



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES

UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES

CENTRO DE CIENCIAS BÁSICAS

Desarrollo de una Metodología para resolver problemas de
Optimización Combinatoria mediante el Sistema Inmune Artificial

TESIS

QUE PARA OBTENER EL GRADO DE
DOCTOR EN CIENCIAS DE LA COMPUTACIÓN
INTELIGENCIA ARTIFICIAL

Presenta:

FATIMA SAYURI QUEZADA AGUILERA

Tutor:

DR. FELIPE PADILLA DÍAZ

Comité Tutorial:

DRA. EUNICE ESTHER PONCE DE LEÓN SENTÍ
DRA. ELVA DÍAZ DÍAZ
DR. CARLOS ALBERTO OCHOA ORTIZ ZEZZATTI
DR. ALEJANDRO PADILLA DÍAZ

Aguascalientes, Ags. México. Julio de 2010

AGRADECIMIENTOS

Agradezco a Dios por la vida, la salud, la familia y todo lo que hasta este momento me fue concedido para dar este gran paso en mi vida profesional y personal.

Agradezco a mi mamá María Graciela Aguilera Guevara, a mis hermanos Andhy Everst Quezada Aguilera y Albret Sinaid Quezada Aguilera a mi sobrino Ian Alejandro Quezada Vázquez por su confianza y fe en mí, así como su amor y su apoyo incondicional, que fueron una gran motivación para seguir adelante.

Agradezco también a mi familia y amigos que de algún modo fueron parte de esto y me dieron su confianza y apoyo.

Agradezco a mis compañeros y amigos Dr. Julio César Ponce Gallegos, Dr. Francisco Javier Ornelas Zapata, Dr. Miguel Ángel Meza de Luna, Dra. Laura Cecilia Rodríguez, Dr. Christian José Correa Villalón, Dra. Aurora Torres Soto, Dra. Dolores Torres Soto, por su amistad, consejos, sonrisas, apoyo y por todo lo que compartimos durante este tiempo, muchas gracias.

Agradezco a mi tutor y comité tutorial, Dr. Felipe Padilla Díaz, Dra. Eunice Esther Ponce de León Sentí, Dra. Elva Diaz Diaz, Dr. Alberto Ochoa Ortiz Zezzatti, Dr. Alejandro Padilla Díaz, por sus consejos, su guía, su paciencia, su confianza, sus conocimientos y su presencia que fueron de suma importancia para lograr este objetivo.

Agradezco al Dr. Francisco Javier Álvarez Rodríguez por el gran apoyo otorgado para cumplir este objetivo y a todos mis profesores que fueron parte importante en mi formación académica.

Agradezco a la Universidad Autónoma de Aguascalientes por darme la oportunidad de estudiar y llegar a este punto en mi formación académica.

DEDICATORIA

Dedico esta tesis a mi mamá María Graciela Aguilera Guevara, a mis hermanos Andhy Everst Quezada Aguilera y Albret Sinaid Quezada Aguilera a mi sobrino Ian Alejandro Quezada Vázquez.



CARTAS DE LIBERACIÓN



**I.S.C. FÁTIMA SAYURI QUEZADA AGUILAR
PASANTE DEL DOCTORADO EN
CIENCIAS DE LA COMPUTACION
P R E S E N T E .**

Estimado (a) Alumno (a) Quezada:

Por medio de este conducto me permito comunicar a Usted que habiendo recibido los votos aprobatorios de los revisores de su trabajo de tesis titulado: **“Desarrollo de una Metodología para Resolver Problemas de Optimización Combinatoria mediante el Sistema Inmune Artificial”**, hago de su conocimiento que puede imprimir dicho documento y continuar con los trámites para la presentación de su examen de grado.

Sin otro particular me permito saludarle muy afectuosamente.

A T E N T A M E N T E
Aguascalientes, Ags., 2 de agosto de 2010
“LUMEN PROFERRE”
EL DECANO

DR. FRANCISCO JAVIER ÁLVAREZ RODRÍGUEZ



c.c.p.- Archivo

CARTAS DE LIBERACIÓN



Aguascalientes, Ags. Julio de 2010

Dr. Francisco Javier Álvarez Rodríguez
Decano del Centro de Ciencias Básicas
Universidad Autónoma de Aguascalientes
PRESENTE.

Por este conducto, hago de su conocimiento que la **ISC FATIMA SAYURI QUEZADA AGUILERA**, egresada del **doctorado en Ciencias de la Computación** del Centro de Ciencias Básicas de la Universidad Autónoma de Aguascalientes, ha integrado satisfactoriamente el documento de tesis titulado: **“Desarrollo de una Metodología para Resolver Problemas de Optimización Combinatoria mediante el Sistema Inmune Artificial”**, por lo que doy mi voto aprobatorio para que continúe con los trámites para presentar el examen de grado reglamentario.

Atentamente:



Dr. Felipe Padilla Díaz

Director de Tesis

CARTAS DE LIBERACIÓN



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES

Aguascalientes, Ags. Julio de 2010

Dr. Francisco Javier Álvarez Rodríguez

Decano del Centro de Ciencias Básicas

Universidad Autónoma de Aguascalientes

P R E S E N T E.

Por este conducto, hago de su conocimiento que la **ISC FATIMA SAYURI QUEZADA AGUILERA**, egresada del **doctorado en Ciencias de la Computación** del Centro de Ciencias Básicas de la Universidad Autónoma de Aguascalientes, ha integrado satisfactoriamente el documento de tesis titulado: **“Desarrollo de una Metodología para Resolver Problemas de Optimización Combinatoria mediante el Sistema Inmune Artificial”**, por lo que doy mi voto aprobatorio para que continúe con los trámites para presentar el examen de grado reglamentario.

Atentamente:

A handwritten signature in blue ink, appearing to read 'Eunice Ponce de León Sentí', written over a horizontal line.

Dra. Eunice Esther Ponce de León Sentí

Asesor de Tesis.

CARTAS DE LIBERACIÓN

Aguascalientes, Ags. Julio de 2010



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES

Dr. Francisco Javier Álvarez Rodríguez

Decano del Centro de Ciencias Básicas

Universidad Autónoma de Aguascalientes

P R E S E N T E.

Por este conducto, hago de su conocimiento que la **ISC FATIMA SAYURI QUEZADA AGUILERA**, egresada del **doctorado en Ciencias de la Computación** del Centro de Ciencias Básicas de la Universidad Autónoma de Aguascalientes, ha integrado satisfactoriamente el documento de tesis titulado: **“Desarrollo de una Metodología para Resolver Problemas de Optimización Combinatoria mediante el Sistema Inmune Artificial”**, por lo que doy mi voto aprobatorio para que continúe con los trámites para presentar el examen de grado reglamentario.

Atentamente:

A handwritten signature in blue ink, appearing to read 'Elva Díaz Díaz', written over a horizontal line.

Dra. Elva Díaz Díaz

Asesor de Tesis.

CARTAS DE LIBERACIÓN



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES

Aguascalientes, Ags. Julio de 2010

Dr. Francisco Javier Álvarez Rodríguez

Decano del Centro de Ciencias Básicas

Universidad Autónoma de Aguascalientes

P R E S E N T E.

Por este conducto, hago de su conocimiento que la **ISC FATIMA SAYURI QUEZADA AGUILERA**, egresada del **doctorado en Ciencias de la Computación** del Centro de Ciencias Básicas de la Universidad Autónoma de Aguascalientes, ha integrado satisfactoriamente el documento de tesis titulado: **“Desarrollo de una Metodología para Resolver Problemas de Optimización Combinatoria mediante el Sistema Inmune Artificial”**, por lo que doy mi voto aprobatorio para que continúe con los trámites para presentar el examen de grado reglamentario.

Atentamente:

A handwritten signature in black ink, appearing to read 'C. Ochoa', written over a horizontal line.

Dr. Carlos Alberto Ochoa Ortiz Zezzatti

Asesor de tesis

CARTAS DE LIBERACIÓN



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES

Aguascalientes, Ags. Julio de 2010

Dr. Francisco Javier Álvarez Rodríguez

Decano del Centro de Ciencias Básicas

Universidad Autónoma de Aguascalientes

PRESENTE.

Por este conducto, hago de su conocimiento que la **ISC FATIMA SAYURI QUEZADA AGUILERA**, egresada del **doctorado en Ciencias de la Computación** del Centro de Ciencias Básicas de la Universidad Autónoma de Aguascalientes, ha integrado satisfactoriamente el documento de tesis titulado: **“Desarrollo de una Metodología para Resolver Problemas de Optimización Combinatoria mediante el Sistema Inmune Artificial”**, por lo que doy mi voto aprobatorio para que continúe con los trámites para presentar el examen de grado reglamentario.

Atentamente:

A handwritten signature in blue ink, appearing to read 'Alejandro Padilla Díaz', written over a horizontal line.

Dr. Alejandro Padilla Díaz

Asesor de Tesis.

RESUMEN

Los algoritmos bio-inspirados han tenido gran atención las últimas décadas tanto por la originalidad de los mecanismos y procesos que son emulados como por su eficiencia, al ser modelado e implementados, para resolver una gran diversidad de problemas, en particular problemas NP-Duros.

En el grupo de los algoritmos bio-inspirados se encuentra el Sistema Inmune Artificial, el cual se considera está inspirado en un sistema sumamente complejo como lo es el Sistema Inmune Natural. Extraer características de este Sistema no es tarea fácil y más aún si no se tiene un área específica de aplicación. Considerando esta aseveración en este trabajo de tesis se realizó un esfuerzo para determinar una metodología que guíe en la resolución de problemas de Optimización Combinatoria considerando las características del Sistema Inmune que mejor se adapten a la solución de los mismos. Los problemas de Optimización Combinatoria son problemas intratables, ya que con algoritmos exactos no es posible resolverlos en un tiempo polinomial, ahí el uso de heurísticas para resolver estos problemas.

Como parte de la investigación se hizo análisis de la teoría del Sistema Inmune Natural del cual se tomaron características que se consideraron factibles para modelar y aplicar a la solución de problemas de Optimización Combinatoria. Los algoritmos desarrollados se implementaron para resolver el problema del Agente Viajero.

Como caso de estudio se considero la aplicación al problema de detección de intrusos el cual es un problema que es de suma importancia para la seguridad informática y que por las características del Sistema Inmune Artificial es factible que pueda ser aplicado para su solución.

Contenido

Contenido.....	XI
Lista de Figuras	XV
Lista de Tablas	XVII
Capítulo 1.....	1
Introducción	1
1.1 Motivación y Marco Teórico	1
1.2 Justificación y Relevancia.....	6
1.3 Estructura de la Tesis	10
En el capítulo 6 se detallan los resultados obtenidos en los experimentos que diseñados para probar el funcionamiento del algoritmo obtenido a partir de la metodología. Los problemas utilizados para estos experimentos son los propuestos en los objetivos de la tesis.	11
En el capítulo 7 se documenta lo que es la detección de malware y como ha sido resuelto como un problema de optimización. También se desarrolla la forma de resolverlo mediante el SIA.....	11

En el capítulo 8 se detallan las conclusiones respecto a la metodología y los resultados de los experimentos así como los que se considera que no quedó cubierto y es tema de investigaciones futuras..... 11

1.4 Publicaciones..... 11

Capítulo 2..... 12

Optimización Combinatoria 12

2.1 Problemas Combinatorios..... 12

2.2 Conceptos de Optimización..... 14

2.3 Optimización Combinatoria..... 17

2.4 Complejidad Algorítmica..... 19

2.5 Problemas Combinatorios..... 20

Capítulo 3..... 21

Sistema Inmune Natural 21

3.1 Definición de Sistema Inmune Natural..... 22

3.2 Órganos y células del Sistema Inmune Natural..... 23

3.2.1 Órganos del Sistema Inmune Natural..... 23

3.2.2 Células del Sistema Inmune Natural 27

3.3 Características de las Respuestas Inmunitarias..... 35

3.4 Fases de las Respuestas Inmunitarias 40

3.5 Respuestas inmunitarias primarias y secundarias..... 43

3.6 Estructura de los linfocitos 44

3.7 La recirculación de los linfocitos y el asentamiento linfocitario 46

3.8 La recirculación de los linfocitos T vírgenes 47

3.9 Migración de los linfocitos T Efectores y de Memoria..... 48

3.10 La recirculación de los linfocitos B..... 49

ÍNDICE DE CONTENIDO

3.12	Características generales de la estructura de los anticuerpos	56
3.13	Modelo del Funcionamiento del Sistema Inmune	68
Capítulo 4	71
Sistema Inmune Artificial	71
4.1	Definición de Sistema Inmune Artificial.....	72
4.2	Características del SIA.....	73
4.3	Marco de trabajo para el SIA.....	75
4.4	Modelos del SIA	78
4.5	Mapeo del SIN y del AIS	87
4.6	Comparación del Sistema Inmune Artificial con otras Heurísticas	87
4.7	Aplicación de algoritmos del SIA en Optimización	89
Capítulo 5	97
Metodología propuesta	97
5.1.	Introducción.....	97
5.2.	Análisis de Metodologías y Marcos de trabajo	98
5.3.	Desarrollo de la Metodología	102
5.3	Relación entre los Elementos del SIA, SIN y el Problema.....	112
Capítulo 6	115
Experimentación	115
6.1.	Desarrollo de la Metodología	115
6.2.	Análisis de los experimentos con el TSP Simétrico	117
Capítulo 7	123
Caso de Estudio	123
7.1	Problema de detección de Malware.....	¡Error! Marcador no definido.

TESIS TESIS TESIS TESIS TESIS

ÍNDICE DE CONTENIDO

7.2	Sistemas de detección de intrusos	123
7.3.	Técnicas de Data mining para detección de Intrusos	126
7.4.	Data mining y los IDS.....	127
Capítulo 8.....	¡Error! Marcador no definido.	
Conclusiones y Trabajo Futuro	¡Error! Marcador no definido.	
Bibliografía.....		168



Índice de Figuras

1: Clasificación de las soluciones de un problema.....	14
2: Clasificación de los problemas de acuerdo a su complejidad algorítmica para ser resueltos.....	20
3 Clasificación de las células que conforman al SIN de acuerdo al tipo de inmunidad al que pertenecen.....	27
4 Clasificación de las células del SIN de acuerdo a la parte del organismo donde se pueden localizar. Tomada del libro Immunobiology (Janeway, 2001).....	28
5 Representación de un linfocito B y su receptor (BCR del inglés B Cell Receptor) o anticuerpo.....	31
6 Duración y principales fases de las respuestas inmunitarias.....	41
7 Grafica de la duración y efectividad de las respuestas inmunitarias primaria y de las respuestas inmunitarias secundarias.....	44
8 Ejemplificación de la teoría de la Selección Clonal. (Janeway, 2001).....	51
9: Modelo aproximado del funcionamiento general del Sistema Inmune Natural de acuerdo a los conceptos tomados del mismo para esta investigación.....	69
10 Marco de Trabajo para el SIA propuesto por Leandro Nunes de Castro y Jonathan Timmis (De Castro et al., 2002; De Castro et al., 2003).....	76
11: Modelos de la generación del repertorio inicial de anticuerpos (De Castro et al., 2000).....	80

ÍNDICE DE FIGURAS

12: Modelo del Algoritmo de Selección Negativa, Forrest et al. 1994	80
13: Modelo del Algoritmo de Selección Clonal utilizado en opitmización, aprendizaje de máquina y reconocimiento de patrones, De Castro et al. 2002.....	83
14 Ejemplo de cadenas modificadas par la Hipermutación Somática.....	86
15. Marco de trabajo Conceptual para Sistema Inmune Artificiales propuesto por Stepney et al., 2004.....	99
16: Correspondencia del funcionamiento del SIN hacia el SIA. Esta es la etapa inicial en la que se considera la formación de las células (Linfocitos T y B).....	105
17: Correspondencia del funcionamiento del SIN hacia el SIA. Esta es la etapa en la que se consideran maduros los linfocitos (buenas soluciones) y se decide forman parte del SIN.	109
18: Correspondencia del funcionamiento del SIN hacia el SIA. Esta es la etapa en que se activan los diferentes mecanismos para atacar al antígeno.	109
19. Diagrama del funcionamiento del algoritmo que se propone, basado en el CLONALG	110
20: Correspondencia del funcionamiento del SIN hacia el SIA. Esta son funciones que se llevan a cabo constantemente independientemente de si se ha activado o no una respuesta inmunitaria.	111
21 Estrategia de Análisis de los diferentes IDS, (Zurutuza et al., 2005).....	128

Índice de Tablas

1 Clasificación de los Órganos Linfoides que constituyen al Sistema Inmune Natural y las funciones que desempeñan los mismos.....	26
2 Clasificación de los elementos que conforman la inmunidad innata y la adaptativa de acuerdo a su comportamiento.....	34
3 Características principales de las respuestas inmunitarias.	38
4 Resumen de trabajos relevantes en los que se utiliza el modelo de Selección Negativa.....	81
5: Mapeo de los elementos que conforman el SI y los que corresponden el AIS, Timmis et. al 2002.	87
6: Comparación entre el AG y el SIA.	88
7: Comparación entre el AG y RN.....	89
8 Clasificación de los mecanismos de exploración y explotación del SIN y que son integrados al AIS.	96
9: Mapeo del Sistema Inmune Natural con el Sistema Inmune Artificial.....	111
10. Correspondencias entre el AIS, el problema a resolver y el SIN	113
11: Valores de cada parámetro para el experimento con el benchmark Burman 14..	118
12 Valores de cada parámetro para el experimento con el benchmark Baygs 29.	120
13 Resultados de los experimentos realizados cone el benchmark Burman 14.....	152

ÍNDICE DE TABLAS

14 Resultados de los experimentos realizados con el benchmark Baygs 29.....157



Capítulo 1

Introducción

En este capítulo se contextualiza el problema de investigación así como el planteamiento del mismo. Se hace una descripción del enfoque de investigación.

