end if
11:      if Todo lo contrario then
12:        Continuar simulando con la ecuación discretizada
13:        nuevoEstado =  $(1 - (4 \times r)) \times ObtenerEstadoActual +$ 
            $r \times EstadoDelVecino(i)$ 
14:      end if
15:    end if
16:  end if
17: end for

```

Algoritmo 9 Algoritmo para el procedimiento Regla de la capa EGF

```

1: Doble nuevoEstado = Obtener Actual Estado EGF
2: for  $i = 0$  hasta  $tamañoVecindario$  do
3:   if Estado actual del index de la capa EGF == Estado actual del index de la
     capa TipoCeldas then
4:     Solo continuar simulando con la ecuación discretizada
5:     nuevoEstado =  $(1 - (4 \times r1)) \times ObtenerEstadoActual + r1 \times EstadoDelVecino(i)$ 
6:   end if
7: end for

```

Algoritmo 10 Algoritmo para el procedimiento Regla de la capa Oxígeno

```
1: Doble nuevoEstado = Obtener Actual Estado Oxígeno
2: for  $i = 0$  hasta  $tamañoVecindario$  do
3:   if Estado actual del index de la capa Oxig == Estado actual del index de la
     capa TipoCeldas then
4:     Solo continuar simulando con la ecuación discretizada
5:     nuevoEstado =  $(1 - (4 \times r2)) \times ObtenerEstadoActual + r2 \times EstadoDelVecino(i)$ 
6:   end if
7: end for
```

3.5 Integración del modelo molecular Al AC

En el proceso de enlace entre los dos modelos, se tiene lo siguiente:

1. El modelo molecular representado por un sistema acoplado de 20 EDOs, ocurre para cada una de las celdas que se encuentran en la rejilla, únicamente para las celdas cancerígenas. El oxígeno no interviene en las ecuaciones, lo que domina la dinámica es la falta de glucosa.
2. Para simular el crecimiento del cáncer, se va a partir de la submatriz de tamaño 5×5 situada en el centro del lattice, la cual representa las células semillas.
3. Dado que la fuente no se encuentra en la mitad de la rejilla, el crecimiento del cáncer no va a tener un crecimiento simétrico.

Por otra parte, la clase miembro utilizada en la integración del modelo, posee las siguientes características:

1. Cada variable continua es representada mediante una clase. Por tanto, se pueden tener tantas clases como variables continuas se tenga. De tal modo que cada una tendría su función de evaluación, denominada feval().
2. El método para resolver las ecuaciones diferenciales va a depender de donde se encuentren las células malas para el instante t . Con esto quiero decir que, con

una célula cancerígena se resuelve las EDOs, creando todas las densidades de las moléculas, las cuales indicarán cual celda se va a reproducir o migrar.

3. Es necesario la creación de una capa donde se encuentren declaradas cada una de estas clases, para llamar el integrador de galatea y de tal forma obtener las densidades. En el presente caso, fueron creadas dentro de la capa tipo de Celdas, ya que ésta es la que mantiene interacción con las demás capas. Habría que decir también que la importancia de una clase miembro, radica en el hecho de que los atributos y métodos tienen facilidad de acceso desde la clase principal a ésta y viceversa. Sin embargo, no existe problema en tener una capa solamente para el Tipo de Celdas y otra exclusivamente para el estado de las ecuaciones.
4. Mediante la función `Addeqs` del software Galatea, la cual es usada para agregar ecuaciones, es usada tanto en la clase miembro para darle resolución al conjunto de ecuaciones, como dentro del constructor de la clase celda. La finalidad de ello, es crear cada una de las celdas con las EDOs.

3.5.1 Acoplar las reglas de crecimiento celular

En los modelos de contagio para el comportamiento proliferativo y migratorio de las células. Algunos puntos a considerar son los siguientes:

- A través del modelo molecular se determinan los valores de los parámetros del algoritmo, dado que éste es gestionado por dos moléculas claves: ERK y PLCy. Por tanto, dichos valores van a indicar si la célula prolifera, migra o permanece inactiva.
- En cada unidad de tiempo el AC debe ver si nace o no otra célula, mediante la aplicación de las 4 reglas del algoritmo de decisión fenotipo.
- Las reglas se programan de acuerdo a las concentraciones obtenidas de la integración de las EDOs.

A continuación se explicará a mayor detalle el modelo de proliferación mediante el algoritmo 11, así como también el comportamiento migratorio celular a través del algoritmo 12.

Algoritmo para modelar la proliferación celular

“El proceso conocido como ciclo celular es de gran importancia para la célula ya que tiene como función la formación completa de una nueva célula, evitando en lo posible la creación de células con múltiples errores, lo cual permite al organismo permanecer en un constante equilibrio. El inicio del ciclo celular se da cuando aparece una nueva célula, descendiente de otra que se divide, y termina en el momento que dicha célula, por división subsiguiente, origina dos nuevas células hijas” [55].

Teniendo en cuenta el conjunto de argumentos mencionados en el párrafo anterior, y dado que en una simulación se debe tratar de lograr la mayor similitud posible al fenómeno que se este analizando, se trato de realizar el comportamiento proliferativo mediante el uso de las siguientes herramientas:

- Vecindario de Von Neumann (4 vecinos), el cual se detallo en el capítulo anterior.
- Uso de bucles y estructuras de decisión.
- Uso de probabilidades.

De tal manera que, que en algoritmo 11 se puede detallar el proceso de modelar este comportamiento. La idea básica tomada es la generación de dos nuevas células en cada paso del tiempo, donde la célula madre en algunos casos se muere, es decir, si su probabilidad es mayor a 0.5, ésta vive. En cuanto al estudio de sus vecinos, en el algoritmo se va comparando sucesivamente, cuales de ellos estan vivos, con el objetivo de que los que no lo estén, en dichos lugares la celda tiene lugar para proliferar, dicho en otras palabras, las células son cancerígenas lo cual implica que su comportamiento es descontrolado al momento de proliferar y en la apoptosis, eso implica que se asumió que van a ir generando dos hijas, con la muerte en algunos casos de la célula madre.

Algoritmo 11 Algoritmo para modelar la proliferación celular