A partir de la sección 1.2 se define el problema así como los objetivos de investigación, las hipótesis. En la sección 1.3 se describe la estructura de la tesis. En la sección 1.4 se mencionan las contribuciones.

1.1 Motivación y Marco Teórico

En la vida cotidiana pudiéramos decir “En el mundo real”, existe una gran cantidad de problemas que están en espera de ser solucionados. Estos se encuentran en las

diferentes áreas del conocimiento, como la Sociología, la Medicina, la Biología y por supuesto, en las Ciencias de la Computación. Muchos de estos problemas se pueden plantear como Problemas de Optimización Combinatoria (OC) NP-Complejos o NP-Duros. Estos problemas han sido atacados y resueltos, hasta ciertas instancias, haciendo uso de diferentes técnicas como las Heurísticas y Metaheurísticas. Dichas técnicas, se inspiran en diversas fuentes como la Biología, surgiendo así los algoritmos Bio-inspirados, los cuales han sido eficientes en la solución de diversos problemas.

El Sistema Inmune Artificial (SIA), siendo una heurística bio-inspirada, ha sido utilizado eficientemente en la solución de ciertos problemas. Sin embargo, el como desarrollar apropiadamente estos algoritmos es todavía un tema de investigación, en el cual han aportado importantes resultados investigadores como Leandro Nunes de Castro, Jonathan Timmis, entre otros (De Castro et al., 2002, 2003), (De Castro et al., 2003) (Stepney et al., 2005). Es por eso que el trabajo de esta tesis está orientado al como desarrollar algoritmos del SIA, que emulen de forma más cercana al SIN, para resolver problemas de OC y que además sean competitivos en relación con lo reportado en la literatura.

1.1.1 Contexto

Los problemas de Optimización Combinatoria (OC) por sus características se clasifican como NP-Complejos o NP-Duros. Dicho adjetivo se les da a los problemas que no pueden ser resueltos en un tiempo polinomial, dada su complejidad, por lo tanto un algoritmo clásico (Ej. Simplex, entre otros) no es el indicado para resolverlos a partir de ciertas instancias. Para encontrar la solución a estos problemas han surgido las heurísticas, que son clasificadas de acuerdo a características como la fuente de inspiración. Las heurísticas bio-inspiradas toman conceptos, modelos, entre otros, de algún sistema biológico y tratan de emularlo para solucionar problemas y ser capaz de optimizar recursos, aprender y adaptarse a nuevos escenarios, como la misma naturaleza lo hace.

El Sistema Inmune Artificial (SIA) surge como una alternativa, en la cual los algoritmos desarrollados se inspiran en el funcionamiento del Sistema Inmune Natural (SIN), principalmente en el del ser humano. Esta heurística bio-inspirada tiene relativamente poco tiempo de haber tomado un mayor auge, siendo aplicada a la solución de diversos problemas, entre ellos los problemas de Optimización Combinatoria.

Dado que es una heurística relativamente reciente, existe la necesidad de describir las bases teóricas del Sistema Inmune Natural tanto para la comprensión, como para el análisis, diseño y la implementación de nuevos algoritmos. Esto lo podemos notar principalmente en la carencia de un algoritmo único, pues no existe una metodología formal para desarrollarlos a diferencia de otras heurísticas bio-inspiradas como los Algoritmos Genéticos (AGs).

Leandro Nunes de Castro y Jonathan Timmis propusieron un marco de trabajo para el desarrollo de algoritmos del SIA, lo que son los primeros cimientos para conducir la investigación en este sentido (De Castro et al., 2002), (De Castro et al., 2003). Otro trabajo relacionado es el realizado por Susan Stepney y colegas (Stepney et al., 2005), en el cuál se propone un marco de trabajo conceptual para el desarrollo de Sistemas Inmunes Artificiales.

1.1.2 Sistema Inmune Natural

El Sistema SIN es el encargado de mantener en buen estado de salud al organismo, es decir libre de enfermedades, infecciones, etc. causadas por la presencia de agentes externos dañinos como sustancias extrañas, microorganismos, macromoléculas (proteínas y polisacáridos). Esto se conoce como Inmunidad que proviene de la palabra latina *immunitas* y que para esta área significa protección frente a la enfermedad infecciosa.

Este es un sistema bastante complejo que ha evolucionado durante miles de años, el cuál está formado por diversos tipos de células y moléculas que al actuar de manera global y coordinada dan origen a la **respuesta inmunitaria**.

El SI tiene las funciones de vigilar, detectar o reconocer una amenaza a la salud y activar los mecanismos necesarios para destruir tal amenaza.

✓ **Inmunidad innata**

La inmunidad innata tiene como características la capacidad de diferenciar un microorganismo de otro y actuar de manera similar ante un agente externo llamado también *antígeno*. Este tipo de inmunidad no requiere de un contacto anterior con el antígeno y está determinada por factores genéticos, raciales, hormonales, humorales, celulares, así como la edad y las barreras de protección mecánica.

✓ **Inmunidad específica o adaptativa**

La respuesta inmunitaria específica o adaptativa se produce frecuentemente después de la exposición de un organismo con un antígeno extraño. La inmunidad específica se puede adquirir de manera activa o pasiva. La primera se da cuando el individuo juega un papel activo, es decir él crea las defensas necesarias para atacar al antígeno. La segunda se lleva a cabo transfiriendo células o suero de un individuo previamente inmunizado. Las principales células de de este tipo de inmunidad son los linfocitos B y los linfocitos T.

Además de esto tiene otras características interesantes como la memoria, la diversidad, la tolerancia, etc.

✓ Teorías del sistema inmune

- Selección Clonal
- Selección Positiva
- Selección Negativa
- Red Inmune
- Teoría del Peligro

1.1.3 Sistema Inmune Artificial

Basado en lo dicho por Leandro Nunes de Castro y Jonathan Timmis (De Castro et al., 2002) el SIA se define de la siguiente manera:

“Los sistemas inmunes artificiales son sistemas adaptativos inspirados por la teoría inmunológica, funciones, principios y modelos inmunológicos observados, los cuales son aplicados a la solución de problemas NP-Complejos”

El SI cuenta con ciertas características que se consideran importantes en un sistema inteligente y que por tanto esta nueva área del SIA ha tratado de emular para la solución a diversos problemas. Las características son robustez, diversidad, memoria, entre otras (Cruz, 2004) que han sido tomadas del SI y que se han tratado de emular en el SIA.

Ya que existen gran variedad de algoritmos y modelos, Leandro Nunes De Castro y Jonathan Timmis hacen una clasificación de estos (De Castro et al., 2002):

- Modelos de timo y de médula ósea
- Algoritmo de selección clonal
- Teoría de red inmune

Aunque el más utilizado para problemas de optimización combinatoria ha sido el modelo basado en la selección clonal.

1.1.4 Optimización Combinatoria

La Optimización Combinatoria se puede definir como la rama de la Optimización que afronta los problemas con un dominio X finito y variables discretas, conocidos como *problemas combinatorios* (Fernández, 1996). Algunos de estos problemas pertenecen a la clase NP-Duros por su complejidad algorítmica para ser resueltos, lo cual ha sido demostrado por Garey y Johnson (Garey et al., 1979).

Estos problemas los podemos encontrar en la realidad en diversas áreas del conocimiento como la medicina, la ingeniería, la sociología, etc. Sin embargo por ser problemas intratables, se han desarrollado las heurísticas para ser resueltos.

1.2 Justificación y Relevancia

Los algoritmos bio-inspirados han resultado muy eficientes para solucionar problemas, sobre todo cuando el sistema biológico usado es bien comprendido y por lo tanto correctamente emulado.

En este contexto los Sistemas Inmune Artificiales son algoritmos que se han utilizado para resolver problemas con resultados muy buenos y que aún pueden ser mejorados incorporando conceptos de la Inmunología que todavía no se han considerado. Además de esto, todavía hay trabajo por hacer en lo que se refiere a la definición clara del cómo desarrollar algoritmos del SIA, incluso cuando se quieren aplicar en la solución de determinado tipo de problemas.

CAPÍTULO I.-INTRODUCCIÓN

En relación a esto, existen ya investigaciones en las que se proponen algoritmos del SIA para la solución de problemas de OC y en los que se demuestran buenos resultados, por lo cual se considera una heurística con un gran potencial y que puede competir con técnicas ya consolidadas como los AG o Colonia de Hormigas (ACO, por sus siglas en inglés). Estas investigaciones son relativamente recientes, aproximadamente 8 años y son los primeros acercamientos a la formalización en el desarrollo de estos algoritmos. En investigaciones posteriores se propone un marco de trabajo para el SIA, aunque no una metodología formal.

Además de ser una técnica relativamente joven en la que hay mucho por proponer, el SIA toma características del SIN que son necesarias en cualquier sistema inteligente como:

- Efectividad
- Eficiencia
- Adaptabilidad
- Robustez, etc.

Es por tales razones que en esta tesis se exploran tanto el funcionamiento del SIN y las interrelaciones entre las diferentes células así como conceptos y modelos para ser emulados de la manera más cercana y poder dar una guía de cómo generar algoritmos del AIS a partir de seguir dicho funcionamiento para la solución de problemas de OC.

Los problemas que se van a tratar en esta tesis son mono-objetivo pues se ha demostrado que en problemas como el Job Shop Scheduling (JSSP, por sus siglas en inglés), que es un problema Np-Duro que ha sido resuelto mediante un AIS, se puede resolver la versión de un solo objetivo y a partir de ahí resolver los demás objetivos de forma más sencilla (Cortes, 2004). Esto puede aplicarse también a problemas como el Travelling Salesman Problem (TSP, por sus siglas en inglés) ya que el JSSP es una instancia de este problema (KeldHelsgaun, 1997).

1.2.1 Descripción General del Enfoque de Investigación

El enfoque de la investigación es de modelación, simulación y experimentación pues se desarrollará la metodología para diseñar el algoritmo para la solución de problemas OC.

Este será implementado en la computadora a través de un lenguaje de programación (C++ Builder, CodeGear 2009). Se diseñarán los experimentos definiendo los parámetros con los que funcione mejor el algoritmo en la solución de cada problema seleccionado.

1.2.2 Descripción del Problema de Investigación

El problema de la investigación se enfoca principalmente en el desarrollo de una metodología formal que pueda ser utilizada como una guía para solucionar problemas Optimización Combinatoria con el paradigma del Sistema Inmune Artificial.

Es por tal razón que el problema involucra el desarrollo de algoritmos más apegados al Sistema Inmune Natural y que estén ajustados al problema a resolver.

Para los cual la será metodología una guía que explique y justifique como se llegó a la creación de un algoritmo par resolver un problema a partir de un concepto o función que se modela del Sistema Inmune Natural.

Considerando que la variedad de problemas es muy amplia, en la tesis se abordarán 2; el problema del agente viajero (Travelling Salesman Problem, TSP en inglés) y el problema para modelar problemas de detección de malware.

1.2.3 Objetivo General

- ✓ Desarrollar una metodología para el diseño e implementación de algoritmos del SIA para atacar problemas de OC.

1.2.4 Objetivos Específicos

1. Describir una metodología para el desarrollo de algoritmos del Sistema Inmune Artificial para la solución de problemas de Optimización Combinatoria.
2. Utilizar mecanismos de otras heurísticas o técnicas de búsqueda local para mejorar el desempeño del Sistema Inmune Artificial en la solución de los problemas.
3. Diseñar e implementar un algoritmo basado en la metodología desarrollada.
4. Aplicar el algoritmo desarrollado al problema de Optimización Combinatoria: Viajante de Comercio (TSP, por sus siglas en inglés).
5. Diseñar experimentos y hacer comparaciones con otras metodologías del Sistema Inmune Artificial o con otros algoritmos.
6. Resolver una aplicación real que pueda ser modelada como algunos de los problemas de OC seleccionados. Esta aplicación se orienta a la detección de software malicioso o malintencionado (malware, en inglés).

1.2.5 Hipótesis

1. ¿Es posible crear una metodología que sirva como base para diseñar algoritmos del SIA para resolver problemas de OC?
2. ¿Es posible diseñar e implementar un algoritmo del SIA para resolver los problemas de OC a partir de la metodología?
3. ¿Los algoritmos diseñados a través de esta metodología pueden mejorar o igualar los resultados reportados en la literatura?

1.3 Estructura de la Tesis

En el capítulo 2 se hace una revisión de los conceptos fundamentales de la Optimización Combinatoria. También se abordan los temas referentes a la complejidad computacional de los problemas de Optimización Combinatoria y se muestra una clasificación de los mismos.

En el capítulo 3 se hace una revisión del estado del arte de l Sistema Inmune Natural, tomando en cuenta principalmente los conceptos, teorías, elementos y métodos que se ha considerando de interés para el desarrollo de la tesis.

En el capítulo 4 se hace una revisión del estado del arte del Sistema Inmune Artificial, mostrando los algoritmos más importantes y haciendo énfasis en los que han sido utilizados en Optimización.

En el capítulo 5 se realiza el desarrollo de la metodología propuesta para la solución de problemas de optimización combinatoria. En dicho capítulo se detallan las fases en que consistirá la metodología y los elementos a considerar en cada fase.

En el capítulo 6 se detallan los resultados obtenidos en los experimentos que diseñados para probar el funcionamiento del algoritmo obtenido a partir de la metodología. Los problemas utilizados para estos experimentos son los propuestos en los objetivos de la tesis.

En el capítulo 7 se documenta lo que es la detección de malware y como ha sido resuelto como un problema de optimización. También se desarrolla la forma de resolverlo mediante el SIA.

En el capítulo 8 se detallan las conclusiones respecto a la metodología y los resultados de los experimentos así como los que se considera que no quedó cubierto y es tema de investigaciones futuras.

1.4 Publicaciones

Sistema Inmune Artificial. F. S. Quezada Aguilera, F. Padilla Díaz, E. E. Ponce de León Sentí, A. Ochoa Zezzatti, C. J. Correa Villalón. COMCEV 2007. pp 85-88. ISBN 970-728-055-7. 2007.

Methodology to solve Combinatorial Optimization Problems with the Artificial Immune System. Fatima Quezada, Felipe Padilla, Eunice Ponce, Alberto Ochoa and Christian Correa. MICAI 2009, Workshop Artificial Intelligence Applications, p.p. 5-10. ISBN978-607-95367-0-1.Ed.Gelbukh.

Artificial Immune System to Solve Symmetric Travelling Salesman Problem. Sometido MICAI 2010.

Capítulo 2

Optimización Combinatoria

En este capítulo se hace una revisión de los conceptos relacionados con la Optimización Combinatoria así como los problemas con comportamiento NP- Duro. Se hace una recapitulación de los que la Optimización y se da la clasificación de los problemas por su complejidad computacional.

2.1 Problemas Combinatorios

Los problemas son circunstancias que se presentan constantemente en los diferentes ámbitos en que se desenvuelven los seres humanos y en las interrelaciones con los elementos que conforman su medio ambiente (natural, social, etc.).

Según lo dicho por Juan José Salazar (Salazar, 1996) un problema se define de la siguiente manera: *“Un problema es la descripción de una situación, más una pregunta*

CAPÍTULO II.- OPTIMIZACION COMBINATORIA

sobre esta. La descripción de la situación está en función de ciertos parámetros ρ que se llaman datos de entrada (“**input**”), y la respuesta puede representarse mediante unas variables σ que constituyen los llamados datos de salida (“**output**”), o propiamente respuesta. Tanto σ como ρ deberán estar codificados como cadenas de caracteres de un cierto alfabeto V (es decir, elementos del cierre de Kleene V^*). En tal sentido, los problemas se pueden representar como conjuntos $\{(\rho; \sigma)\}$ en el producto $V^* \times V^*$. Para cada ρ fijo, cuando el subconjunto $\{(\rho; \sigma)\}$ es no vacío se conoce como ejemplo (“**instance**”) del problema, aunque en ocasiones se confunde con el propio término “problema”. Se llama tamaño del ejemplo definido por los datos de entrada ρ a la longitud de esta cadena (es decir, al número de caracteres iguales o distintos para codificar ρ con el alfabeto V).”

Existe una gran cantidad de problemas tipo, entre ellos encontramos los Problemas Combinatorios (PC) que son en los que se trata de encontrar grupos, ordenamientos, o asignaciones de un conjunto discreto, finito de objetos que satisfacen ciertas condiciones o restricciones. Las combinaciones de estos componentes forman las soluciones potenciales de estos problemas (Hoos et al., 2005).

Este tipo de problemas combinatorios los podemos encontrar en diferentes áreas del conocimiento como la bioinformática, investigación de operaciones, medicina, biología, sociología, etc.

Para estos problemas las soluciones que se generan y que no cumplen todas las restricciones definidas se les conoce como *soluciones candidatas* (Hoos et al., 2005). Las *soluciones factibles* son aquellas que cumplen las restricciones establecidas, en las cuales, obtiene el óptimo o solución de problema.

Según Hoos y Stützle (Hoos et al., 2005) los problemas combinatorios se clasifican en problemas de Decisión y problemas de Optimización. Los primeros se subdividen en búsqueda, en los que dada una instancia del problema, se desea encontrar la solución (o determinar que la solución no existe); y los de decisión, en los que dada una instancia del problema se desea contestar una pregunta independientemente de si la solución existe o no. Los segundos se verán en la sección 2.3

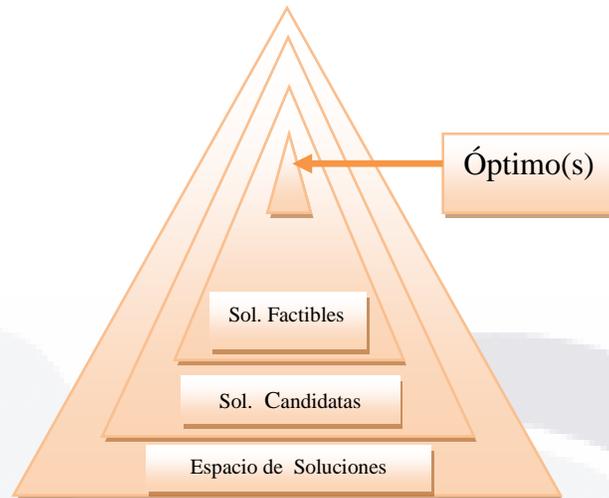


Figura 1: Clasificación de las soluciones de un problema

2.2 Conceptos de Optimización

La optimización o programación matemática tiene su origen desde épocas antiguas ya que era aplicada en la solución de problemas de geometría y mecánica. Hoy día las aplicaciones de la optimización son muchas y una de las áreas que más ha hecho uso de ella es la economía, aunque también es aplicada en control, producción, etc.