```

1: Procedimiento ComportamientoProliferacion()
2: for  $i = 1$  hasta  $filas - 1$  do
3:   for  $j = 1$  hasta  $columnas - 1$  do
4:     Poner fuente en la posición i-fuente, j-fuente del tablero
     Sumar estado actual de los vecinos de Neumann
     Entero Suma =  $U_{i,j-1}^{(t)} + U_{i+1,j}^{(t)} + U_{i,j+1}^{(t)} + U_{i-1,j}^{(t)}$ 
5:     if ( $i \neq$ i-fuente) AND ( $j \neq$ j-fuente) then
6:       if (Suma==2) then
7:         if  $U_{i,j+1}^{(t)} + U_{i-1,j}^{(t)} == 2$  then
8:           if (Probabilidad >0.5) then
9:              $U_{i,j}^{(t+1)} = 1$ 
10:          end if
11:           $U_{i,j-1}^{(t+1)} = 1$  ,  $U_{i+1,j}^{(t+1)} = 1$ 
12:        end if
13:        if  $U_{i,j-1}^{(t)} + U_{i+1,j}^{(t)} == 2$  then
14:          if (Probabilidad >0.5) then
15:             $U_{i,j}^{(t+1)} = 1$ 
16:          end if
17:           $U_{i,j+1}^{(t+1)} = 1$  ,  $U_{i-1,j}^{(t+1)} = 1$ 
18:        end if
19:        if  $U_{i,j-1}^{(t)} + U_{i,j+1}^{(t)} == 2$  then
20:          if (Probabilidad >0.5) then
21:             $U_{i,j}^{(t+1)} = 1$ 
22:          end if
23:           $U_{i+1,j}^{(t+1)} = 1$  ,  $U_{i-1,j}^{(t+1)} = 1$ 
24:        end if
25:      end if
26:      if (Suma ==4) then
27:         $U_{i,j}^{(t+1)} = 0$ 
28:      end if
29:    end if
30:  end for
31: end for

```

Algoritmo para modelar la migración celular

En cuanto el modelado del comportamiento migratorio, se usaron las mismas herramientas que el proliferativo. Es relevante recordar que, la parte de vascularización no lo comprende el presente proyecto, lo que significa que la activación de las células cancerígenas solo se dará entre sus vecinos más cercanos. De tal modo que en el algoritmo se va comparando si existe entre el vecindario algún lugar disponible, para que la célula pueda moverse. En caso de que los cuatro vecinos estén vivos, la célula permanece inactiva hasta el paso siguiente del tiempo. En el algoritmo 12, se puede apreciar a mayor brevedad lo enunciado.

www.bdigital.ula.ve

C.C. Reconocimiento

Algoritmo 12 Algoritmo para modelar la migración celular

```

1: Procedimiento ComportamientoMigracion()
2: for  $i = 1$  hasta  $filas - 1$  do
3:   for  $j = 1$  hasta  $columnas - 1$  do
4:     Poner fuente en la posición  $i$ -fuente,  $j$ -fuente del tablero
     Entero Suma =  $U_{i,j-1}^{(t)} + U_{i+1,j}^{(t)} + U_{i,j+1}^{(t)} + U_{i-1,j}^{(t)}$ 
5:     if ( $i \neq i$ -fuente) AND ( $j \neq j$ -fuente) then
6:       if (Suma==3) then
7:         if  $U_{i+1,j}^{(t)} + U_{i,j+1}^{(t)} + U_{i-1,j}^{(t)} == 3$  then
8:            $U_{i,j-1}^{(t+1)} = 2$ 
9:         end if
10:        if  $U_{i,j-1}^{(t)} + U_{i,j+1}^{(t)} + U_{i-1,j}^{(t)} == 3$  then
11:           $U_{i+1,j}^{(t+1)} = 2$ 
12:        end if
13:      end if
14:      if (Suma==2) then
15:        if  $U_{i,j-1}^{(t)} + U_{i+1,j}^{(t)} == 2$  then
16:          if (Probabilidad >0.5) then
17:             $U_{i,j+1}^{(t+1)} = 2$ 
18:          end if
19:           $U_{i-1,j}^{(t+1)} = 2$ 
20:        end if
21:        if  $U_{i,j-1}^{(t)} + U_{i-1,j}^{(t)} == 2$  then
22:          if (Probabilidad >0.5) then
23:             $U_{i+1,j}^{(t+1)} = 2$ 
24:          end if
25:           $U_{i,j+1}^{(t+1)} = 2$ 
26:        end if
27:      end if
28:      if (Suma ==4) then
29:         $U_{i,j}^{(t+1)} = 4$ 
30:      end if
31:    end if
32:  end for
33: end for

```

3.6 Organización de códigos fuentes

El presente proyecto fue estructurado en cuanto a los códigos fuentes elaborados, de la siguiente manera:

- Programa principal: `CancerPulmonAutomata.java`. Se realiza la declaración de las capas y agregación de éstas al AC, la declaración de las constantes usadas en la resolución del modelo molecular, así como también la declaración de archivos para almacenar los datos de la simulación.
- Perfil de glucosa: `GlucoseLayer.java`, `CancerPulmonCell.java`. El primero corresponde a la capa de dicha señal química, constituido por las funciones de consulta, función de inicialización del AC y las funciones para solución de condiciones de bordes. En cuanto al segundo archivo, se encuentra conformado por la regla de la celda y las funciones respectivas para esa celda en particular. De manera similar para los otros dos compuestos químicos.
- Perfil del factor de crecimiento: `Egf-Layer.java`, `Egf-LayerCell.java`.
- Perfil de Oxígeno: `Oxigeno-Layer.java`, `Oxigeno-LayerCell.java`
- Capa de interacción: `TipoCeldas.java`, `TipoCeldasCell.java`. El primero de ellos posee funciones de consulta con respecto a otras capas. El segundo, es el más importante, ya que posee la clase miembro para dar solución a las EDOs, es decir, está conformado por las funciones para la celda de esta capa, las clases para cada una de las variables continuas dentro de la clase miembro.