La Optimización busca solucionar los problemas encontrando el máximo o el mínimo de una cierta función objetivo en un cierto dominio X . Las teorías clásicas de Optimización (Cálculo Diferencial, Teoría del Control Óptimo, etc.) afrontan los problemas cuando el dominio X es infinito (y la función objetivo tiene determinadas características específicas). (Salazar, 1996).

La forma más simple de representar un problema de optimización es (Bector et al, 2005):

$$\min f(x)$$

sujeito a $x \in C$ (también llamado problema de programación matemática, PM)

CAPÍTULO II.- OPTIMIZACION COMBINATORIA

Donde $f : \mathfrak{R}^n \rightarrow \mathfrak{R}$ y $C \subseteq \mathfrak{R}^n$. f es la función objetivo y C el conjunto de restricciones. Cuando C es un conjunto propio de \mathfrak{R}^n se dice que es un problema con restricciones. Si $C = \mathfrak{R}^n$ entonces es un problema con restricciones.

Se dice que de A el punto $x_0 \in C$ es un mínimo local del problema si existe un $\delta > 0$ tal que $f(x) \geq f(x_0)$ para todo $x \in B_\delta(x_0) \cap C$. Se dice que de A el punto $x_0 \in C$ es un mínimo o mínimo global del problema si $f(x) \geq f(x_0)$ para todo $x \in C$. Si $C = \emptyset$ entonces $\min_{x \in C} \max f(x) = -\min(-f(x))$.

$\min f(x)$

sujeto a $g_i(x) \leq 0, i = 1, 2, \dots, m$

$$h_j(x) = 0, j = 1, 2, \dots, k \quad x \in X$$

Las funciones f y cada g_i y h_j son funciones evaluada en los números reales, definidas en \mathfrak{R}^n y $X \subseteq \mathfrak{R}^n$. Este problema es llamado por lo general problema de programación matemática general. El conjunto de restricciones C está dado por $C = \{x \in X : g_i(x) \leq 0, \forall i, h_j(x) = 0, \forall j\}$

Las restricciones $g_i(x) \leq 0, i = 1, 2, \dots, m$ son llamadas restricciones de desigualdad, las restricciones $h_j(x) = 0, j = 1, 2, \dots, k$ son llamadas restricciones de igualdad y $x \in X$ como una restricción abstracta. Si la función objetivo y las funciones de restricciones son diferenciables entonces el problema anterior es **programa diferenciable** y si el objetivo y las restricciones de desigualdad son funciones convexas, y las restricciones de igualdad funciones afines y X un conjunto convexo, entonces el problema anterior es **problema de programación convexo**.

Los problemas en los que se aplica la optimización son variados pues dependen del tipo de variable, la función objetivo, las restricciones, etc., como se puede ver en los ejemplos anteriormente enunciados. A continuación se muestra una clasificación posible según Fernández (Fernández, 1996).

2.2.1 Modelos según:

1. La naturaleza de los datos:

- ✓ Modelos Deterministas: Problemas donde se conocen con exactitud los datos que intervienen en el modelo
- ✓ Modelos Estocásticos: Problemas donde algunos o todos los datos dependen de fenómenos aleatorios.

2. La variable tiempo:

- ✓ Modelos Estáticos: La variable tiempo no se toma en consideración. Se tienen: Optimización o programación estática, Programación Clásica, Programación No Lineal, Programación Lineal y Teoría de Juegos.
- ✓ Modelos Dinámicos: Cuando se considera la variable tiempo de forma explícita en el modelo. Se tienen: Optimización o programación dinámica, El principio del máximo, Juegos diferenciales, etc.

3. Los objetivos del problema:

- ✓ Modelos de un único objetivo
- ✓ Modelos Multiobjetivo

4. Las restricciones:

- ✓ Modelos libres
- ✓ Modelos con restricciones

5. La linealidad:

- ✓ Modelos Lineales: Todas las funciones que intervienen son lineales
- ✓ Modelos No Lineales: Cuando al menos una de las funciones que interviene es no lineal.

6. El tipo de variables:

- ✓ Modelos Continuos: Todas las variables son continuas.
- ✓ Modelos Discretos: Al menos una de las variables únicamente puede tomar valores enteros.

La gran mayoría de los problemas que se desean resolver por sus características caen dentro de más de un modelo. De acuerdo al tipo de problemas (modelos) en que la Optimización es utilizada para resolverlos, esta se puede clasificar como Optimización, Multiobjetivo, Optimización Discreta, Optimización Combinatoria (OC), etc., esta última es de interés en esta tesis.

La OC se puede definir como la rama de la Optimización que afronta los problemas con un dominio X finito y variables discretas, conocidos como *problemas combinatorios* (Fernández, 1996), mencionados en la sección anterior.

2.3 Optimización Combinatoria

Actualmente los Problemas de Optimización Combinatoria (POC) se pueden encontrar en todas partes que no necesariamente están directamente relacionadas con las matemáticas o la ciencia de la computación. Las aplicaciones son tan diversas como se pueda imaginar, por ejemplo, secuenciación de máquinas, planificación de producción, diseño y localización de factorías, problemas de ventas, localización de plantas industriales, estudios de códigos genéticos, clasificación de plantas y animales,

CAPÍTULO II.- OPTIMIZACION COMBINATORIA

diseño de nuevas moléculas, asignación controlada de ondas de radios, construcción de compiladores y bases de datos, trazado de redes de comunicación robustas, posicionamiento de satélites, diseño y producción de circuitos VLSI, tamaño de flotas de camiones y planificación del transporte, desarrollo de sistemas de transportes de masas y secuenciación de autobuses y trenes, asignación de trabajadores a tareas tales como pilotos a aviones, embalaje de mercancías, codificación de información, incluso en áreas como deportes, arqueología o psicología. (Salazar, 1996)

Anteriormente los problemas combinatorios eran vistos como simples y triviales en la forma en que estos podían ser resueltos, sin embargo, una característica de estos problemas es que al crecer el número de variables involucradas crece de manera exponencial el número de posibles soluciones. Por ejemplo, si una supervisor tiene que visitar 5 escuelas y regresar a la escuela en que inició el recorrido, las soluciones candidatas serían 120 ($5!$). Enumerar estas soluciones para una persona resulta laborioso pero factible, es algo con lo que ella puede tratar. Pero, si en lugar de 5 escuelas necesitara visitar 15, las soluciones candidatas serían 1307674368000 ($15!$), lo cual es impráctico enumerar manualmente, por lo que consideraríamos la herramienta perfecta un equipo de cómputo. No obstante, como podemos apreciar, usando un modelo exhaustivo sería difícil resolver problemas con un número mayor de variables, aún para una computadora de gran poder de procesamiento, pues estos problemas se vuelven intratables a partir de ciertas instancias y tamaños.

Desafortunadamente, los POC que se encuentran en áreas como la Biología, la Economía, la Sociología, etc. son muy grandes, por lo cuál es necesario recurrir a técnicas de resolución diferentes y que tengan un costo tiempo-cómputo (recursos) razonable.

Estos problemas constan de una o varias funciones objetivo las cuáles se desean maximizar o minimizar y sus respectivas restricciones. Estas son representadas o definidas como se describió en la sección 2.2. Las restricciones permiten evaluar si la

solución generada es factible y la función objetivo, evaluar que tan buena es dicha solución (*calidad*).

Estos problemas pueden ser vistos como grafos en los cuales se busca la solución....

2.4 Complejidad Algorítmica

Como se mencionó en la sección 2.2, los problemas de OC tienen un grado de complejidad alto, ya que lograr solucionar ciertos tamaños de dichos problemas resulta prácticamente imposible para un algoritmo exhaustivo, considerando que no siempre se cuenta con equipo de cómputo de gran poder y aún teniéndolo estos problemas pueden volverse intratables.

Un problema se llama polinomial cuando se conoce un algoritmo polinomial que lo resuelva. La familia de los problemas polinomiales se conoce como clase P. Un problema se dice NP (o que está en la clase NP) cuando existe un algoritmo polinomial que confirma la pertenencia al problema de cualquier elemento $(\rho; \sigma)$ que efectivamente lo sea. Tales algoritmos se llaman certificados de factibilidad. Es evidente que $P \subseteq NP$, y la gran pregunta de la Complejidad Algorítmica es “¿ $NP \subseteq P$?”. Un problema se llama NP-difícil si dicho problema es NP y para cada problema de la clase NP existe un algoritmo polinomial que transforma cualquier ejemplo de este en un ejemplo de aquél. Por tanto configura una subclase de NP (representada por NPC) que contiene sus problemas más difíciles. Se conoce que tal subclase es no vacía, y que dentro destacan por su simplicidad precisamente muchos problemas combinatorios. Dentro de la clase NPC se puede a su vez distinguir entre los problemas NP-difíciles en sentido débil, y los problemas NP-difíciles en sentido fuerte, según que se conozca un algoritmo pseudo-polinomial o no, respectivamente, para su resolución. Los problemas que se tratan en esta memoria pertenecen a este último subgrupo, esto es, son problemas NP-difíciles en sentido fuerte.

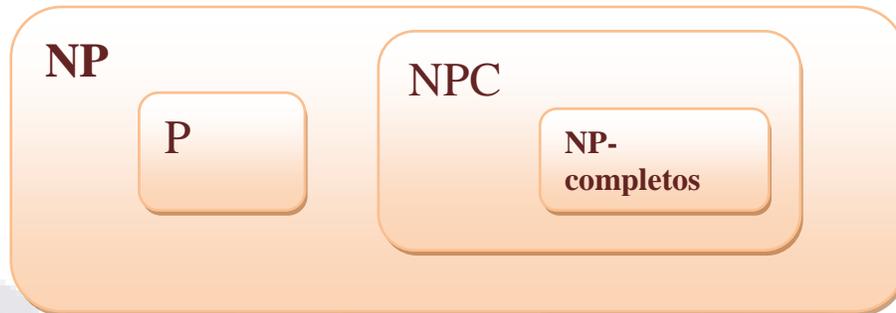


Figura 2: Clasificación de los problemas de acuerdo a su complejidad algorítmica para ser resueltos

2.5 Problemas Combinatorios

Ya se ha mencionado que los Problemas Combinatorios se encuentran en diferentes áreas como la Ciencia de la Computación, investigación de operaciones, bioinformática y comercio electrónico.

En este tipo de problemas se trata de encontrar grupos, ordenamientos, o asignaciones de un conjunto discreto, finito de objetos que satisfacen ciertas condiciones o restricciones. Las combinaciones de estos componentes forman las soluciones potenciales de estos problemas (Hoos et al., 2005).

Estos problemas se clasifican en:

Problemas de decisión

Problemas de optimización, son vistos como una generalización de los problemas de decisión.

En los 2 casos existen 2 variantes la de búsqueda y la de evaluación.

Capítulo 3

Sistema Inmune Natural

La inmunología es una ciencia que ha sufrido grandes cambios en su forma de estudiar el Sistema Inmune gracias al desarrollo de nuevas técnicas como la cristalografía por rayos X, la inmunoquímica, etc. Dichos avances han permitido un conocimiento más completo y profundo de los que es el Sistema Inmune Natural y sus interacciones.

La explicación de las funciones, mecanismos, células, etc. que conforman lo conforman ha sido la inspiración para la creación del Sistema Inmune Artificial.

En este capítulo se define lo que es el Sistema Inmune Natural Vertebrado específicamente el del ser Humano. Se hace una descripción un nivel de mayor detalle en cuanto a los órganos, moléculas y células que lo constituyen así como de los diferentes mecanismos con los que detecta, ataca y elimina a las amenazas que ponen en peligro al organismo.

Se hace mayor énfasis en etapas como la recirculación linfocitaria, monitoreo y muestreo, así como en algunas funciones y células de la inmunidad innata.

3.1 Definición de Sistema Inmune Natural

Existen diferentes definiciones para el Sistema Inmune, pero para iniciar, ¿Qué significa inmune? Según inmunidad proviene de la palabra latina “inmunitas”, que significaba que los senadores romanos estaban exentos de ciertas obligaciones civiles y procesos legales.

Así pues la palabra **inmunidad**, desde un punto de vista biológico o médico, denota la protección o resistencia que poseen los individuos frente a ciertas enfermedades, en especial las infecciosas.

Dado que un sistema es un conjunto de cosas o elementos que relacionadas entre sí ordenadamente contribuyen a alcanzar un objetivo común, para el caso del Sistema Inmune, dichos elementos son las células y las moléculas que proporcionan la **inmunidad**. La **respuesta inmunitaria** corresponde a la colaboración entre estas células y moléculas para defender al organismo de lo extraño y/o dañino.

Los agentes extraños dañinos y/o infecciosos que pueden ser macromoléculas, proteínas, microorganismos, virus, etc. son llamados antígenos, palabra que los engloba a todos para tratarlos de una forma más general, cuando no es necesario especificar su naturaleza.

Claramente podemos ver que las palabras inmune, inmunidad e inmunitario, no necesariamente tienen un mismo significado aunque si están relacionadas. Inmune, en este caso quiere decir que tiene la capacidad para enfrentar amenazas, tenga éxito o no.

Inmunidad, es cuando ya se tiene la protección contra alguna amenaza. Inmunitaria(o), es cuando se tienen los mecanismos y la posibilidad de que las diferentes células puedan colaborar para eliminar la amenaza que ha sido detectada y que disparó el estado de alerta máxima. Por tal razón es que en ocasiones el Sistema Inmune es llamado Sistema Inmunitario, para efectos de esta tesis se le llamará Sistema Inmune Natural (SIN).

Tomando en cuenta que la protección que brinda el Sistema Inmune Natural está muy enfocada a las enfermedades infecciosas, pero que también actúa contra enfermedades no infecciosas., una definición más completa de la inmunidad dada por Abbas (Abbas et al, 2002) es la siguiente:

“Reacción a sustancias extrañas, incluidos los microorganismos así como macromoléculas tales como proteínas y polisacáridos, cualesquiera que sean las repercusiones fisiológicas o patológicas de esta reacción”

3.2 Órganos y células del Sistema Inmune Natural

Como ya se ha mencionado en la sección anterior, la función del SIN es la defensa ante microorganismo dañinos. Para ello cuenta con un “ejército” de células que se clasifican por sus características y funciones.

Estas células tiene un proceso de generación, maduración y tiempo de vida sean activadas o no, además de que estas células llevan a cabo sus funciones y circulan a través del organismo gracias a los órganos y tejido linfoides.

A continuación se define la función de cada órgano que conforma el SIN.

3.2.1 Órganos del Sistema Inmune Natural

En los órganos del Sistema Inmune es donde se desarrollan, maduran, circulan las diferentes células que pertenecen al SIN. Es también donde sufren cambios en su estructura las células activadas para ser mejoradas por medio de la hipermutación somática, proceso que forma parte de lo que se conoce como Maduración de la Afinidad.

Los órganos se dividen por la función que realizan en órganos centrales y órganos periféricos.

CAPÍTULO III.- SISTEMA INMUNE NATURAL

A continuación se describe cada uno de ellos y su principal función.

Médula Ósea

En la médula ósea se generan los grupos celulares de la sangre como las células rojas, etc. así como los linfocitos B y T macrófagos, etc.

Este proceso inicia a partir de una **célula madre** común que se diferencia en elementos de estirpes celulares específicas que son las linfocítica, megacariocítica, eritroide, granulocítica y monocítica, que dan origen a los diferentes tipo celulares.

La generación, maduración y proliferación se lleva a cabo constantemente gracias a los estímulos recibidos de los linfocitos T activados así como de macrófagos.

También en la médula ósea se pueden encontrar células plasmáticas efectoras que llegan allí después de que son generadas por la activación de células B.

Timo

El timo es el órgano linfoide donde se maduran los linfocitos T provenientes de la médula ósea.

Sistema Linfático

El sistema linfático tiene una función muy importante en la defensa y eliminación de los agentes extraños dañinos o infecciosos.

El sistema linfático es el encargado de transportar a los antígenos desde el punto en que han ingresado al organismo hasta los ganglios linfáticos.

Ganglios Linfáticos

Los antígenos (específicamente proteicos) son llevados a los ganglios linfáticos ya que es allí donde se llevan a cabo inicialmente las respuestas inmunitarias adaptativas, es

CAPÍTULO III.- SISTEMA INMUNE NATURAL

decir en la que actúan principalmente los linfocitos. Para cada linfocito existe una zona para alojarse. La segregación anatómica es la que permite que las diferentes clases de linfocitos interactúen tanto con las células accesorias correspondientes como con otros linfocitos de la manera adecuada en el momento en que sea necesario. Esto se logra gracias a las señales que son enviadas a través de diversas sustancias químicas.

Bazo

El bazo lleva a cabo una función importantísima, “La limpieza de la sangre”, ya que este funciona como un filtro.

En él es donde se logran eliminar antígenos que están en la sangre pues se encuentran diferentes células agrupadas en ciertos lugares del bazo, estas son: la zona marginal, la pulpa blanca y la pulpa roja, donde se pueden encontrar macrófagos, células dendríticas, linfocitos, etc.

El bazo es donde se fagocitan los microorganismos recubiertos de anticuerpos.