- `TableroAC.java`. Constituido por algunas funciones que trabajan exclusivamente con la parte de la interfaz gráfica. Por tanto, posee la parte de inicialización del tablero con los siguientes colores:
 - Negro: Corresponde a las células semilla que son cancerígenas en estado de proliferación.
 - Rojo: Corresponde a las células semilla que son cancerígenas en estado de migración.

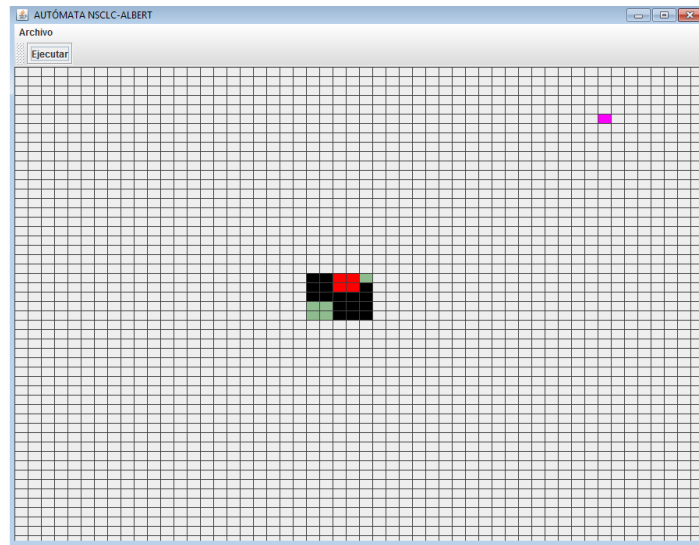


Figura 3.5: Interfaz Gráfica del AC

- Verde claro: Aquellas células que por el momento se encuentran en estado inactivo.
- Rosado: Para diferenciar la fuente de nutrientes.
- Gris: Células que se encuentran sanas.

Además de ello, se programo las funciones de comportamiento de la proliferación y migración en el mismo.

- `VentanaAutomata.java`. Es la interfaz gráfica diseñada para el AC, construida en un `Jpanel`, a su vez se hace uso de la clase `Timer`, para manipular el paso del tiempo del crecimiento celular.

3.7 Aportes del presente proyecto

1. Resolución de EDPs a través de aproximaciones numéricas. Esto es, convertir un problema continuo en un problema discreto, representado a través de una malla de puntos mediante el proceso de discretización. Por lo tanto, de una manera u otra se ésta combinando temas continuos con discretos.
2. Ingresar EDOs acopladas en una capa perteneciente al AC, para posteriormente resolverlas con el integrador de ecuaciones que posee la plataforma.
3. Comprensión de procesos biológicos relacionados con el cáncer de pulmón: proceso de transducción de señales, entre uno de los más importantes.
4. A través de la construcción de la simulación, darnos cuenta de los aspectos que se le podrían mejorar a la plataforma. Por ejemplo, parametrizar las celdas.
5. Tomar un modelo multiescala con la dinámica desarrollada por Wang y construir un primer prototipo de AC sobre el software GALATEA, que represente estos tipos de procesos que han sido mencionados a lo largo de todos los capítulos.
6. Ser un modelo base para la realización de futuros proyectos en el ámbito de la simulación biológica.

Capítulo 4

RESULTADOS

Para realizar un modelo de cualquier realidad que se desee representar, es necesario tener claro la dimensión del fenómeno físico que se está queriendo modelar. A su vez, tener presente una gran variedad de detalles, por ejemplo, trabajar en mismas unidades de tiempo o espacio, por mencionar algunos. Se simula para explicar, comprender o simplemente para mejorar el sistema [49]. Por tanto, la realización del presente proyecto tenía como uno de sus propósitos, evaluar el desempeño de la plataforma de simulación y de esta manera, continuar con la línea de investigación en procesos que involucren difusión en el Centro de Simulación y Modelos.

A continuación se presenta una pequeña comparación de las reglas de mayor incidencia en el comportamiento fenotípico celular, éstas actúan de manera recursiva y haciendo uso de probabilidades, ya que como se mencionó al comienzo del presente trabajo, el cáncer es un proceso complejo que involucra fenómenos emergentes o auto-organización. De tal manera que en la figura 4.1 se puede contrastar con las figuras 4.2 y 4.3, la cual corresponde a la simulación del AC diseñado con la plataforma de simulación Galatea, mayormente la librería de Ac. Por otra parte, en la figura 4.4 se puede apreciar el comportamiento de la molécula PLCy, donde en el artículo base del proyecto, decae mucho más rápido en virtud de que toman en consideración el comportamiento de una celda específica, la número 48. A diferencia del modelo realizado en Galatea, donde se gráfica de manera general no especificando una en particular.

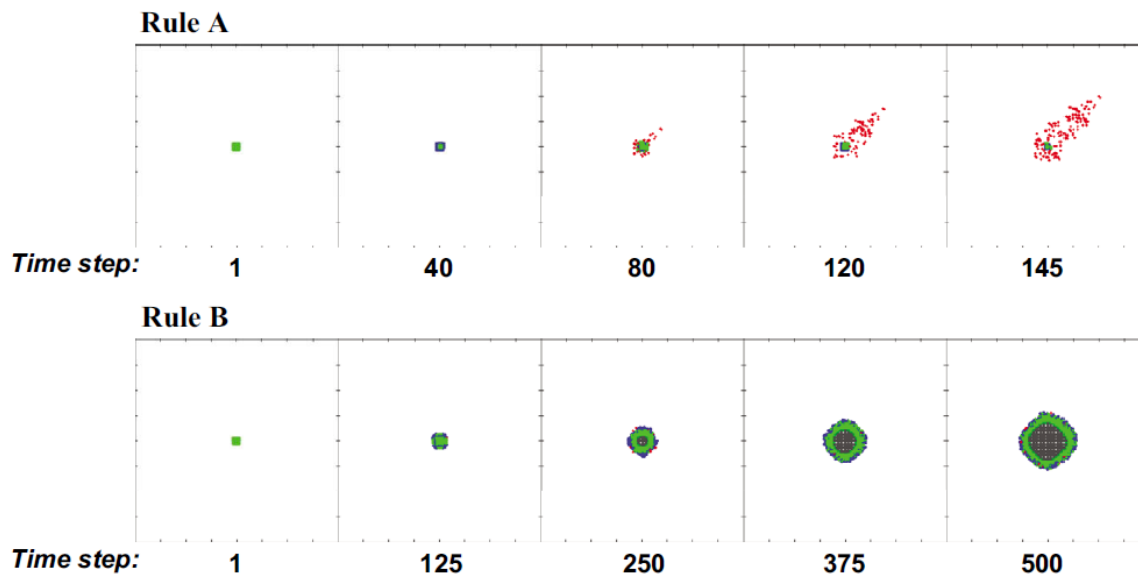


Figura 4.1: Influencias de la proliferación y migración en el artículo base

www.bdigital.ula.ve

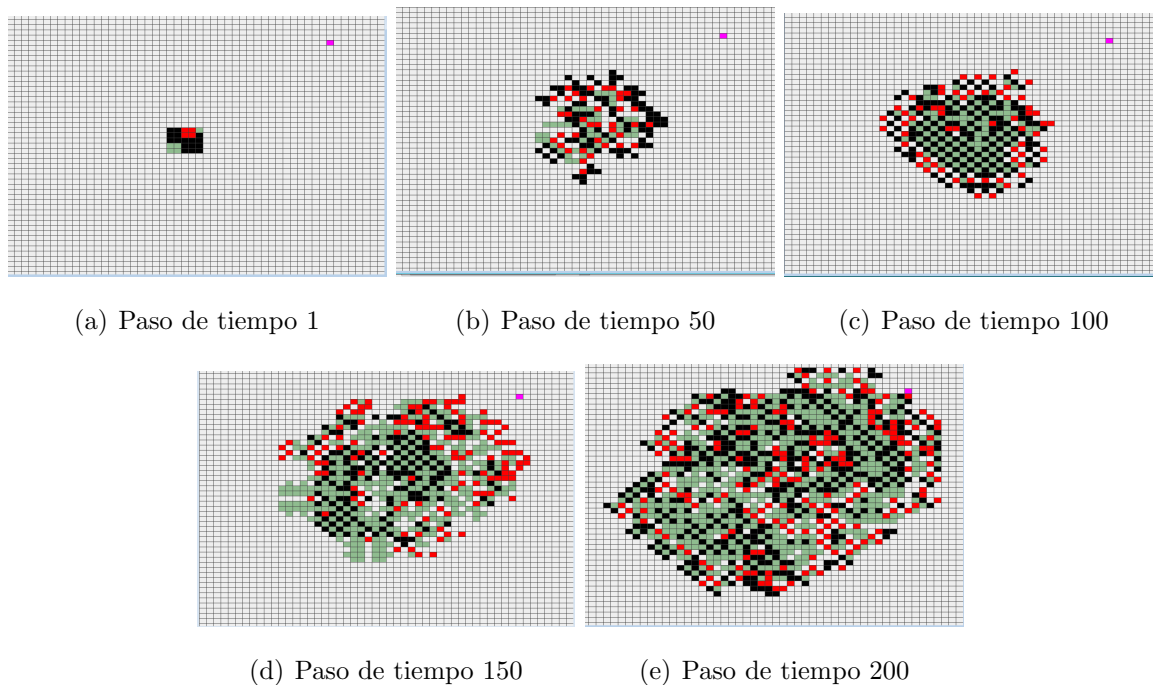


Figura 4.2: REGLA A: Migración

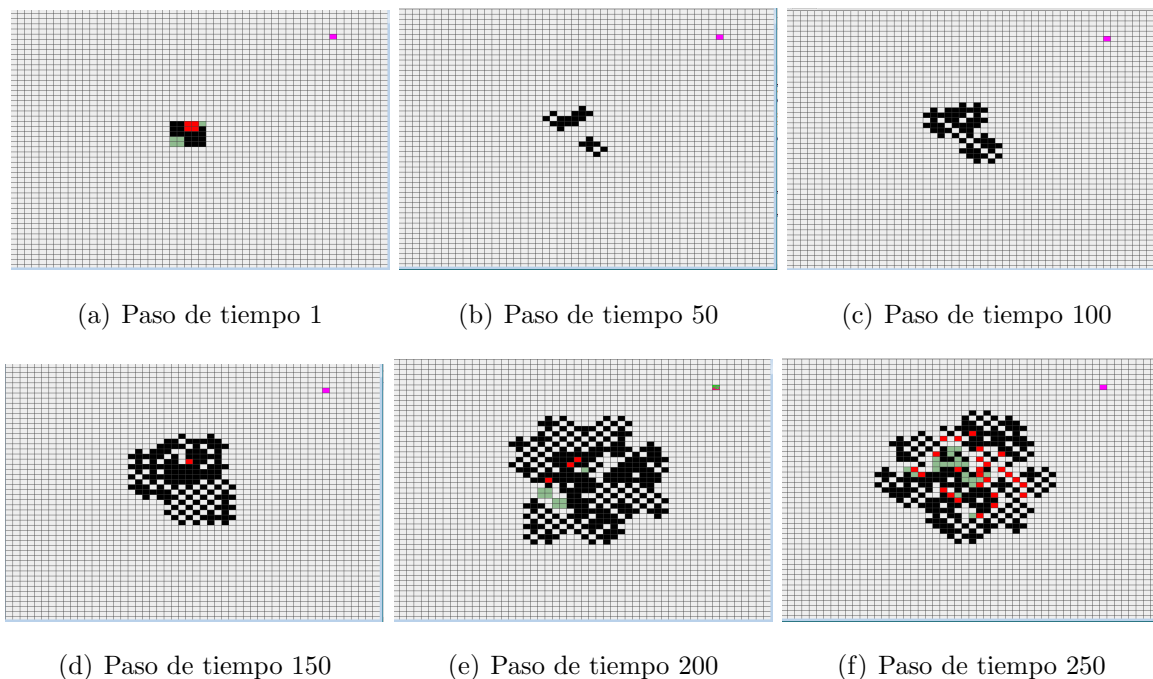
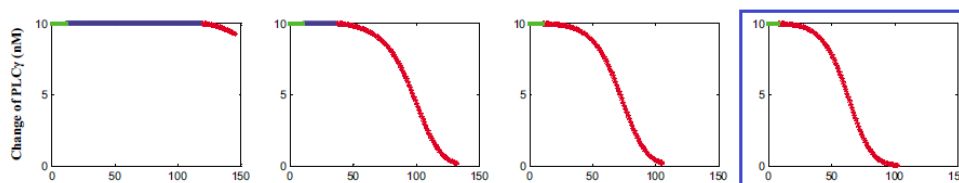


Figura 4.3: REGLA B: Proliferación

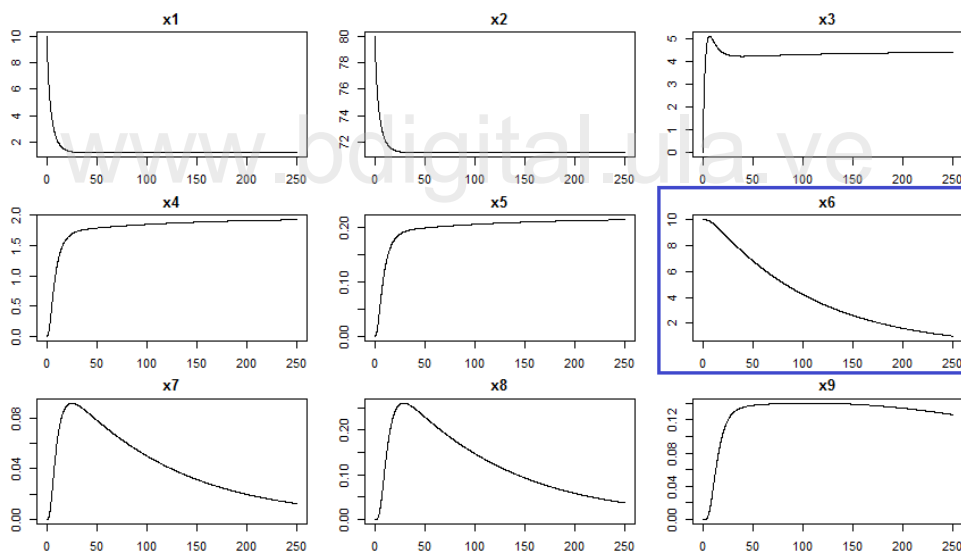
4.1 CONCLUSIONES

Con la elaboración del presente proyecto, se logró los objetivos planteados en el capítulo I. Entre los procesos de señalización molecular que ocurren en la célula, uno de los más importantes es el llamado transducción de señales, donde se produce el intercambio de información desde la parte externa de la célula a su interior; tampoco debemos menospreciar la proliferación y migración ya que son los dos rasgos fenotípicos del cual se basa principalmente el comportamiento del AC.

Por otra parte, también fue necesario la comprensión de los conceptos referentes a Galatea, fundamentalmente en lo que se refiere a las funciones contenidas en el módulo de autómatas celulares y gode, la cual da solución al sistema de 20 EDOs acopladas, dicho en otras palabras, el modelo de señalización molecular. De tal manera que, uno de los objetivos al comienzo del presente trabajo, era probar la capacidad de la plataforma en un proyecto de dicha magnitud. Por tanto, se hizo una comparación de los resultados proporcionados por la plataforma en el modelo molecular, y resultados obtenidos con el software RStudio. En la figura 4.5 se puede apreciar a mayor detalle



(a) Velocidad de cambio de PLCy



(b) Velocidad de cambio de las moléculas

Figura 4.4: Comparación comportamiento de PLCy

la similitud de los mismos.

En cuanto a la capacidad de la plataforma Galatea en la resolución de las EDOs.