Sistema Inmune cutáneo

Como es bien sabido, la piel es el órgano más grande con el que cuenta el ser humano y es a través de él que se interactúa con el medio ambiente. Por lo tanto, este tiene sus células específicas (linfocitos entre otras) y estas van a constituir la primera línea de defensa contra antígenos que puedan entrar al organismo por este medio. Algunas de las células que están presentes en este sistema son las células de Langerhans de la epidermis, Queratinocitos, entre otras (Abbas, 2002).

Sistema Inmune de las mucosas

CAPÍTULO III.- SISTEMA INMUNE NATURAL

Otro medio en el que pueden habitar y acceder al organismo los antígenos son las mucosas, tanto del tracto respiratorio como del gastrointestinal. En estos sitios se pueden encontrar una gran cantidad de células (linfocitos, etc.) que también constituyen una primera línea de defensa.

Órganos Centrales: Médula Ósea y Timo

Órganos Periféricos: Sistema Linfático, Ganglios Linfáticos, Bazo, Sistema Inmune cutáneo, Sistema Inmune de las mucosas.

Tabla 1 Clasificación de los Órganos Linfoides que constituyen al Sistema Inmune Natural y las funciones que desempeñan los mismos.

Órganos Linfoides	Nombre de los Órganos	Función
Centrales	Médula Ósea	Generación de células de las diferentes células y maduración de los linfocitos B
	Timo	Maduración de los linfocitos T
Periféricos	Sistema Linfático	transportar a los antígenos a los ganglios linfáticos
	Ganglios Linfáticos	Es donde se llevan las respuestas inmunitarias adaptativas
	Bazo	Filtrar la sangre, lugar donde se fagocitan microorganismos cubiertos de anticuerpos
	Sistema Inmune Cutáneo	Proteger de antígenos que pueden entrar por este medio
	Sistema Inmune de la Mucosas	Proteger de antígenos que pueden entrar por este medio

3.2.2 Células del Sistema Inmune Natural

Las células, moléculas y/o sustancias que constituyen al SIN se clasifican de acuerdo a sus características y funciones. Las células principales del sistema inmune son los linfocitos, las células accesorias y las células efectoras (Abbas, 2002).

Aunque todas interactúan para la eliminación del antígeno en cuestión, ciertas células van a pertenecer a lo que se conoce como la inmunidad innata y otras a la inmunidad adquirida o adaptativa. Estas se encuentran tanto en la sangre como en la linfa que forma parte del sistema linfático. (Abbas, 2002)

Dentro de cada tipo de inmunidad los elementos que actúan se clasifican, según su estructura, en **Humoral** o **Celular**.

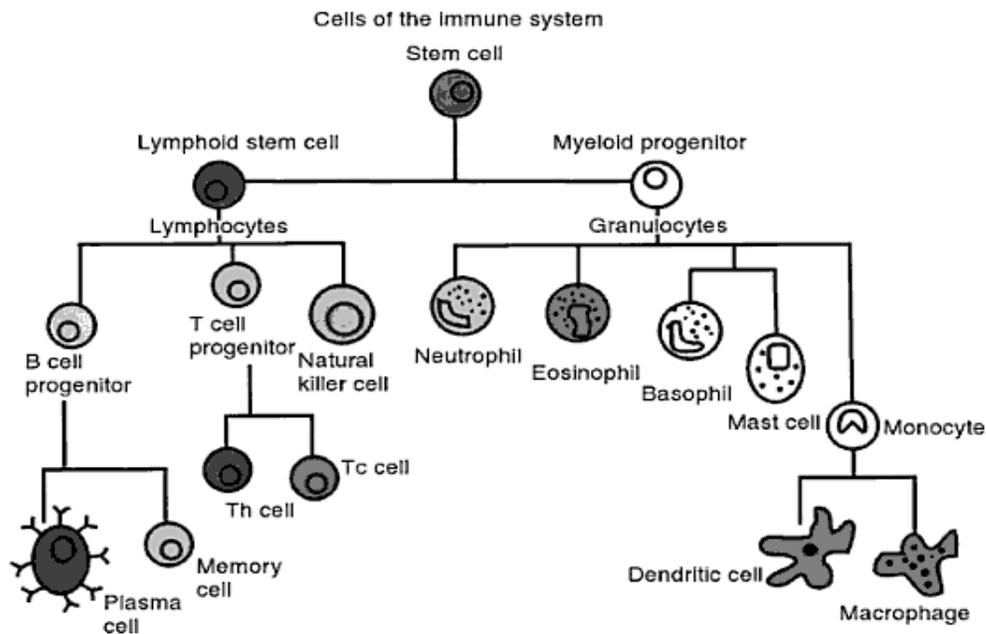


Figura 3 Clasificación de las células que conforman al SIN de acuerdo al tipo de inmunidad al que pertenecen.

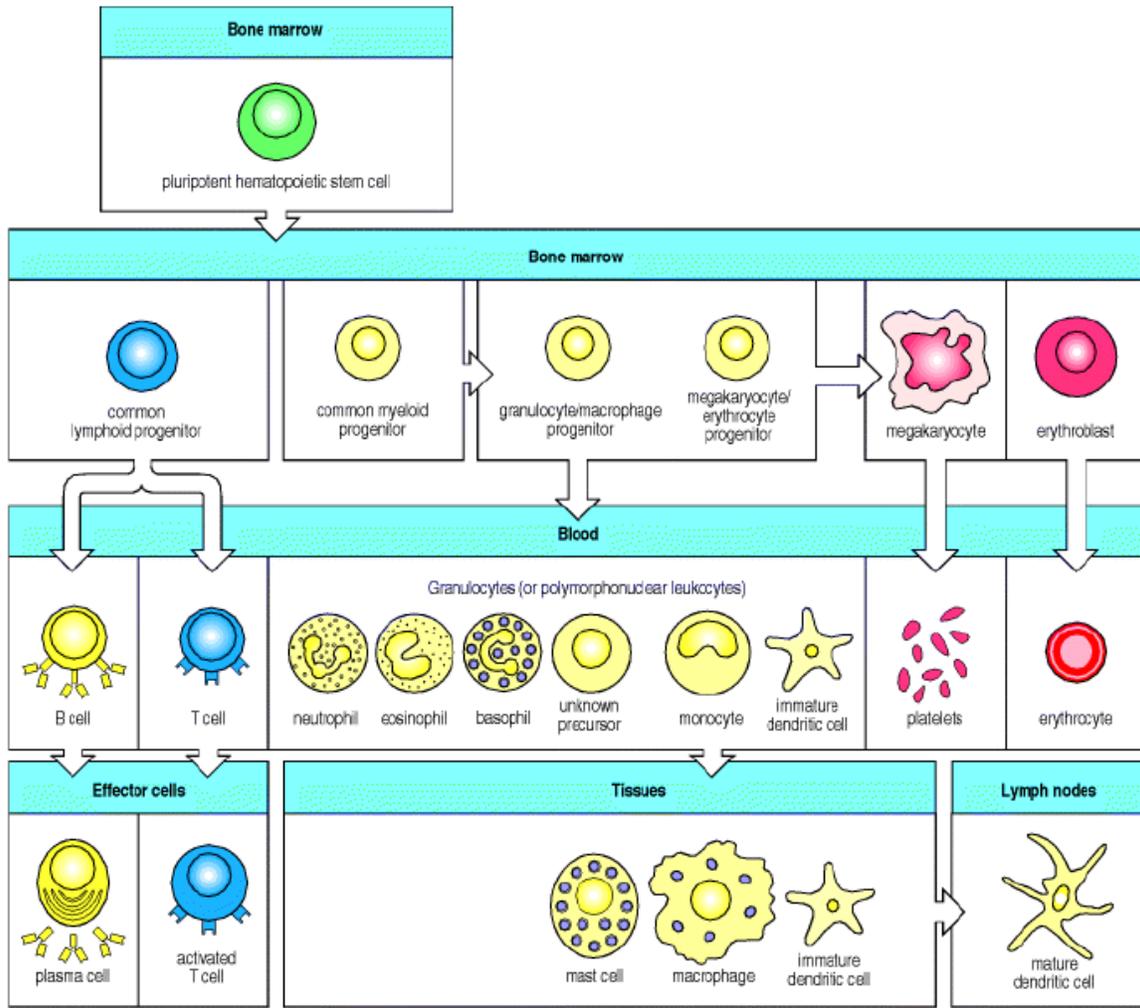


Figura 4 Clasificación de las células del SIN de acuerdo a la parte del organismo donde se pueden localizar. Tomada del libro Immunobiology (Janeway, 2001)

Inmunidad innata

La inmunidad innata es la que podríamos decir tenemos predefinida de fábrica y que tiene la capacidad de reconocer ciertas estructuras de antígenos sin haber tenido un encuentro previo. Anteriormente se consideraba que la inmunidad innata era inespecífica, sin embargo, con los estudios realizados se ha comprobado lo contrario (Akira, 2006)

CAPÍTULO III.- SISTEMA INMUNE NATURAL

Como se ha mencionado anteriormente la piel constituye una primera línea de defensa y funciona como una barrera contra los antígenos. Sustancias como el sudor, ácidos grasos, participan en la eliminación de ciertos antígenos. Otros elementos como la mucosa que se encuentra en los orificios naturales así como cilios del epitelio, el estornudo, etc. trabajan para la protección del organismo. Así mismo la lisozima contenida en saliva y lágrimas, el PH ácido del estomago, la fiebre, la flora intestinal, tensión del oxígeno pueden destruir cierto tipo de agentes dañinos. Todos estos elementos se clasifican como **Barreras físico-químicas**.

Las células que forman parte de la inmunidad innata son las siguientes:

Células fagocíticas inespecíficas: Existen diferentes clases de estas células que son los **macrófagos, neutrófilos, eosinófilos, basófilos, mastocitos**. Además de realizar la fagocitosis los macrófagos también se encargan de presentar el antígeno a las células T.

Células asesinas naturales (NK, del inglés): son una subpoblación de linfocitos T, sin embargo se diferencian principalmente en que no tienen receptores específicos de antígeno como el caso de los linfocitos T. Su función es la de ingerir microorganismos intracelulares así como reconocer y eliminar las células que tienen un comportamiento fuera de lo normal como las células tumorales, las células infectadas por virus (Janeway, 2001) así como detectar infecciones microbianas (Akira et al., 2006).

Otros elementos de la inmunidad innata que difieren de las células son:

Ácido láctico: puede eliminar ciertos antígenos por su efecto bacteriostático.

Interferón: Es una sustancia que es secretada por las células que han sido infectadas de inmediatamente después de la infección para que las células que son vecinas

CAPÍTULO III.- SISTEMA INMUNE NATURAL

produzcan proteínas antivirales. La función del interferón es minimizar la replicación viral.

Sistema Complemento: Está compuesto por proteínas que cuando entran en acción causan inflamación en la zona afectada. Consta de 2 vías

Vía clásica: es mediante la cuál se detectan anticuerpos que estén unidos a la superficie de ciertos microorganismos, esto se lleva a cabo mediante una proteína plasmática.

Vía alternativa: en este caso se reconocen estructuras o patrones de la superficie del microorganismo en cuestión.

Los elementos que forma parte de la clasificación Humoral o Celular se mencionan a continuación:

Humoral: Barrear físico-químicas, ácido láctico, interferón, Sistema Complemento.

Celular: Células fagocíticas inespecíficas, células NK.

Inmunidad adquirida o adaptativa

La inmunidad adaptativa (Roitt, 1997) es llamada así porque se obtiene después de la exposición a determinado antígeno y tiene características como la memoria que la hacen más efectiva a segundas exposiciones al mismo antígeno (maduración de la respuesta). Esta es puesta en marcha cuando la inmunidad innata no elimina al antígeno.

Las células que la forman son los linfocitos T y los linfocitos B también llamado glóbulos blancos. Cada una de estas células consta de receptores para detectar o reconocer al antígeno, TCR (T cell receptor), para los linfocitos T y BCR (B cell receptor) o anticuerpos, para los linfocitos B.

CAPÍTULO III.- SISTEMA INMUNE NATURAL

A continuación se describen las células que pertenecen a la inmunidad adaptativa así como las que participan para potenciar este tipo de inmunidad aunque en realidad nos son parte de ella.

Linfocitos B: son las células que pueden detectar/reconocer, por medio de los receptores (Jerne 1955; Burnet 1957), antígenos solubles o que están sobre la superficie de alguna célula del organismo invadido, pero que son extracelulares.

Anticuerpos: son las moléculas (proteínas) producidas por los linfocitos. Tienen la capacidad de reconocer al antígeno para el cuál son específicos, uniéndose a él para su eliminación. Su función principal es la de activar los mecanismos efectores según el tipo de antígeno. Algunos anticuerpos producen la eliminación del antígeno por medio de la fagocitosis y otros por la inflamación.

Los anticuerpos están constituidos por dos parátopes y 2 idiotopes como se muestra en la figura.

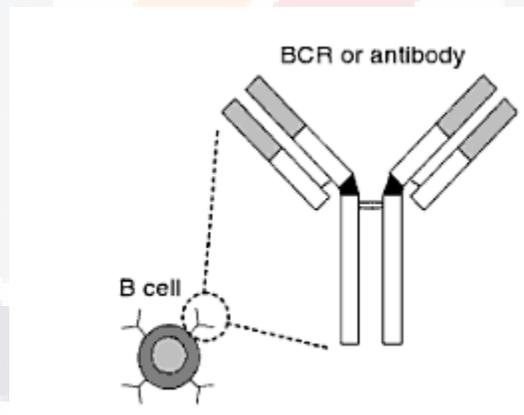


Figura 5 Representación de un linfocito B y su receptor (BCR del inglés B Cell Receptor) o anticuerpo.

Linfocitos T: son capaces de detectar los antígenos que se encuentran en el interior de las células del huésped, ej. Virus, bacterias, etc. Estos linfocitos tiene la capacidad de reconocer solamente antígenos peptídicos decodificados por genes del complejo

CAPÍTULO III.- SISTEMA INMUNE NATURAL

principal de histocompatibilidad (MHC, Major Histocompatibility Complex) que estén en la superficie celular a diferencia de los linfocitos B que pueden reconocer antígenos extracelulares (solubles).

Los linfocitos T se dividen en poblaciones de acuerdo a su función y a continuación se mencionan

Colaboradores (T-helper): entre sus funciones están el estimular a los linfocitos B a través de las citocinas que producen. Esta sustancia promueve la diferenciación de las células como los macrófagos, linfocitos B entre otras.

Supresores: son un subgrupo de los linfocitos colaboradores, su función es suprimir o inhibir las respuestas inmunitarias. La sustancia que producen es también la citocina.

Citocidas (CTL, Cytolytic T Lymphocyte): Su principal función es la de destruir células infectadas por virus o algún otro antígeno intracelular.

Asesinas Naturales (Natural Killer): pertenecen a la inmunidad innata.

Moléculas de Histocompatibilidad.

Estas moléculas así como otras células (linfocitos T y B) se encuentran en los seres vertebrados y en particular el Sistema Inmune de los humanos cuenta con ellas. Estas moléculas son codificadas por genes del locus llamado **Complejo Principal de Histocompatibilidad** (MHC, Major Histocompatibility Complex) (Abbas, 2002). Las moléculas se dividen en Clase I, II y III, esta última relacionada con factores de complemento (Litman, 1997).

CAPÍTULO III.- SISTEMA INMUNE NATURAL

Células de Memoria: Son células que se crean cuando un antígeno ha activado a un linfocito (es decir, se ha activado una respuesta inmunitaria). Tienen la capacidad para responder de manera más rápida y efectiva a una segunda exposición al mismo antígeno. Las células de memoria de los linfocitos B **cambian de isotipo** (IgG, IgE, IgA), en las de los linfocitos T aumenta el número de moléculas de adhesión; moléculas que fomentan que está células viajen hacia los lugares del cuerpo donde se presente una infección. Pueden sobrevivir aún si encontrase nuevamente con el antígenos durante años.

Células Efectoras: En algunos casos algunas son llamadas células plasmáticas. Las células efectoras son las que tienen la función de eliminar el antígeno que ha puesto a andar los mecanismo de defensa del SIN. Cuando estas células se han diferenciado de a partir de un linfocito B, producen anticuerpos para la eliminación del antígeno. Las que provienen del linfocito T colaborador producen citocinas para activa otras células y las provenientes de un linfocito T citolítico producen gránulos que dañan a las células tumorales y las infectadas por virus.

Células Accesorias: son células que no son específicas a un antígeno ya que no cuentan con receptores que funcionen como los de los linfocitos T y B. Estas células tienen la función de presentar al antígeno a los linfocitos T para que sea reconocido y enviar la segunda señal (generación de proteínas de membrana) para activar al linfocito en cuestión, así como potenciar la respuesta inmunitaria.

Células presentadoras: son células que presentan al antígeno a los linfocitos T, como las células dendríticas maduras y foliculares, fagocitos mononucleares. Las células dendríticas son las que tienen una mayor capacidad para activar las respuestas de las células T vírgenes a través de los ganglios linfáticos y el bazo. Esto hace que la

CAPÍTULO III.- SISTEMA INMUNE NATURAL

posibilidad de encontrarse con el antígeno sea muy alta y se de la respuesta inmunitaria adaptativa.

Los elementos que forma parte de la clasificación Humoral o Celular se mencionan a continuación:

Humoral: Anticuerpos

Celular: Linfocitos B y T.

Tabla 2 Clasificación de los elementos que conforman la inmunidad innata y la adaptativa de acuerdo a su comportamiento.

Inmunidad	Comportamiento	Nombre del elemento	Función
Innata	Humoral	Barreras físico-químicas	Interactúa con el medio ambiente (Ej. piel) y funciona como una primer línea de defensa contra los antígenos que intentan ingresar o que ya han ingresado al organismo.
	Celular	Macrófago	Fagocitar. También funciona como célula accesoria para presentar el antígeno a los linfocitos T.
		Célula Dendrítica	Funciona como célula accesoria para presentar el antígeno a los linfocitos T.
		Asesinas Naturales (NK)	Detectar y eliminar células tumorales, infectadas por virus, etc.
Adaptativa	Humoral	Anticuerpos	Son los receptores de los linfocitos B que se unen a antígenos para su

			eliminación.
	Celular	Linfocitos B	Son las células que se especializan en la detección de antígenos extracelulares. Cuando son activado se diferencian en células de memoria y efectoras (productoras de anticuerpos)
		Linfocitos T	

3.3 Características de las Respuestas Inmunitarias

En la inmunidad Innata como en la inmunidad Adaptativa se tienen diversos mecanismo tanto para el reconocimiento como para la eliminación (mecanismos efectores) del antígeno. En cuanto a estos últimos, en ambas inmunidades son similares. Sin embargo, la inmunidad adaptativa adiciona la característica de tener la expansión clonal de las células productoras de los anticuerpos específicos al antígeno invasor.

A pesar de que la inmunidad adaptativa tiene mecanismos muy evolucionados siempre estará estrechamente relacionada con la inmunidad innata para llevar a cabo su función. La regulación (supresión/estimulación) de las respuestas inmunitarias es un área de investigación de interés para la inmunología (Janeway, 2001).

Son muchas las características y propiedades con las que el SIN cuenta (Abbas, 2002) y que son relevantes para los investigadores del área de Inteligencia Artificial (Dasgupta 1998; Timmis et al., 2002a) por lo cual ha acaparado atención como una fuente de inspiración para el desarrollo de algoritmos.

Dichas características se describen a continuación:

Reconocimiento: El SIN puede reconocer entre lo propio y lo no propio, además puede reconocer cuando una célula propia es disfuncional, un elemento extraño no es dañino y cuando si lo es.

Extracción de características: El SIN tienen la capacidad de detectar características a nivel molecular para lograr detectar un patógeno a través de los receptores de los linfocitos, las células presentadoras y las moléculas como el Complejo Principal de Histocompatibilidad. Esto lo hace tomando los elementos moleculares que le interesan y eliminando el ruido o elementos no relevantes.

Especificidad: Cada respuesta inmunitaria que se activa está dirigida hacia un antígeno en específico o hacia un antígeno que forma parte de un grupo con el que comparte una determinada estructura.

Los linfocitos reconocen específicamente al antígeno a través de los epítomos, los que tienen la capacidad de detectar las diferencias, por pequeñas que sean, en la estructura de los diferentes patógenos.

Este concepto es esencial en la teoría de la selección clonal, ya que el linfocito que se expande es el que resulta específico al antígeno extraño dañino.

Diversidad: Los elementos y células del sistema Inmune Natural tienen la capacidad de reconocer una gran cantidad de diferentes antígenos. En particular, el repertorio de linfocitos es muy amplio dando lugar a una gran diversidad, ya que el sitio de unión con el antígeno es diferente en cada uno de ellos. Se considera que el Sistema Inmune Natural de los mamíferos puede reconocer entre 10^9 y 10^{11} determinantes antigénicos diferentes (Abbas, 2002).

Memoria: Esta es una característica propia de la inmunidad adaptativa pues cuando se enfrenta a un antígeno, en la fase de activación, se crean células de memoria específicas al patógeno. Gracias a estas células, las segundas respuestas inmunitarias

CAPÍTULO III.- SISTEMA INMUNE NATURAL

son más eficaces, rápidas y duraderas tanto en el reconocimiento del antígeno como en la eliminación del mismo, y son también diferentes en cuanto a la cantidad de células participantes. Por su estructura las células de memoria son más eficaces que los linfocitos vírgenes y pueden acudir más eficazmente a la zona donde está el daño.

Especialización: Esta característica se refiere a la forma en que pueden interactuar tanto de la inmunidad celular y/o humoral de las respuestas inmunitarias innata o adaptativa, que mejoran los mecanismos de defensa, los cuales van a estar determinados por el antígeno que ha invadido el organismo. Cada mecanismo que se desencadene en cada respuesta inmunitaria o en las diferentes fases estará especializado al microorganismo que afecte al huésped.

Detección distribuida: El SIN no tiene un punto central que controle todas las células y moléculas. Es naturalmente distribuido y cada célula tiene la poder de reconocer, responder y activar los mecanismos necesarios para eliminar al patógeno.

Autolimitación/Autoregulación: Las respuestas inmunitarias inician con una baja intensidad y esta va aumentando en el tiempo hasta llegar a su punto más alto. Después tienden a disminuir, ya que el estímulo del antígeno va desapareciendo pues los mecanismos efectores logran eliminar poco a poco al invasor dañino.

El proceso para llegar al estado de reposo es conocido como **homeostasis**, la cual se mantiene debido a que no se activa alguna respuesta inmunitaria o que disminuye el estímulo del antígeno después de que la respuesta se ha iniciado.

Tolerancia: Esta es una característica sumamente importante y que de alguno modo define la esencia del SIN. La auto-tolerancia es importante ya que de esta manera el sistema inmune no reacciona contra tejidos o sustancias que pertenecen al organismo huésped. Cuando hay un problema con la tolerancia sobrevienen enfermedades auto-inmunes. La auto-tolerancia es un aprendizaje que se lleva a cabo tanto en los órganos linfoides principales (tolerancia central) como en los órganos linfoides

CAPÍTULO III.- SISTEMA INMUNE NATURAL

periféricos (tolerancia periférica). En ambos casos existen mecanismos para eliminar las células que reaccionan contra los antígenos propios del huésped. Durante la maduración de los linfocitos se ponen en contacto con los antígenos propios de manera que presentan auto-reactividad son eliminados o inactivados.

Aprendizaje: Como se mencionó en la Tolerancia, el SIN aprende a través de diversos mecanismos. El mecanismo de la hipermutación somática, que viene después de una fuerte presión de selección, mejora los linfocitos de manera que la respuesta a determinado patógeno sea mejor. Este proceso es llamado maduración de la afinidad (Berek et al., 1993) el cual conduce a que las respuestas inmunitarias adaptativas sean cada vez más eficaces. Otro mecanismo de aprendizaje es el de la red inmune.

Metadinámica: La metadinámica es el proceso de la producción continua de nuevas células y moléculas, el reclutamiento, y la eliminación de aquellas que no son activadas, que son demasiado viejas o que no son de gran utilidad. (Varela et al., 1988)

Además el SIN interactúa fuertemente con otros sistemas como el Sistema Nervioso de una manera extraordinariamente compleja y que funciona muy bien, lo cuál es otra razón por la cual es de mucho interés para los investigadores.

Tabla 3 Características principales de las respuestas inmunitarias.

Características	Inmunidad Innata	Inmunidad Adaptativa
Especificidad	No	Si
Diversidad	Limitada	Si
Memoria	No	Si
Especialización	Limitada	Si
Autolimitación	Si	Si

CAPÍTULO III.- SISTEMA INMUNE NATURAL

Discriminación entre lo propio y lo no propio	Si	Si
Tolerancia	Si	Si
Aprendizaje	Si, central	Si
Reconocimiento	Si	Si
Extracción de características	Si	Si
Aprendizaje	Si	Si
Detección distribuida	Si	Si
Metadinámica	Si	Si

De acuerdo a las características descritas anteriormente, en esta tabla se especifica si están presentes o no en cada tipo de inmunidad. Se puede ver que dichas características se encuentran en la inmunidad adaptativa, ya que es la que cuenta con más mecanismos para llevar a cabo la eliminación del antígeno. Sin embargo la realidad es que la mayoría de ellas también se encuentran en la inmunidad innata sólo que de una forma aparentemente limitada o menos especializada, sin embargo son también de interés.

Los elementos que forman la inmunidad innata tienen una gran capacidad para reconocer lo propio y lo no propio. Sus receptores no tienen especificidad para un determinado antígeno sin embargo si pueden detectar antígenos que pertenecen a un grupo, es decir, que comparten ciertas características. Determinado tipo de receptores se encontrarán en células del mismo grupo.

La respuesta inmunitaria innata se puede activar antes de que el antígeno entre al organismo. Pero, cuando estos logran ingresar causando algún daño, la respuesta inmunitaria innata activa mecanismos que logran eliminar al invasor rápidamente, ya

sean humorales o celulares, pues no se incorpora el proceso de clonación. Cuando la inmunidad innata es evadida o superada por el antígeno, es entonces que se activa la inmunidad adaptativa.

3.4 Fases de las Respuestas Inmunitarias

Las fases en las que en general se divide la respuesta inmunitaria son 3: Fase de Reconocimiento, Fase de Activación, Fase de Eliminación / Efectora.

Estas van a variar, dependiendo del tipo de respuesta inmunitaria en cuanto a los mecanismos que se llevan a cabo y el tiempo que tardar en llegar a la eliminación del antígeno.

3.4.2 Fases de la Respuesta Inmunitaria Innata

Los antígenos o patógenos tienen diversas puertas de entrada al organismo ocasionando que se desencadenen diferentes mecanismos

La respuesta inmunitaria innata inmediata se lleva a cabo en un tiempo de 0 a 4 horas (Janeway, 2001) las fases son las siguientes:

Fase Reconocimiento y Activación: Después de haberse dado la alarma de que existe una infección, los elementos no específicos reconocen al antígeno y se activan para pasar a la siguiente fase.

Eliminación/Efectora: Los elementos que se activaron son capaces de eliminar al antígeno por ejemplo mediante el cambio de temperatura.

La respuesta que sigue a la innata es la temprana, queda entre la innata y la adaptativa. Esta toma un tiempo de entre 4 y 96 horas (Janeway, 2001). Las fases con las que cuenta

CAPÍTULO III.- SISTEMA INMUNE NATURAL

Reconocimiento y Activación: En este caso primero se hace el reclutamiento de las células que reconocerán al antígeno para luego ser activadas, todo esto en base al tipo de antígeno que causa la infección.

Efectora/Eliminación: Las células que son activadas llevan a cabo la eliminación del antígeno.

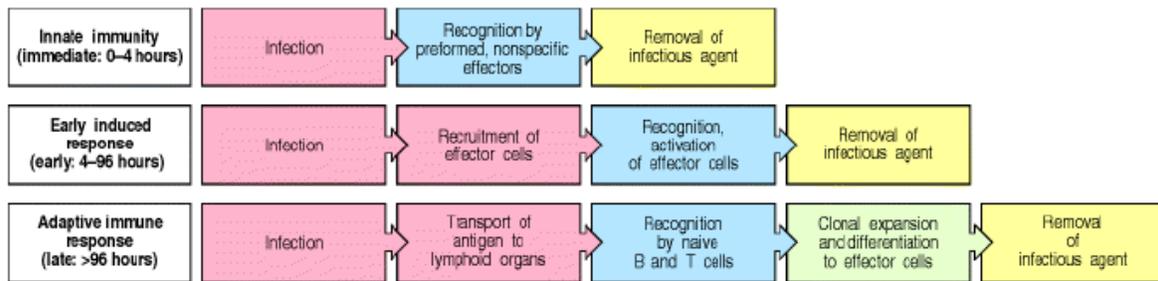


Figura 6 Duración y principales fases de las respuestas inmunitarias.

3.4.3 Fases de la Respuesta Inmunitaria Adaptativa

Es el tipo de respuesta que toma más tiempo ya que el principal mecanismo que utilizar para lograr la eliminación es la clonación.

La teoría de la selección clonal (Jerne, 1955; Burnet 1957) es la teoría central de la inmunidad adaptativa (Janeway, 2001), por medio de la cual se explica el funcionamiento de dicha inmunidad y la capacidad para reconocer antígenos de manera específica y la eliminación de aquellos que son autoreactivos.

Como ya se ha mencionado, la inmunidad adaptativa también se divide 3 fases:

Reconocimiento: Es la fase en la cuál el o los linfocitos afines al antígeno que invade al organismo, se unen a él para desencadenar los mecanismos necesarios para su

CAPÍTULO III.- SISTEMA INMUNE NATURAL

eliminación. Dependiendo del antígeno este va a ser reconocido por los linfocitos T o B específicos.

Activación: La presencia del antígeno y su reconocimiento se considera una primera señal y los componentes de las respuestas inmunitaria innatas cuando se enfrentan por ejemplo a microbios son una segunda señal, las cuales son necesarias para la activación de los linfocitos. A esto se le conoce como la **hipótesis de las 2 señales** (Abbas, 2002).

La primera señal se da cuando está presente un antígeno extraño que causa algún daño, garantizando así una respuesta específica. La segunda señal garantiza que la respuesta no sea auto-reactiva o contra sustancias o elementos no dañinos.

Otros procesos que se llevan a cabo en esta fase son: la síntesis de nuevas proteínas, la proliferación celular, la diferenciación en células efectoras y en células de memoria y finalmente la homeostasis. Todas ellas características de la inmunidad adaptativa.

Efectora / Eliminación: En esta fase la eliminación del antígeno se lleva a cabo por diversos mecanismos en los que participan tanto células y elementos de la inmunidad innata como de la adaptativa así como de células efectoras no linfoides. Los anticuerpos producidos por los linfocitos B recubren antígenos extracelulares facilitando así la fagocitosis por células como los macrófagos. Los linfocitos T eliminan antígenos intracelulares haciendo uso de las citocinas o citoquinas producidas por los linfocitos T citocidas.

En este punto se puede ver claramente la colaboración estrecha que hay entre los elementos y mecanismos de la inmunidad innata y adaptativa, ya que en esta fase la inmunidad adaptativa potencia los mecanismos efectores de la inmunidad innata dirigiéndolos a la zona afectada para eliminar al antígeno.

3.5 Respuestas inmunitarias primarias y secundarias

Cuando se activan respuestas inmunitarias adaptativas, al SIN le toma cierto tiempo el lograr eliminar al antígeno. Si es la primera vez que se enfrenta a un determinado patógeno la respuesta se le conoce como respuesta inmunitaria primaria (Tizard, 1988a) y a las subsecuentes respuestas al mismo patógeno se les conoce como respuestas inmunitarias secundarias.

En la respuesta primaria se reconoce al antígeno, se activa el linfocito que es más afín y se generan las células efectoras y de memoria, por lo cual toma más tiempo la eliminación del antígeno. Las respuestas secundarias suelen ser más rápidas y efectiva debido a que hay células de memoria y estas respuestas pueden ser más específicas y eficientes. Esto se ilustra en la figura

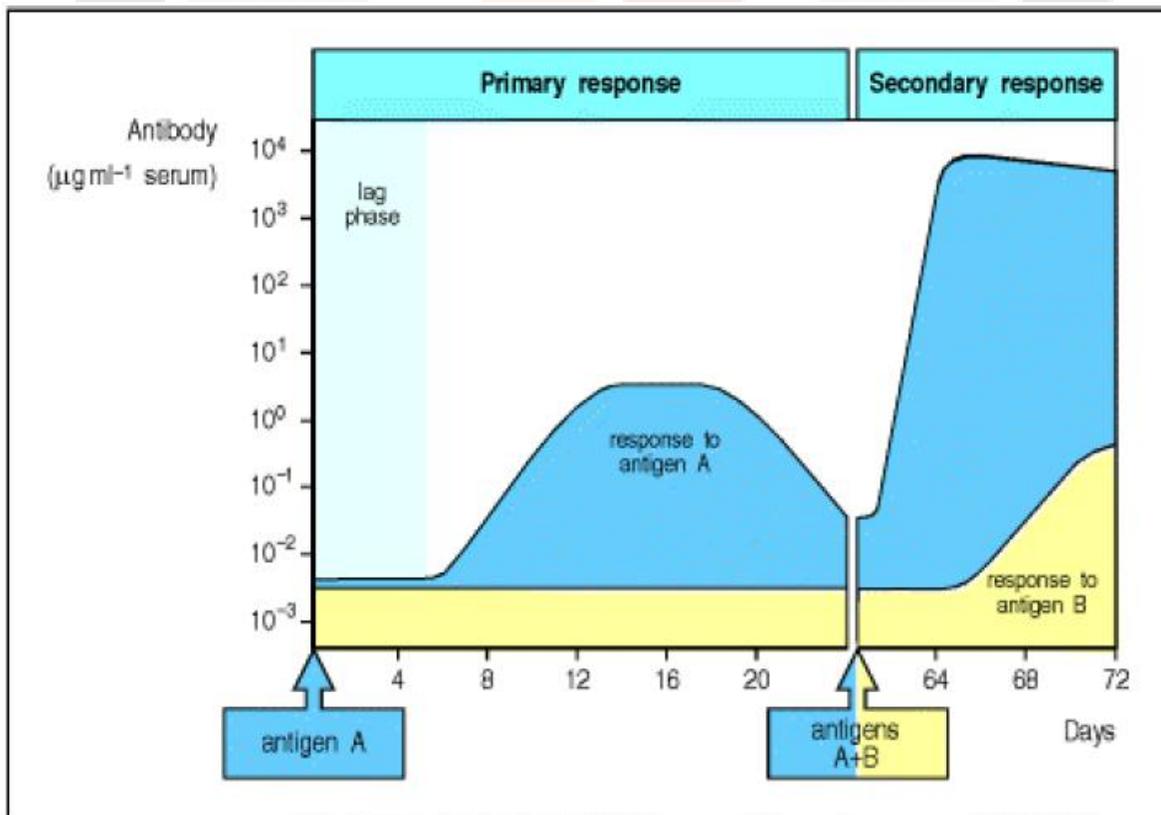


Figura 7 Grafica de la duración y efectividad de las respuestas inmunitarias primaria y de las respuestas inmunitarias secundarias.

3.6 Estructura de los linfocitos

Como se ha mencionado en secciones anteriores, los linfocitos son las células que tiene una mayor capacidad para reconocer y diferenciar de manera específica los determinantes antigénicos. Es a través de ellos y sus capacidades que el Sistema Inmune tiene memoria y especificidad.

Los linfocitos se clasifican según sus funciones y los productos proteicos sin embargo morfológicamente también se pueden diferenciar (Abbas, 2002). Los linfocitos reciben su nombre por el lugar donde maduran, como se mencionó, todas las células son generadas en la médula ósea, sin embargo los **linfocitos T** maduran en el Timo, por eso es la T. Los linfocitos B maduran en la bolsa de Fabricio en la aves, y aunque no hay un órgano equivalente en los mamíferos y la maduración es en la médula ósea, se les dejó ese nombre.