```
> head(out)
      time      x1      x2      x3      x4      x5      x6
[1,] 0.00 10.000000 80.00000 0.0000000 0.000000e+00 0.000000e+00 10.000000
[2,] 0.08  9.810508 79.81051 0.1894718 1.004406e-05 2.127405e-07 10.000000
[3,] 0.16  9.625938 79.62594 0.3739109 7.304667e-05 2.310181e-06 10.000000
[4,] 0.24  9.446144 79.44614 0.5533595 2.383733e-04 9.695741e-06 10.000000
[5,] 0.32  9.270985 79.27098 0.7278635 5.481969e-04 2.660101e-05  9.999999
[6,] 0.40  9.100325 79.10033 0.8974738 1.039836e-03 5.745302e-05  9.999997
      x7      x8      x9      x10 x11      x12 x13
[1,] 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 10 0.000000e+00 100
[2,] 1.977695e-09 2.786731e-11 2.750628e-14 2.590836e-16 10 5.539354e-16 100
[3,] 5.145193e-08 1.586605e-09 1.175206e-11 2.370526e-13 10 4.502071e-13 100
[4,] 3.084214e-07 1.363632e-08 1.412771e-10 4.258631e-12 10 7.597680e-12 100
[5,] 1.093072e-06 6.449881e-08 8.707112e-10 3.676062e-11 10 6.136929e-11 100
[6,] 2.876013e-06 2.126507e-07 3.525040e-09 1.930785e-10 10 3.010951e-10 100
      x14 x15      x16      x17 x18      x19      x20
[1,] 0.000000e+00 120 0.000000e+00 0.000000e+00 100 0.000000e+00 0.000000e+00
[2,] 1.603870e-19 120 1.835777e-25 1.164601e-51 100 1.372059e-62 0.000000e+00
[3,] 1.818229e-14 120 1.930848e-18 1.541089e-38 100 3.881658e-49 3.710077e-103
[4,] 4.828811e-13 120 1.013320e-16 4.114275e-33 100 1.001304e-39  5.461185e-82
[5,] 5.511805e-12 120 1.783514e-15 1.370186e-30 100 9.186964e-37  1.207902e-73
[6,] 3.572386e-11 120 1.598673e-14 1.040377e-28 100 1.221217e-34  4.893784e-69
```

(a) Modelo Molecular resuelto en RStudio