El repertorio de linfocitos es muy amplio debido a que durante su maduración los receptores se forman mediante la recombinación de segmentos de ADN. Después en el proceso de recombinación somática crea millones de genes codificadores de receptores diferentes de manera tal que el SIN tiene puede tener un gran repertorio diverso de linfocitos.

Cada población de linfocitos expresa una proteína diferente en la membrana para poder ser diferenciados unos de otros. Las subpoblaciones de linfocitos T expresan una proteína CD (Clustes of differentiation) que cambia en cada uno para poder ser diferenciado. Los linfocitos T colaboradores expresan la proteína CD4 mientras que los linfocitos T citocidas expresan la CD8.

Los linfocitos que no son estimulados se diferencian de los que si han sido estimulados y los de memoria en los marcadores de superficie. Los receptores además

de identificar a los antígenos, también participan en la supervivencia de los linfocitos especialmente en los que no han sido activados, que se cree que algunos autoligandos se unen a sus receptores para mantener su estancia en el SIN. Sin embargo, si no tienen un encuentro con el antígeno al que son específicos, los linfocitos suelen morir por un proceso de apoptosis.

3.6.1 Estadio de los linfocitos

La médula ósea constantemente está generando nuevos linfocitos, entre otras células, que en conjunto con el proceso de apoptosis logran mantener en estado de equilibrio la población de linfocitos maduros circulantes.

Es a través de las segundas señales que los linfocitos se activan y van a los órganos linfoides para proliferar y diferenciarse en células de memoria y efectoras. En el caso de las células T citocidas, las células efectoras producen gránulos que lisan las células infectadas por un virus o células tumorales.

En el caso de las células B, las células efectoras, que algunas se identifican como células plasmáticas, producen anticuerpos para eliminar al antígeno. Las células plasmáticas por lo general migran a la médula ósea y viven por un largo periodo de tiempo aún y cuando ya se haya eliminado el antígeno.

Las células de memoria son diferentes para cada linfocito sin embargo su función es la misma, potenciar las respuestas inmunitarias secundarias. Estas células tienen una morfología diferente y pueden diferenciarse de los linfocitos no estimulados y de los linfocitos efectoras. Los linfocitos B de memoria cambian de isotipo (IgG, IgE, IgA) a diferencia de los no estimulados que expresan IgM e IgD. Los linfocitos T de memoria expresan más moléculas de adhesión que las no estimuladas, moléculas que promueven la migración al punto en que el organismo sufra un infección.

3.7 La recirculación de los linfocitos y el asentamiento linfocitario

Como ya se ha mencionado en repetidas ocasiones, el SIN tiene la capacidad de reconocer antígenos extraños dañinos, los cuales pueden entrar al organismo por diferentes puntos. Para lograr detectar a estos antígenos el SIN realiza una especie de patrullaje viajando del torrente sanguíneo hacia el sistema linfático o a los órganos linfoides mecanismo conocido como recirculación linfocitaria (Abbas, 2002). El proceso mediante el cual se selecciona que población de linfocitos accede a ciertas zonas del organismo se conoce como asentamiento linfocitario (Abbas, 2002). El reclutamiento es el movimiento regulado de un tejido a otro.

Por tanto la recirculación, con asentamiento selectivo, logra por un lado que los linfocitos viajen a través del organismo para poder encontrarse con el antígeno al que son específicos independientemente del lugar por el que ingrese. Por otro lado, permite que los linfocitos puedan llegar las zonas donde se va a llevar a cabo una respuesta inmunitaria en la que hace falta su participación.

La recirculación linfocitaria sigue ciertos patrones que van a variar dependiendo de las moléculas de adhesión de los linfocitos. También va a ser diferente para linfocitos vírgenes, efectores o de memoria, determinando como se van a llevar a cabo las respuestas inmunitarias.

3.7.1 Las moléculas de adhesión

La adhesión de las moléculas va a determinar cuanto tiempo estarán retenidos los linfocitos en cierto tejido para luego regresar a los vasos linfáticos a través de la sangre. Por lo tanto de esto de la recirculación linfocitaria.

Las moléculas de adhesión conocidas como receptores de asentamiento, son las que determinan las vías de recirculación y asentamiento. Estos receptores de asentamiento son parte de las selectinas, integrinas y la super-familia de las Ig, la cuales son diferentes de los receptores para el antígeno.

El papel que juega el reconocimiento del antígeno en la recirculación es hacer más afín las integrinas de los linfocitos activados, permitiendo que se queden por más tiempo en el lugar donde se encontró el antígeno.

3.8 La recirculación de los linfocitos T vírgenes

Aún cuando no se tienen demasiados linfocitos T vírgenes, la forma en como se lleva a cabo la recirculación permite que estos muestren los órganos linfoides de todo el organismo. Los órganos donde principalmente se concentran los antígenos son en los ganglios linfáticos y el bazo y es ahí donde las células dendríticas los presentan a los linfocitos T vírgenes.

Por el torrente sanguíneo circulan los linfocitos T vírgenes que van teniendo contacto con las paredes vasculares así como con las células endoteliales, quedando unidos a ellas por un periodo muy corto (aprox. 1 seg.). Otros sitios como las vénulas endoteliales altas (HEV, por sus siglas en inglés) por donde circulan duran unidos por un periodo más largo. Al ir avanzando por dicha zona y hasta llegar al estroma del órgano linfoide periférico algunas células T este pueden aumentar su fuerza de unión, lo cuál puede ser propiciado por quimioquinas que se encuentran en su recorrido.

La selectina-L es el receptor de asentamiento que se une a las integrinas de las HEV. Durante el día circulan a través del bazo aproximadamente la mitad de los linfocitos debido a la selectividad.

Los linfocitos T cuando entran a un ganglio linfático y encuentran al antígeno al que son afines pueden durar un mayor tiempo en esa zona y en todo caso aumentar el

número de células que arriben ya que el flujo sanguíneo es mayor. Esto puede darse por efecto de una reacción inflamatoria.

Si el linfocito no encuentra el antígeno al cual es afín entonces desaparece la unión con la célula accesoria o presentadora y sale del ganglio linfático, volviendo a la circulación y asentándose en otro ganglio linfático.

En resumen cada linfocito puede visitar cada ganglio linfático 1 vez al día y aproximadamente pasan $25 \cdot 10^9$ células por los ganglios linfáticos. Mostrando así una recirculación muy activa.

3.9 Migración de los linfocitos T Efectores y de Memoria

Este tipo de linfocitos preferentemente viajan hacia los lugares donde se encuentra un foco de infección o hacia los tejidos periféricos inflamados, ya que indica la presencia de un patógeno y son necesarios para la eliminación del mismo. La circulación de los linfocitos de memoria es necesaria para que sea posible que se lleven a cabo las respuestas inmunitarias en la zona que se requiera independientemente de donde se hayan generado estas células.

El cambio en el proceso de circulación se modifica después de que se reconoce al antígeno.

Las células activadas tienen una mayor fuerza de unión con las células accesorias y para permanecer en el ganglio linfático. Esto es porque aumenta la afinidad de las integrinas y en las moléculas de adhesión. Esto es lo que principalmente diferencia a las células no activadas de las efectoras y de memoria.

El hecho de que los linfocitos de memoria migren a la lámina del intestino, al epitelio de la mucosa o a la dermis tiene que ver con las moléculas de adhesión responsables de este tipo de migración.

3.10 La recirculación de los linfocitos B

El patrón de recirculación de los linfocitos B es diferente de los linfocitos principalmente porque su permanencia en los órganos linfoides es más duradera. Dado que los anticuerpos pueden actuar aún y cuando no se encuentre el linfocito B activado, no es necesario que estas células se encuentren en los focos de inflamación para la defensa del organismo afectado.

3.11 Teoría y conceptos de la Inmunidad Adaptativa

La inmunidad Adaptativa consta de diferentes características así como células y moléculas que han servido como inspiración para el desarrollo de algoritmos.

Para llevar a cabo la función de eliminación del patógeno que afecta al organismo huésped fue enunciada la Teoría de la Selección Clonal, la cual fue propuesta primero por Niels Jerne en 1955 (Jerne, 1955) y claramente definida por Macfarlane Burnet en 1957 (Burnet, 1957).

Este es el principio fundamental de la inmunidad adaptativa (Janeway, 2001). Es mediante esta teoría que se explica el cómo se seleccionan los linfocitos (B o T) que son afines al antígeno extraño dañino y cómo estos proliferan, además de otros aspectos como la memoria y el aprendizaje.

Aún cuando la inmunidad innata puede reconocer y eliminar distintos patógenos, tiene ciertas limitaciones cuando el patógeno evoluciona más rápido que el mismo huésped. Mostrando así que la inmunidad innata puede reconocer patógenos que han conservado cierta estructura durante la evolución. Cuando un antígeno no puede ser reconocido por la inmunidad innata entonces puede hacerlo la inmunidad adaptativa a través de los linfocitos ya que de algún modo supera las limitaciones de la inmunidad innata.

Los postulados de la teoría de la selección clonal descritos en (Janeway, 2001) son:

1. Cada linfocito tiene un único receptor con una especificidad única.
2. La interacción entre una molécula del exterior y una capacidad de unión de linfocitos a esa molécula con una alta afinidad conduce a la activación de los linfocitos.
3. Las células efectoras diferenciadas derivadas de un linfocito activadas expresan receptores con especificidad idéntica a las del clon padre del cual provienen.
4. Los linfocitos que expresan receptores específicos que reconocen moléculas propias del huésped son borrados en las fases tempranas de desarrollo y maduración de esta manera estos desaparecen del repertorio de linfocitos.

Mediante estos postulados se puede explicar el proceso de la selección clonal, la cual se ejemplifica en la figura .

Este proceso inicia cuando algún linfocito del amplio repertorio existente en ese momento se activa por el encuentro con el antígeno al cual es específico dando lugar un número de procesos (Abbas, 2001) como los siguientes:

Proliferación celular: es también conocido como expansión clonal ya que los linfocitos activados se multiplican generándose clones con idéntica especificidad aumentando en número. En la presencia de cierto tipo de virus las células T pueden aumentar hasta 100000 veces.

Los linfocitos que reconocen a antígenos o células propias son eliminadas para que no proliferen y no formen parte del repertorio linfocitario.

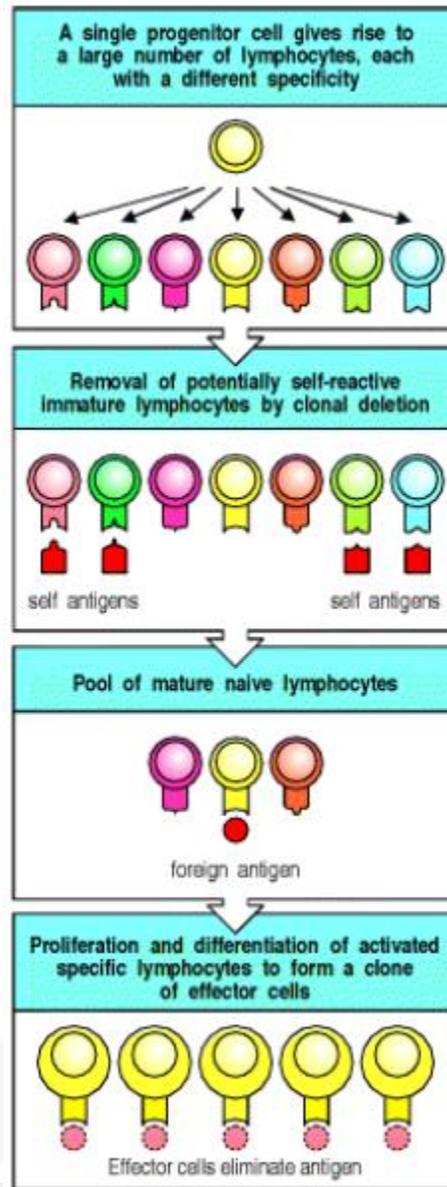


Figura 8 Ejemplificación de la teoría de la Selección Clonal. (Janeway, 2001)

Diferenciación en células efectoras: parte de los linfocitos que han proliferado se diferencian en linfocitos efectores que como ya se mencionado tienen la función de eliminar al antígeno.

Las sustancias producidas por los linfocitos T colaboradores son las citocinas para activar otras células y los linfocitos T citocidas producen sustancias para eliminar al antígeno. Los linfocitos B efectores secretan activamente anticuerpos.

Diferenciación en células de memoria: Otra parte de las células diferenciadas se convierten en células de memoria. Eso sucede tanto con los linfocitos T como en los B. Esta es la principal característica de la inmunidad adaptativa. Gracias a estas células el organismo permanece protegido por periodos largos de tiempo. Las células de memoria de los linfocitos B través del organismo, torrente sanguíneo, órganos linfoides probablemente sin producir anticuerpos (Perelson et al., 1978)

Homeostasis: Es un estado en el que el organismo vuelve a su equilibrio después de haber eliminado al patógeno que ha causado una infección o daño. Este estado se alcanza al morir la mayor parte de las células que fueron activadas por medio del proceso llamado apoptosis. Este es un proceso de muerte celular programado el cual no activa ningún tipo de respuesta inmunitaria. Se considera que esta muerte se da porque ya no existe el estímulo del antígeno así como la no presencia de ciertas proteínas que inducen el crecimiento y proliferación.

Otros conceptos importantes son que se mencionan como parte de la hipótesis de la selección son los siguientes:

Afinidad: Este es un concepto que se utiliza cuando se refiere al reconocimiento del antígeno, ya que el linfocito más afín al es el que se activa. Cuanto más fuerte es la interacción Anticuerpo- Antígeno más alta es la afinidad.

Cuando es activado un linfocito B, se da un proceso de mejora para aumentar la afinidad. Esto se logra modificando la estructura de las regiones V de los anticuerpos durante la respuesta inmunitaria adaptativa humoral. Estos cambios logran que el anticuerpo se una con mayor afinidad al antígeno que el linfocito progenitor. Por lo tanto estos nuevos linfocitos serán más estimulados en segundas respuestas inmunitarias.

CAPÍTULO III.- SISTEMA INMUNE NATURAL

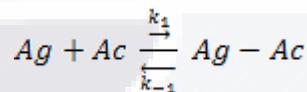
Los nuevos clones de los linfocitos B activados son copias exactas que pasan por el proceso de hipermutación somática (Berek et al., 1993). Este es el proceso encargado de modificar la afinidad del linfocito clonado.

Otros autores definen la afinidad como:

“La energía de enlace potencial contenida en cada uno de los sitios de combinación del anticuerpo y en ella participan todas las posibles interacciones con el antígeno. Estas, a su vez, están determinadas por el número y el tipo de grupos químicos reactivos presentes en el sitio de combinación” (Rojas, 2001)

“La afinidad de anticuerpo es una medida cuantitativa de la fuerza de unión. Es la fuerza combinada de las interacciones no covalentes entre un sitio de unión de antígeno único en un anticuerpo y un epítipo único. Los anticuerpos de afinidad baja se unen al antígeno de manera débil y tienden a disociarse con facilidad, en tanto que los de afinidad alta unen con antígeno con más fuerza y la unión se mantiene por más tiempo” (Kindt et al, 2007)

La relación entre el sitio de unión en un anticuerpo (Ac) y un antígeno (Ag) se describe como sigue (Kindt et al, 2007)



Ecuación 1 Representación de la relación entre un antígeno y un anticuerpo que está relacionado con la afinidad.

En ocasiones el anticuerpo puede reaccionar de forma cruzada con un antígeno que no está relacionado. Es **reactividad cruzada** sucede si dos antígenos diferentes comparten un epítipo idéntico o muy similar. Siendo así la afinidad del anticuerpo por el epítipo de reacción cruzada suele ser menor que para el epítipo original. (Kindt et al, 2007)

CAPÍTULO III.- SISTEMA INMUNE NATURAL

PUNTOS IMPORTANTES

La organización anatómica de las células y los tejidos del sistema inmunitario es de importancia crucial para la generación de las respuestas inmunitarias. Esta organización permite al reducido número de linfocitos específicos de cada antígeno localizar y responder a su antígeno con independencia del lugar por el que éste penetre en el organismo.

Las células accesorias tienen la función de promover la activación de los linfocitos y de realizar funciones efectoras que pueden potenciarse por repuestas inmunitarias adaptativas humorales o celulares. Son células accesorias los fagocitos mononucleares, las células dendríticas y las células dendríticas y las células dendríticas foliculares.

La recirculación linfocitaria es el proceso por el cual los linfocitos se desplazan continuamente entre distintas localizaciones del cuerpo a través de la sangre y los vasos linfáticos, y es crucial para las fases de iniciación y efectora de las respuestas inmunitarias.

Las células T vírgenes recirculan en condiciones normales entre los diversos normales entre los diversos órganos linfoides periféricos, lo que aumenta la probabilidad de que se encuentren con el antígeno presentado por las células presentadoras de antígenos tales como las células dendríticas maduras. Las células T de memoria y efectoras suelen ser reclutadas en focos de inflamación periféricos, donde se localizan los agentes microbianos.

El proceso de recirculación linfocitaria está gobernado por moléculas de adhesión, que son responsables de las interacciones entre los linfocitos y las células del endotelio vascular, y el patrón de recirculación difiere entre las distintas subpoblaciones de linfocitos. Las moléculas de adhesión expresada en los linfocitos que median la recirculación se denominan receptores de asentamiento. Las células

CAPÍTULO III.- SISTEMA INMUNE NATURAL

endoteliales de los diferentes tejidos pueden expresar distintos ligandos para estos receptores de asentamiento, denominados diriginas, que promueven el asentamiento linfocitario específico de tejidos.

El asentamiento selectivo de las células T vírgenes en los ganglios linfáticos está mediado en gran medida por la unión de la selectina-L presente en las células T a la dirigina de los ganglios linfáticos periféricos presente en la vénulas endoteliales altas (HEV) de los ganglios linfáticos. Las células T efectora y de memoria que son generadas por estimulación antigénica de las células T vírgenes se caracterizan por una menor expresión de selectina-L, pero una mayor expresión de intreginas y de los ligandos de la selectina-E y la selectina-P, y estas moléculas median la unión al endotelio de las células T en los focos de inflamación periféricos.

ANTICUERPO Y ANTÍGENOS

Los anticuerpos, las moléculas del complejo de histocompatibilidad (MHC, major histocompatibility complex) y los receptores para el antígeno de la célula T son las tres clases de las moléculas que utiliza la inmunidad adaptativa para el reconocimiento de antígenos. Los anticuerpos se unen a una gama más amplia de estructuras antigénicas, muestran una mayor capacidad para diferenciar los distintos antígenos y se unen a los antígenos con mayor fuerza.

Los anticuerpos se unen específicamente al antígeno en las fases de reconocimiento y efectora de la inmunidad humoral.

DISTRIBUCIÓN NATURAL Y PRODUCCIÓN DE ANTICUERPOS

Los anticuerpos se encuentran distribuidos en los líquidos biológicos de todo el organismo, y están presentes en la superficie de un número reducido de tipos celulares. En el interior de las células B, los anticuerpos se encuentran en los

compartimentos unidos a la membrana citoplasmática y en la superficie, donde se expresan como proteínas integrales de membrana.

Las concentraciones de moléculas de anticuerpo en el suero específicas de un antígeno determinado se suele calcular determinando el número de diluciones seriadas del suero que es posible hacer hasta que deje de observarse la unión al antígeno; los suero con una elevada concentración de moléculas de anticuerpo específicas de un determinado antígenos se dice que son “fuertes” o que tienen un “título elevado”.

Un ser humano adulto sano de 70kg produce aproximadamente 3g de anticuerpos al día. Casi las dos terceras partes corresponden a un anticuerpo denominado inmunoglobulina A (IgA).

Los anticuerpos que entran en la circulación tienen una semivida limitada, que para el tipo de anticuerpo más común del suero, la inmunoglobulina G (IgG), se sitúa en torno a las 3 semanas.

El suero contiene mezclas de distintos anticuerpos producidos por muchos clones linfocitos B (denominados anticuerpos policlonales). “En 1975 Georges Köhler y Cesar Milstein idearon un método para inmortalizar células secretoras de anticuerpos individuales”

3.11 Características generales de la estructura de los anticuerpos

Elvin Kabat y colaboradores demostraron que la mayor parte de los anticuerpos se encuentran en el tercer grupo de las globulinas de migración más rápida, llamadas **gammaglobulinas** por la tercera letra del alfabeto griego. Otro nombre común para los anticuerpos es inmonuglobulina (Ig), que hace referencia a la parte de la fracción de las gammaglobulinas que confieren inmunidad. Los términos **inmunoglobulina** y **anticuerpo** se usan indistintamente.

CAPÍTULO III.- SISTEMA INMUNE NATURAL

Todas las moléculas de anticuerpos comparten las mismas características estructurales básicas, pero presentan una notable variabilidad en las regiones de unión al antígeno. Esta variabilidad justifica la capacidad de los diferentes anticuerpos de unirse a un asombroso número de antígenos estructuralmente distintos. Existen más de 10^7 (quizá hasta 10^9) moléculas de anticuerpos diferentes en cada individuo, cada una con secuencias de aminoácidos únicas en sus sitios de unión al antígeno.

Todos los anticuerpos tienen una estructura básica simétrica compuesta de dos cadenas ligeras idénticas (cada una de aproximadamente 24 kilodalton [kD]) y dos cadenas pesadas idénticas (de unos 55 o 70 kD).

Las cadenas pesadas y ligeras de las Ig se componen de regiones variables (V) amino terminales (N-Terminales) o de regiones constantes (C) carboxi terminales.

La región V de la cadena pesada consta de un dominio de Ig, y la región C de la cadena pesada consta de tres o cuatro dominios de Ig. La región V de la cadena ligera tiene un dominio de Ig, y la región C de la cadena ligera también tiene un dominio de Ig. Las regiones variables reciben este nombre debido a que contienen regiones de variabilidad en su secuencia de aminoácidos que diferencian los anticuerpos sintetizados por un clon de células B de los anticuerpos sintetizados por otro clon. La región V de una cadena pesada (V_H) se yuxtapone con la región V de una cadena ligera (V_L) para formar un sitio de unión al antígeno.

Las regiones C de la cadena pesada interactúan con otras moléculas y células efectoras del sistema inmunitario y por lo tanto median la mayoría de las funciones biológicas de los anticuerpos. Por consiguiente, el reconocimiento antigénico y las funciones efectoras de las moléculas de anticuerpos están separadas espacialmente en la región V y la región C, respectivamente. Las regiones C de las cadenas ligeras no participan en las funciones efectoras y no se fijan a las membranas celulares.

Características estructurales de las regiones variables y su relación con la unión al antígeno

La mayoría de las diferencias de secuencia en los distintos anticuerpos está confinada a tres cortos segmentos localizados en las regiones V de las cadenas pesadas y ligeras, que reciben el nombre de segmentos hipervariables. Cada uno de estos tres tramos sumamente divergentes dentro de cada región V tiene una longitud aproximada de unos 10 aminoácidos, y se mantiene en posición mediante regiones marco más conservadas del dominio de Ig de la región V. Los mecanismos genéticos que conducen a la variabilidad de los aminoácidos se analizan en el capítulo 7. Las tres regiones hipervariables de un dominio V_L y las tres regiones hipervariables de un dominio V_H .

Las regiones hipervariables también se denominan regiones determinantes de la complementariedad (CDR). Las regiones CDR3 de los segmentos tanto V_L como V_H son las regiones CDR más variables y, existen mecanismos genéticos especiales responsables de la generación de una mayor diversidad de secuencia en la región CDR3 que en las regiones CDR1 y CDR2.

Las diferencias de secuencia entre las regiones CDR de las diferentes moléculas de anticuerpo originan estructuras químicas únicas que se muestran en las superficies de los bucles proyectados, y son responsables por lo tanto de las diferentes especificidades para los antígenos. Al quedar confinada la variabilidad de secuencia a tres cortos tramos, la estructura básica de todos los anticuerpos se conserva a pesar de la variabilidad existente entre los distintos anticuerpos.

La unión del anticuerpo de las moléculas de anticuerpo al antígeno es una función principalmente de las regiones hipervariables de V_H y V_L . El contacto más extenso corresponde a la tercera región hipervariable (CDR3), la cual, como se comentó anteriormente, es también la región variable de las tres.

CAPÍTULO III.- SISTEMA INMUNE NATURAL

Características estructurales de las regiones constantes y su relación con las funciones efectoras.

Las distintas clases de moléculas de anticuerpos también reciben el nombre de isotipos, y se denominan IgA, IgD, IgE, IgG e IgM. En el ser humano, los isotipos IgA e IgG pueden subdividirse en subclases estrechamente relacionadas, o subtipos, llamados IgA1 e IgA2, e IgG1, IgG2, IgG3 e IgG4 respectivamente. Las regiones C de las cadenas pesadas de todos los anticuerpos de un isotipo o subtipo comparten esencialmente la misma secuencia de aminoácidos. Las cadenas pesadas se designan por las letras del alfabeto griego correspondiente al isotipo general del anticuerpo: la IgA1 contiene cadenas pesadas α 1; la IgA2, α 2; la IgD, δ ; la IgE ϵ ; la IgG1, γ 1; IgG2, γ 2; IgG3, γ 3; IgG4, γ 4; IgM, λ . Los anticuerpos pueden actuar como antígenos cuando se introducen en huéspedes extraños, desencadenando la producción de anticuerpos. Por consiguiente es posible obtener anti-anticuerpos específicos de una clase o subclase de Ig.

Diferentes isotipos y subtipos de anticuerpos realizan distintas funciones efectoras.

El motivo es que la mayoría de las funciones efectoras de los anticuerpos está mediada por la unión de las regiones C de las cadenas pesadas a distintos receptores de la superficie celular y macromoléculas tales como las proteínas del complemento. Los isotipos y tanto y subtipos de anticuerpos difieren en su regiones C y por lo tanto en la funciones efectoras mediadas por cada isotipo de anticuerpo.

Las moléculas de antígeno son flexibles, lo que les permite unirse a distintas series de antígenos multivalentes. Cada anticuerpo posee al menos dos sitios de unión al antígeno, cada uno formado por un par de dominios V_H y V_L . Muchas

moléculas de Ig pueden orientar estos sitios de unión forma que se pueden unir al mismo tiempo a dos antígenos situados sobre una superficie plana. Esta flexibilidad depende en gran medida de la **región bisagra** situada entre C_{H1} y C_{H2} en ciertos isotipos. Le región bisagra puede contener desde unos 10 aminoácidos hasta más de 60 en los distintos isotipos. Partes de estas secuencias adoptan una conformación **aleatoria** y flexible, lo que les proporciona libertad de movimiento molecular entre C_{H1} y C_{H2} . Algunas de las diferencias principales de las subclases de IgG se concentran en la bisagra. Esto origina formas globales muy diferentes de los distintos subtipos de IgG. Además, parte de la flexibilidad de las moléculas de anticuerpo se debe a la capacidad de cada dominio V_H de rotar con respecto al dominio C_{H1} .

Dos clases o isotipos de cadenas ligeras, denominadas k y λ , se distinguen por sus por sus regiones C carboxiterminales. Toda molécula de anticuerpo tiene dos clases ligeras λ , pero nunca una de cada una. Las regiones C de las cadenas ligeras k en todos los anticuerpos que contienen cadenas k tienen una secuencia de aminoácidos idéntica y son distintas de las secuencias de la región C de las cadenas ligeras λ , que son idénticas en todos los anticuerpos que contienen cadenas λ .

Las regiones C_k y C_λ se pliegan en un único dominio de Ig. En el ser humano, aproximadamente el 60% de las moléculas de anticuerpo tienen cadenas ligeras k , y alrededor del 40% tienen cadenas ligeras λ . Se producen cambios notables de esta proporción cuando tiene lugar una expansión monoclonal de células B neoplásicas, produce moléculas de anticuerpo con la misma cadena ligera. A diferencia de los isotipos de las cadenas pesadas, no existen diferencias de función conocidas entre los anticuerpos que contienen cadenas k y los que contienen cadenas λ .

Unión de los Anticuerpos a los Antígenos

Características de los antígenos

CAPÍTULO III.- SISTEMA INMUNE NATURAL

Se puede definir un **antígeno** como cualquier sustancia a la que se unen específicamente una molécula de anticuerpo o un receptor de la célula T. Los anticuerpos tienen la capacidad de reconocer como antígeno prácticamente cualquier clase de molécula biológica, entre otras los metabolitos intermediarios los azúcares, los lípidos, los autacoides y las hormonas, macromoléculas del tipo de los hidratos de carbono complejos, los fosfolípidos, los ácidos nucleicos y las proteínas. Esta característica los diferencia de las células T, que sólo reconocen péptidos. Sólo las macromoléculas pueden estimular a los linfocitos B para iniciar la respuesta inmunitaria humoral. Las moléculas que estimulan la generación de respuestas inmunitarias se llaman **inmunógenos**. Los agentes químicos pequeños, como el dinitrofenol, no pueden activar por sí mismos a las células B (es decir, no son sustancias inmunogénicas). En este sistema, el agente químico de pequeño tamaño se llama **hapteno**, y la macromolécula recibe el nombre de **transportador**. El complejo hapteno-transportador, puede actuar como inmunógeno.

Por lo tanto, un anticuerpo se une sólo a una parte específica de la macromolécula, denominada **determinante** o **epítipo**. Las macromoléculas contienen típicamente determinantes múltiples, algunos de los cuales pueden repetirse, y cada uno de los cuales, por definición, puede unirse a un anticuerpo. La presencia de múltiples determinantes idénticos en un antígeno se conoce como **polivalencia** o **multivalencia**.

Los epítipos formados por varios residuos de aminoácidos adyacentes se llaman **determinantes lineales**.

Los **determinantes conformacionales** están formados por residuos de aminoácidos que nos están en secuencia, pero que quedan yuxtapuestos en el espacio cuando la proteína está plegada.

Las proteínas pueden ser objetos de modificaciones tales como fosforilación o proteólisis. Estas modificaciones, al alterar la estructura covalente, pueden dar origen a nuevos epítipos. Estos epítipos reciben el nombre de **determinantes neoantigénicos** y también pueden ser reconocidos por anticuerpos específicos.

Bases estructurales y químicas de la unión al antígeno

Los sitios de unión al antígeno de la mayoría de los anticuerpos son superficies planas que aceptan epítomos conformacionales de macromoléculas. Las moléculas del MHC (moléculas del complejo principal de histocompatibilidad), que contienen hendiduras de unión al antígeno que unen pequeños péptidos, pero no proteínas globulares nativas.

La fuerza de enlace entre un único sitio de combinación de un anticuerpo y un epítomo de un antígeno, que puede determinarse experimentalmente mediante diálisis de equilibrio, se denomina **afinidad** del anticuerpo. La afinidad se suele representar mediante una constante de disociación (K_d), que describe la concentración de antígeno necesaria para ocupar los sitios de combinación de la mitad de las moléculas de anticuerpo presentes en una solución de anticuerpo. Cuando menor es K_d , mayor es la afinidad de la interacción, puesto que se necesita una menor concentración de antígenos para ocupar los sitios de unión. La K_d de los anticuerpos específicos de los antígenos naturales suele oscilar entre $10^{-7}M$ y $10^{11}M$. En el suero de un individuo inmunizado hay una mezcla de anticuerpos con distintas afinidades por el antígeno, dependiendo principalmente de la secuencia de aminoácidos de las CDR.

Debido a que la región bisagra de los anticuerpos les proporciona flexibilidad, un único anticuerpo puede unirse a un único antígeno multivalente a través de más de un sitio de unión. Para las IgG o IgE, esta unión puede realizarse como máximo por dos sitios de unión, uno en cada Fab, pero para la IgM pentamérica, un solo anticuerpo puede unirse hasta a diez sitios diferentes. Los antígenos polivalentes tienen más de una copia de un determinante específico. Aunque la afinidad de cualquier sitios de unión polivalente, la fuerza de la unión del anticuerpo al antígeno debe tener en cuenta la unión de todos los sitios a todos los epítomos disponibles. Esta fuerza total de la unión recibe el nombre de **avidez**, y es mucho mayor que la afinidad de cualquier sitio de unión al antígeno. Matemáticamente, la avidez aumenta casi

geométricamente (en lugar de aritméticamente) para cada lugar ocupado. De esta manera, una molécula de IgM con baja afinidad puede unirse fuertemente a un antígeno polivalente, ya que muchas interacciones de baja afinidad producen una única interacción de gran afección.

Si un antígeno polivalente se mezcla con un anticuerpo específico en un tubo de ensayo, ambos se unirán formando **inmunocomplejos**. En las concentraciones correctas (lo que recibe el nombre de “zona de equivalencia”), antígeno y anticuerpo forman una amplia red de moléculas unidas de forma no covalente, de tal manera que la mayoría o todas las moléculas de antígeno y anticuerpo quedan unidas en grandes masas. Los inmunocomplejos que quedan atrapados o se forman en los tejidos pueden iniciar una reacción inflamatoria que provoque una “enfermedad por inmunocomplejos”.

Complejo Principal de Histocompatibilidad

Las funciones más importantes de los linfocitos T son la defensa frente a organismos intracelulares y la activación de otras células tales como los macrófagos y los linfocitos B. La especificidad de los linfocitos T contrasta con la de los linfocitos B y sus productos de secreción, los anticuerpos, que pueden reconocer antígenos solubles y antígenos asociados a células. La tarea de mostrar los antígenos de los microorganismos asociados a células para su reconocimiento para las células T es realizada por proteínas especializadas codificadas por genes situados en un locus denominado **complejo principal de histocompatibilidad** (MHC, Major histocompatibility complex).

La función fisiológica de las moléculas del MHC es parte integral de los ligandos que reconocen la mayoría de las células T, ya que los receptores de antígeno de las células T realmente son específicos para los complejos de antígenos peptídicos extraños y moléculas de MHC. Hay 2 tipos diferentes de productos de los genes del MHC, las **moléculas de clase I** y las **moléculas de clase II**, que “**muestran**” conjuntos diferentes de antígenos proteicos (antígenos extracelulares que han sufrido

CAPÍTULO III.- SISTEMA INMUNE NATURAL

endocitosis y antígenos intracelulares citosólicos) y presentan péptidos a los diferentes subgrupos de células T, los linfocitos T colaboradores CD4⁺ y los linfocitos T citolíticos CD8⁺, respectivamente.

El polimorfismo genético, en otro tipo de genes existen formas alternativas, o variantes, con frecuencias estables en diferentes miembros de la población. A este tipo de genes se les denomina **polimórficos**, y cada variante común de un gen polimórfico se llama **alelo**. En el caso de los genes polimórficos, cada individuo puede tener el mismo alelo en ese locus genético en ambos cromosomas del par, denominándose homocigótico, o dos alelos diferentes, uno en cada cromosoma, denominándose heterocigótico.

Cuando se injerta un tejido u órgano pueden ocurrir dos cosas, la piel injertada sobrevive y funciona normalmente; o el sistema inmunitario destruye el injerto, proceso denominado **rechazo**.

Los genes responsables de que se considere a un tejido injertado como similar o diferente a los tejidos propios se llaman **genes de histocompatibilidad**, y las diferencias entre lo extraño y lo propio se atribuyeron al polimorfismo genético entre los diferentes alelos de los genes de histocompatibilidad.

El locus identificado por el grupo de Snell estaba ligado a un gen ubicado en el cromosoma 17 que codifica un antígeno polimórfico de grupo sanguíneo llamado Antígeno II, motivo por el cual recibió el nombre de región de histocompatibilidad-2 o simplemente, H-2.