```
 t      x1      x2      x3      x4      x5      x6      x7      x8      x9      x10 x11      x12 x13
0.00 9.810509 79.81051 0.1894714 9.464571e-06 1.519038e-07 10.000000 2.359296e-09 0.000000e+00 0.000000e+00 0.000000e+00 10 0.000000e+00 100
0.08 9.625939 79.62594 0.3739107 7.303051e-05 2.201179e-06 10.000000 4.592941e-08 1.134112e-09 6.588130e-12 1.013492e-13 10 1.970884e-13 100
0.16 9.446144 79.44614 0.5533591 2.385383e-04 9.694618e-06 10.000000 2.911896e-07 1.204705e-08 1.150875e-10 3.313629e-12 10 5.961322e-12 100
0.24 9.270984 79.27098 0.7278630 5.483644e-04 2.676709e-05 9.999999 1.063871e-06 6.022259e-08 7.766856e-10 3.193357e-11 10 5.347404e-11 100
0.32 9.100325 79.10033 0.8974732 1.040215e-03 5.755404e-05 9.999997 2.850934e-06 2.045517e-07 3.280463e-09 1.738880e-10 10 2.729062e-10 100
0.40 8.934035 78.93403 1.0622448 1.747590e-03 1.059499e-04 9.999993 6.279393e-06 5.464202e-07 1.041359e-08 6.749607e-10 10 9.973251e-10 100
      x14 x15      x16      x17 x18      x19      x20
0.000000e+00 120 0.000000e+00 0.000000e+00 100 0.000000e+00 0.000000e+00
6.304960e-15 120 0.000000e+00 0.000000e+00 100 0.000000e+00 0.000000e+00
3.599352e-13 120 7.788312e-17 2.050652e-33 100 4.126976e-40 6.239503e-80
4.736188e-12 120 1.609049e-15 8.672532e-31 100 5.101098e-37 2.113570e-73
3.161964e-11 120 1.428758e-14 6.572827e-29 100 6.552733e-35 4.121904e-69
1.425153e-10 120 7.989700e-14 2.004879e-27 100 2.762328e-33 7.505406e-66
```

(b) Modelo Molecular resuelto en Galatea

Figura 4.5: Comparación de soluciones

En el proceso de integración del modelo mencionado y el microscópico, el uso de clases anidadas fue de vital importancia, ya que cada celda en su interior estaría resolviendo el conjunto de EDOs. De tal modo, que las 20 variables continuas son representadas por clases dentro de la clase miembro, resueltas a un paso de tiempo inferior al paso de tiempo de crecimiento del AC.

Con respecto a la solución para el proceso difusivo, se tiene un estado estacionario para la solución de este, dicho en otras palabras, la célula crece muy lentamente en comparación con la difusión de nutrientes en el lattice, habiendo un paso de tiempo para la reproducción mucho más largo que el paso de difusión. Además, dado que

la plataforma trabaja con diferentes mallas para cada compuesto quimiotáctico, se elaboro funciones de consulta, de tal manera que si se desea conocer la cantidad de Glucosa, oxígeno o factor de crecimiento de una celda en particular, basta con pasar las coordenadas de ésta a la función, como puede ser observado para cada una de las capas en las figuras: 4.6, 4.7 y 4.8. Donde posterior a tener resuelto el perfil de concentración del nutriente sólo para la matriz del cáncer, se extrae el valor de las celdas de interés.

```

17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0
17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0
17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0
17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0
17.0 17.0 17.0 16.998235243395303 16.99860622074294 16.999449607946275 17.0 17.0
17.0 17.0 17.0 16.99970521301405 16.999767182252967 16.999908062260825 17.0 17.0
17.0 17.0 17.0 16.999975379321448 16.99998055500922 16.999992321339473 17.0 17.0
17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0
17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0
17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0
17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0
17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0
17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0 17.0

```

(a) Parte de los resultados de difusión para la capa Glucosa

Concentración de las celdas de interes:

```

Concentración de Glucosa de la celda P(23,23) = 16.998235243395303
Concentración de Glucosa de la celda P(23,25) = 16.999449607946275
Concentración de Glucosa de la celda P(25,23) = 16.999975379321448
Concentración de Glucosa de la celda P(25,25) = 16.999992321339473
Agregando nodo CAPA 1: GLUCOSA[1]

```

(b) Prueba de la función de consulta

Figura 4.6: Difusión Capa Glucosa

En definitiva, se pudo lograr acoplar dos modelos que ocurren a escalas espacio-tiempo distintos en una misma simulación, evaluando el modelo y el desempeño de la plataforma de simulación Galatea, utilizando términos biológicos y AC, lo cual

```

0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017
0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017
0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017
0.0017 0.001698424093500711 0.00170000959637796407 0.0017000077451972296 0.0017 0.001
0.0017 0.0017000001604600148 0.001699999994431647 0.0016999999986277933 0.0017 0.001
0.0017 0.0016999999999824922 0.0017000000000003072 0.001699999999999856 0.0017 0.00
0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017
0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017
0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017
0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017 0.0017

```

(a) Parte de los resultados de difusión para la capa Oxígeno

Concentración de las celdas de interes:

```

Concentración de Oxígeno de la celda P(23,23) = 0.001698424093500711
Concentración de Oxígeno de la celda P(23,25) = 0.0017000077451972296
Concentración de Oxígeno de la celda P(25,23) = 0.0016999999999824922
Concentración de Oxígeno de la celda P(25,25) = 0.001699999999999856
Agregando nodo CAPA 2: OXÍGENO[2]

```

(b) Prueba de la función de consulta

Figura 4.7: Difusión Capa Oxígeno

```

2.56 8.587843274304303E-4 3.242026380568109E-14 1.3773276569654115E-31 6.584855993964582E-
[
8.587843274304303E-4 2.8809004728130233E-7 1.087578689354639E-17 4.620419552884198E-35 2.2
[
3.242026380568109E-14 1.087578689354639E-17 4.105755879804454E-28 1.7442705463156213E-45 8
[
1.3773276569654115E-31 4.620419552884198E-35 1.7442705463156213E-45 7.410279197819859E-63
r

```

(a) Parte de los resultados de difusión para la capa Oxígeno

```

Concentración de EGF de la celda P(0,0) = 2.56
Concentración de EGF de la celda P(0,3) = 1.3773276569654115E-31
Concentración de EGF de la celda P(3,3) = 7.410279197819859E-63
Concentración de EGF de la celda P(3,0) = 1.3773276569654115E-31
Agregando nodo CAPA 3: FACTOR EPIDERMICO DE CRECIMIENTO[3]

```

(b) Prueba de la función de consulta

Figura 4.8: Difusión Capa Oxígeno

constituye un pequeño aporte para trabajos posteriores en esta área para el Centro de Simulación y Modelos (CeSiMo). Las conclusiones a las que se pueden llegar usando un modelo no describirán con rigurosa exactitud lo que ocurre en el sistema real (especialmente en sistemas con un alto grado de incertidumbre) pero, cuando menos, aportarán un conocimiento significativamente mejor que el que obtendríamos sin aplicar modelo alguno. El valor de un modelo no se limita necesariamente al análisis del sistema real para el cual el modelo fue diseñado. El conocimiento adquirido mediante el estudio de un determinado modelo puede trascender a otros sistemas reales cuya esencia quede capturada por la misma abstracción, incluso aunque el modelo no se desarrolle inicialmente con estos otros sistemas reales en mente [45].

www.bdigital.ula.ve

C.C. Reconocimiento

Bibliografía

- [1] G.R. Anderson, D.L. Stoler, and B.M. Brenner. *Cancer: the evolutionary consequence of a destabilized genome*. BioEssays, 23:1037-1046, 2001.
- [2] Solé, R. V. (2003). *Phase transitions in unstable cancer cell populations*. The European Physical Journal B-Condensed Matter and Complex Systems, 35(1), 117-123.
- [3] López Romero, Mónica (). *Estudio de un modelo matemático para la inhibición de angiogénesis en tumores secundarios*. Tesis, p 102.
- [4] *Información sobre cáncer*. Boletín informativo de la Asociación Española Contra el Cáncer. Disponible en <http://www.aecc.es>.
- [5] Organización Mundial de la Salud. *Cáncer*. Disponible en: <http://www.who.int/mediacentre/factsheets/fs297/es/>, Informe mundial sobre el cáncer (2014). [Consultado el 25 de Marzo de 2016]
- [6] Ortiz, M .A,González, M. A.,Lorenzo, J.L.,Hernández, M., Cabrera, C. (2005). *El cáncer de pulmón: Algunos aspectos sociales y bioéticos*. Gaceta Médica espirituaana 7(3).
- [7] The American Cancer Society Inc., *The History of Cancer* . 25 Mar. 2002. 13-17 June. 2005.
- [8] From Wikipedia, the free encyclopedia,(2016). *Lung Cancer* .Disponible en http://wikipedia.org/wiki/lung_cancer. [consultada 15.03.2016].

- [9] Menchón, S. A. (2007). *Modelado de las diversas etapas del crecimiento del cáncer y de algunas terapias antitumorales*. Universidad Nacional de Córdoba.
- [10] Zapata Peña, J., Ortiz, P., and Cristina, A. (2010). *Uso de modelos matemáticos para la descripción del crecimiento de tumores cancerosos*. NOVA publ. cient, 8(14),140-147.
- [11] OSPINA, M. E. E., POR, D., and PERDOMO, D. J. G. *Modelo de crecimiento de las variantes del virus de papiloma humano tipo 16*. Tesis, p 106.
- [12] Ferlay, J., and al.,(2008). *Estimates of worldwide burden of cancer in 2008: GLOBOCAN*. International journal of cancer. Journal international du cancer, 2010. 127(12): p. 2893-917.
- [13] Wang, Z., Zhang, L., Sagotsky, J., and Deisboeck, T. S. (2007). *Simulating non-small cell lung cancer with a multiscale agent-based model*. Theoretical Biology and Medical Modelling,4,50.
- [14] Hernández, M. E., and Bagilet, L. R. (2015). *Una mirada a los modelos basados en agentes*. Revista de Educación Matemática, 30(1).
- [15] Dávila, J. Uzcátegui, M. Tucci, K. *Simulación Multi-agente con GALATEA*. Universidad de Los Andes. (2004).
- [16] Uzcátegui, M. *Diseño de la plataforma de simulación de sistemas multi-agentes GALATEA, Tesis de Maestría*. Universidad de Los Andes. (2002).
- [17] M. Uzcátegui, J. Dávila, K. Tucci, *GALATEA: Plataforma de Simulación de Sistemas Multi-Agentes*, Memorias de VI Jornadas científico técnicas de la Facultad de Ingeniería, Universidad de Los Andes, Mérida, Venezuela (2007).
- [18] b. P. Zeigler, Theory of modelling and simulation, ser. Interscience. New York: Jhon Wiley and sons, (1976).
- [19] Wang, Z., Birch, C. M., Sagotsky, J., and Deisboeck, T. S. (2009). *Cross-scale, cross-pathway evaluation using an agent-based non-small cell lung cancer model*. Bioinformatics, 25(18), 2389-2396.

- [20] Wang, Z., Birch, C. M., and Deisboeck, T. S. (2008). *Cross-scale sensitivity analysis of a non-small cell lung cancer model: linking molecular signaling properties to cellular behavior*. Biosystems, 92(3), 249-258.
- [21] Pennisi, M., Pappalardo, F., Palladini, A., Nicoletti, G., Nanni, P., Lollini, P. L., and Motta, S. (2010). *Modeling the competition between lung metastases and the immune system using agents*. BMC bioinformatics, 11(7), 1.
- [22] Zhang, L., Athale, C. A., and Deisboeck, T. S. (2007). *Development of a three-dimensional multiscale agent-based tumor model: simulating gene-protein interaction profiles, cell phenotypes and multicellular patterns in brain cancer*. Journal of theoretical biology, 244(1), 96-107.
- [23] Deisboeck, T. S., and Stamatakis, G.S. (2011). *Multiscale cancer modeling*. Mathematical and computational biology series, (Chap. 9, pp. 173-187).
- [24] Wang, Z., Butner, J. D., Kerketta, R., Cristini, V., and Deisboeck, T. S. (2015, February). *Simulating cancer growth with multiscale agent-based modeling*. In Seminars in cancer biology (Vol. 30, pp. 70-78). Academic Press.
- [25] Gómez, Á., Valle, A. A. E., and Aldaco, F. (2004). *Receptor de factor de crecimiento epidérmico y sus implicaciones terapéuticas en oncología*. Revista de Especialidades Médico-Quirúrgicas, 9(2), 12-16.
- [26] Respiratory health association of metropolitan chicao tipos de cancer pulmonar (2007). *Tipos de cáncer de Pulmón*.
- [27] Editorial medica Panamericana. *Comunicación celular*. Disponible en <http://www.curtisbiologia.com/node/124>. [Consultada el 13.04.2016].
- [28] Universidad de Oviedo. *Regulación del crecimiento celular normal*. Disponible en http://www10.uniovi.es/anatopatodon/modulo5/tema04_regeneracion/03receptores.htm. [Consultada el 13.04.2016].
- [29] Vargas, C. (2008). *Receptor del factor de crecimiento epidérmico y cáncer*. (Doctoral dissertation, Facultad de Ciencias Médicas).

- [30] Martínez Iglesias, O. A. (2007). *Efecto de la expresión del receptor de hormonas tiroideas beta (TR β) sobre la transformación celular, tumorigénesis y formación de metástasis.*
- [31] López-Ayllón, B. D. (2013). *Estudio de la resistencia a quimioterapia en células con características de cáncer stem cells (CSC) derivadas de cáncer no microcítico de pulmón.* Universidad Autónoma de Madrid.
- [32] Carreón, O. Z., Mendiola, A. V., Steider, B. W., and Cruz, I. S. (2012). *El receptor para el factor de crecimiento epidérmico (EGFR) y su relación con el cáncer.* 15(1), 15-25.
- [33] Anido Folgueira, J. (2006). *Nuevos mecanismos de tumorigénesis del protooncogén HER2. Implicaciones terapéuticas en cáncer de mama.* Universidad de Barcelona.
- [34] Charriel B., P. *Estudio de la activación de PI3K/AKT/MTOR, RAS/RAF/MEK/ERK, WNT/B-CATENINA Y SONIC-HEDGEHOG en osteosarcoma.* Trabajo de maestría. Máster Universitario en Investigación en Medicina (Valencia/Spain) Clínica. 1-22.
- [35] Pinzón, C. E., Serrano, M. L., and Sanabria, M. C. (2009). *Role of phosphatidylinositol 3-kinase pathway (PI3K/Akt) in humans.* Revista Ciencias de la Salud, 7(2), 47-66.
- [36] Encuentros en la biología (2015). *Genómica del cáncer de pulmón.* Disponible en <http://www.encuentrosenlabiologia.es/2015/09/genomica-del-cancer-de-pulmon/>. [Consultada el 14.04,2016].
- [37] Moncho-Amor, V. (2011). *Papel de DUSP1 en progresión tumoral y en respuesta a CDDP en cáncer no microcítico de pulmón.* Tesis doctoral.
- [38] S. A Ana (2013). *Modelización. Computación y matemáticas contra el cáncer.* Universidad de Alcalá.
- [39] Goñi, J., and García, J. M. (2006). *Dinámica de Sistemas Biológicos: modelando complejidad.* Iniciación a la Investigación, (1).

- [40] Parrilla, A. A., and Váldez, A. E. *Autómatas Celulares: La Regla 150*.
- [41] Caligaris, M. G., Rodriguez, G. B., Marinsalta, M. M., Schivo, M. E., Romiti, M. R., Laugero, L. F., and Tello, J. M. *Asistencia computacional en la enseñanza en carreras de Ingeniería*. JEIN 2012, 8.
- [42] Cortes, F., and Ventura, S. (2016). *Modelado de reproducción e invasión de neoplasias malignas usando autómatas celulares*.
- [43] Martínez, G. J., and México, D. F. (2006). *Introducción a la simulación de procesos con autómatas celular*.
- [44] Cardoso, C., Bert, F., and Podestá, G. *Modelos Basados en Agentes (MBA): definición, alcances y limitaciones*.
- [45] Izquierdo, L. R., Ordax, J. M. G., Santos, J. I., and Martínez, R. D. O. (2008). *Modelado de sistemas complejos mediante simulación basada en agentes y mediante dinámica de sistemas*. Empiria. Revista de metodología de ciencias sociales, (16), 85-112.
- [46] Martínez Ortega, M. A. (2008). *Agentes inteligentes aplicados al reconocimiento de imágenes en el campo de la agricultura*.
- [47] Gisbert, M. G. (2015). *¿Qué pueden ofrecer los modelos basados en agentes vivos en el contexto docente?*. Modelling in Science Education and Learning, 8(2), 5-26.
- [48] Sancho, C. F. *Sistemas Multiagente y Simulación*. Disponible en <http://www.cs.us.es/~fsancho/?e=57> [Consultada el 31.08.2016]
- [49] Hoeger, H. *Pasos en un Estudio de simulación*. Curso de Modelado y Simulación I. Universidad de Los Andes. (2013).
- [50] Ortigosa, E. M. (2002). *Simulación en arquitecturas paralelas de la propagación de ondas en medios excitables*. Málaga.
- [51] Universidad Tecnológica Nacional. Facultad Regional San Nicolas (FRSN). *Métodos numéricos*. Disponible en <http://www.frsn.utn.edu.ar/GIE/AN/EDP/Conceptos.html>. [Consultada el 15.08.2016]

- [52] Universidad Complutense Madrid. Departamento de Bioquímica y Biología Molecular. *Sistemas con difusión*. Disponible en <http://solea.quim.ucm.es/fmoran/publications/librobiophysica/chap8.pdf>. [Consultada el 14.07.2016]
- [53] Hernández, D. P. D., and Herrera, L. C. B. (2002). *¿Cómo se transporta la glucosa a través de la membrana celular?*. *Iatreia*, 15(3), 179-189.
- [54] Mañas, J. A. (1997). *Análisis de algoritmos: Complejidad*.
- [55] Lomanto Díaz, L. D., Ortiz Cala, Ó. L., Bretón Pinto, C. O., Gómez Lizcano, Á. I., and Mesa Cornejo, V. M. (2010). *El ciclo celular*. *MedUNAB*, 6(16).

www.bdigital.ula.ve