La región genética que controla el rechazo de injertos y que contenía varios genes ligados recibió el nombre de complejo principal de histocompatibilidad o MHC.

Otros genes que contribuyeron en menor medida al rechazo de injertos reciben el nombre de genes del sistema menor de histocompatibilidad.

CAPÍTULO III.- SISTEMA INMUNE NATURAL

En los decenios de 1960 y 1970 se descubrió que los genes del MHC son esenciales para todas las respuestas inmunitarias frente a antígenos proteicos. Baruj Bencerraf, Hugh McDevitt y colaboradores observaron que las cepas endogámicas de cobayas y ratones se diferenciaban en su capacidad para producir anticuerpos frente a polipéptidos sintéticos simples, y que su capacidad de respuesta se heredaba como rasgo mendeliano dominante. Los genes relevantes recibieron el nombre de **genes de respuesta inmunitaria (Ir, Inmune response)** y todos ellos están localizados en el MHC.

Los alelos del MHC cuyos productos se unen a esos péptidos, formando complejos péptido-MHC que pueden ser reconocidos por células T colaboradoras. Estas células T ayudan a las células B a producir anticuerpos. Las cepas que no responden expresan moléculas del MHC que no son capaces de unirse a péptidos derivado del polímero sintético, por lo que no pueden generar células T colaboradoras ni anticuerpos específicos del antígeno.

Las moléculas del MHC humano reciben el nombre de antígenos leucocitarios humanos (HLA, del inglés human leukocyte antigens)

El trasplante de órganos proporcionó un fuerte impulso para la detección y definición de los genes que controlan las reacciones de rechazo en el ser humano. Jean Dausset, Jan van Rood y colaboradores mostraron por primera vez que los pacientes que rechazan riñones trasplantados o sufren reacciones transfusionales frente a leucocitos a menudo presentan anticuerpos circulantes contra antígenos de los leucocitos de la sangre o del órgano del donante. Los sueros que reaccionan contra las células de individuos aloagénicos se llaman **aloantisueros** y se dice que contienen **aloanticuerpos**, cuyas dianas moleculares se denominan **aloantígenos**. Se supuso que estos aloantígenos son los productos de genes polimórficos que distinguen los tejidos extraños de los propios.

CAPÍTULO III.- SISTEMA INMUNE NATURAL

Estos aloantígenos se llamaron **antígenos leucocitarios humanos** (HLA) debido a que se expresan en los leucocitos humanos.

El uso de anticuerpos par el estudio de las diferencias aloantigénicas entre los donantes y los receptores se completó con la reacción leucocitaria mixta (MLR. Mixed leukocyte reaction), una prueba para el reconocimiento de células antigénicas por las células T.

Los genes identificados como determinantes del rechazo en los humanos definidos serológicamente (HLA-A, HLA-B, HLA-C) y se agrupan como **genes de clase I del MHC**. Los genes **Ir** de los humanos identificados mediante respuestas linfocitarias en la MLR (HLA-DR, HLA-DP y HLA-DQ) y se agrupan como **genes de clase II del MHC**.

Las moléculas de clase I del MHC presentan péptidos a las células T CD8⁺ y son reconocidas por estas, mientras que las moléculas de clase II presentan péptidos a las células TCD4⁺.

Los genes del MHC son los genes más polimórficos de los que se encuentran en el genoma de todas las especies analizadas.

La característica más sobresaliente que surgió a partir de los estudios de los genes del MHC humano es la extensión sin precedentes e inesperada de su polimorfismo. Se han identificado más de 150 alelos para algunos de los loci del HLA mediante análisis serológicos.

Los genes del MHC se expresan de manera codominante en cada individuo. Cada individuo expresa los alelos del MHC en los dos cromosomas heredados de los padres. Para el individuo, esto aumenta al máximo el número de moléculas del MHC disponibles para ligar péptidos y presentarlos a las células T.

El conjunto de alelos del MHC presentes en cada cromosoma se denomina **haplotipo del MHC**. En los seres humanos algunos alelos de HLA que se encuentran en distintos loci se heredan juntos con mayor frecuencia de la que cabría esperar por azar, fenómeno denominado **desequilibrio de ligamento**. La conclusión de que los genes

CAPÍTULO III.- SISTEMA INMUNE NATURAL

del MHC controlan no sólo el rechazo del injerto, sino también las respuestas inmunitarias a todos los **antígenos proteicos**.

Los dominios tipo Ig no polimórficos de las moléculas del MHC contienen sitios de unión para las moléculas CD4 y CD8 de las células T.

Las moléculas CD4 y CD8 son “correceptores” de las células T. Este es el motivo de que las células T CD4+ reconozcan únicamente péptidos presentados por moléculas de clase II, y que las células T CD8+ reconozcan péptidos presentados por moléculas del clase I. La mayoría de las células T CD4+ actúan como células colaboradoras, mientras que las células CD8+ son CTL.

Cada molécula del MHC consta de una hendidura (o surco) extracelular de unión al péptido seguida de un par de dominios tipo inmunoglobulina (Ig), y está anclada a la célula mediante dominios transmembrana y citoplásmicos.

Las moléculas de clase I están constituidas por dos cadenas: una cadena α codificada por el MHC (o cadena pesada) de 44 47 kD y una subunidad no codificada por el MHC de 12 kD denominada microglobulina β fig 4-4. Esta estructura forma la hendidura de unión al peptido de las moléculas de clase I. Su tamaño es suficientemente grande ($\sim 25 \text{ \AA} \times 10 \text{ \AA} \times 11 \text{ \AA}$) para unirse a péptidos de entre 8 y 11 aminoácidos en una conformación flexible y extendida. Todo individuo normal (heterocigótico) expresa seis moléculas de clase I diferentes en cada célula, que contienen cadenas α derivadas de los alelos de los genes HLA-A, HLA-B y HLA-C heredados de los padres.

Las moléculas de la clase II del MHC están compuestas por dos cadenas polipeptídicas unidas de forma no covalente, una cadena α de 32 a 34 kD y una cadena β de 29 a 32 kD. A diferencia de lo que sucede en las moléculas de clase II están codificadas por genes polimórficos del MHC.

CAPÍTULO III.- SISTEMA INMUNE NATURAL

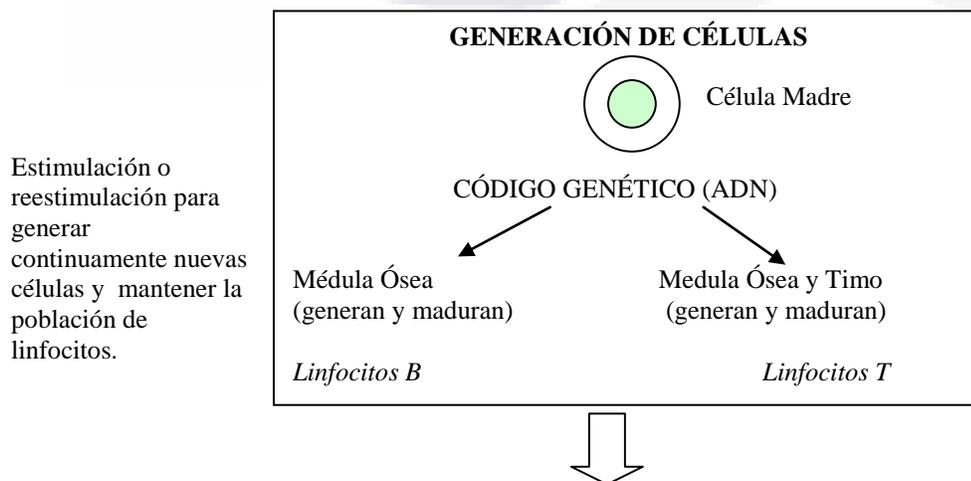
Un polipéptido no polimórfico denominado **cadena invariable** (I_i) se asocia a las moléculas de clase II recién sintetizadas. El I_i , desempeña papeles importantes en el tráfico de las moléculas de clase II y en la determinación del sitio de la célula en el que se unen los péptidos a las moléculas de clase II.

Características de las interacciones péptido-complejo principal de histocompatibilidad.

Las moléculas del MHC muestran una amplia especificidad para la unión a péptidos, y esta fina especificidad del reconocimiento de antígenos reside en gran medida en los receptores de antígenos de los linfocitos T.

Todo péptido frente al cual puede generarse una respuesta inmunitaria debe contener ciertos residuos que contribuyen a su unión a las hendiduras de las moléculas del MHC, así como otros residuos que se extienden desde las hendiduras, donde son reconocidos por las células T. Existen varias características importantes de las interacciones de las moléculas del MHC con los péptidos antigénicos.

3.12 Modelo del Funcionamiento del Sistema Inmune



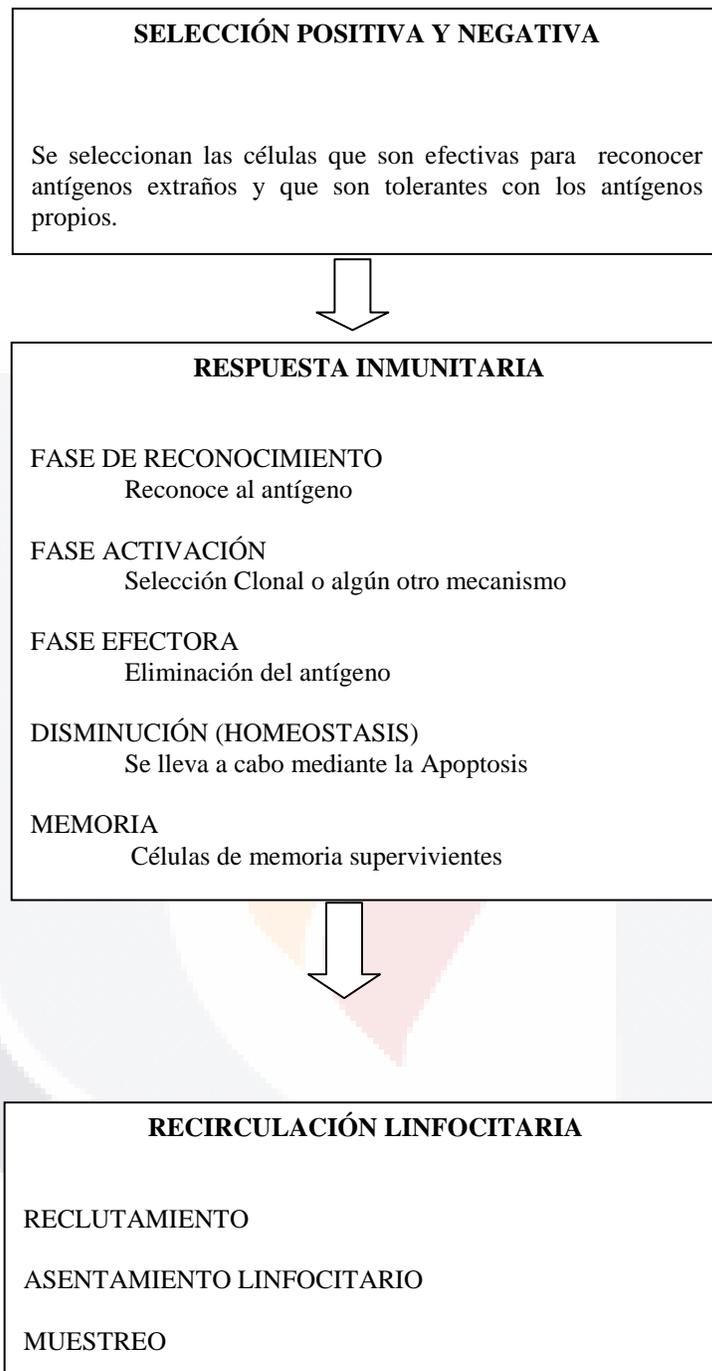


Figura 9: Modelo aproximado del funcionamiento general del Sistema Inmune Natural de acuerdo a los conceptos tomados del mismo para esta investigación.



Capítulo 4

Sistema Inmune Artificial

En este capítulo se hace una descripción de lo que es el Sistema Inmune Artificial, de algunas de las definiciones que se han dado a lo largo del desarrollo de esta área de investigación. Como parte de la teoría se analizan algoritmos que se han desarrollado basados en diferentes modelos y funciones del Sistema Inmune Natural, algunas de las aplicaciones. Se realizó una clasificación de algunos elementos del SIN en mecanismos de exploración y explotación, así como también de un análisis de cómo interpretar y utilizar la afinidad en problemas de Optimización Combinatoria y Detección de Malware.

4.1 Definición de Sistema Inmune Artificial

El Sistema Inmune Artificial (AIS del inglés Artificial Immune System) es una heurística bio-inspirada ya que se basa en el funcionamiento del SIN, básicamente en el del ser humano.

El SIA es un área de investigación relativamente nueva ya que su aparición dentro de la Inteligencia Artificial (IA) fue en el año de 1986 (Farmer et al., 1986), aunque en los últimos años ha tomado gran importancia. (Yoo et al., 1999), (De Castro et al., 2002), (De Castro et al., 2002a), (De Castro et al., 2003). Esto se debe principalmente a las características que tiene el SIA (Dasgupta, 1998; De Castro et al., 2002)

Algunos investigadores han dado definiciones de lo que un Sistema Inmune Artificial. A continuación se muestran algunas de las más relevantes, siendo la tercera definición la más divulgada pues se considera más completa.

"El sistema inmune artificial es un sistema computacional basado en los principios del sistema inmune natural" (Timmis et al., 2000)

"Los sistemas inmunes artificiales son metodologías inteligentes inspiradas en el sistema inmune, enfocadas a resolver problemas del mundo real" (Dasgupta, 1999)

"Los sistemas inmunes artificiales son sistemas adaptativos, inspirados en la teoría, modelos y principios inmunológicos y las funciones inmunes observadas, que son aplicados a la solución de problemas" (De Castro et al., 2002)

4.2 Características del SIA

El SIN cuenta con ciertas características que se consideran importantes en un sistema inteligente y que por tanto esta nueva área del SIA ha tratado de emular para la solución a diversos problemas y que diferentes investigadores del área han exaltado (Dasgupta, 1998; Timmis 2002; Cruz, 2004).

Las características se resumen a continuación:

Reconocimiento: El SIN puede reconocer entre lo propio y lo no propio, además puede reconocer cuando una célula propia es disfuncional, un elemento extraño no es dañino y cuando si lo es.

Extracción de características: El SIN tienen la capacidad de detectar características a nivel molecular para lograr detectar un patógeno a través de los receptores de los linfocitos, las células presentadoras y las moléculas como el Complejo Principal de Histocompatibilidad. Esto lo hace tomando los elementos moleculares que le interesan y eliminando el ruido o elementos no relevantes.

Diversidad: El SI tiene una gran capacidad para detectar cualquier intruso y generar el anticuerpo específico para ese antígeno.

Memoria: Puede recordar a través de sus células de memoria, en encuentro con algún antígeno. Por su estructura las células de memoria son más eficaces que los linfocitos no activados.

Distribuido: Esto es porque las células del SI se encuentran en todo el cuerpo.

Aprendizaje: El SIN aprende a través de diversos mecanismos. En primera instancia se da en los encuentros repetidos contra un mismo antígeno ya que la respuesta es

cada vez mejor y más rápida. En este proceso el mecanismo de la hipermutación somática, que viene después de una fuerte presión de selección, mejora los linfocitos de manera que la respuesta a determinado patógeno sea mejor. Este proceso es llamado maduración de la afinidad (Berek et al., 1993) el cual conduce a que las respuestas inmunitarias adaptativas sean cada vez más eficaces. Otro mecanismo de aprendizaje es el de la red inmune.

Tolerante a fallos: El SI se puede recuperar después de un fallo. . La auto-tolerancia es un aprendizaje que se lleva a cabo los órganos linfoides que cuentan con mecanismos para eliminar las células que reaccionan contra los antígenos propios del huésped.

Autónomo: El SI no tiene un control externo aún cuando es parte de organismo y tampoco tiene un control interno ya que actúa localmente.

Adaptable: El SI es capaz de enfrentarse a nuevos antígenos y recordar esos encuentros.

Reconocimiento: El SI tiene la capacidad de reconocer y responder a una gran cantidad de patrones. Además puede diferenciar entre células propias y extrañas.

Robusto: Se considera robusto ya que el aprendizaje, la memoria, la tolerancia a errores, etc., lo hacen capaz de resolver exitosamente problemas como la intrusión de un agente externo.

Dado que el SIA es muy flexible como muchas de las heurísticas bionspiradas, además de las características como las anteriormente mencionadas lo convierten en un área con gran potencial para aplicarlo a la solución de problemas complejos.

El área de aplicación es muy grande ya que se ha utilizado en áreas como Robótica, Seguridad en sistemas de cómputo y redes, Optimización, Control, Aprendizaje de máquina, etc.

4.3 Marco de trabajo para el SIA

Para crear un algoritmo basado en algún modelo o función del Sistema Inmunitario para la solución de algún problema, es posible utilizar el marco de trabajo propuesto por de Castro y Timmis (De Castro et al., 2002; De Castro et al., 2003).

Ellos proponen este marco de trabajo como un intento de unificar o estandarizar la forma de desarrollar los algoritmos inmunoinspirados como es el caso de la Redes Neuronales (RN), en las cuales existe un marco de trabajo bien definido para el diseño de RN en el cuál primero se tiene un conjunto de neuronas que son interconectadas y un algoritmo de aprendizaje con el que se entrena la red y se modifican los parámetros de la misma. En el caso de los Algoritmos Evolutivos (AE) se define una población de cromosomas (individuos) que durante un número de generaciones sufren un proceso de selección, mutación y reproducción o cruzamiento para obtener mejores individuos al terminar las iteraciones.

De Castro y Timmis identifican los elementos con los que al menos debe contar un algoritmo bioinspirado:

- 1.- Una representación para los elementos del sistema.
- 2.- Un conjunto de mecanismos para evaluar la interacción de los individuos con el ambiente y con los demás elementos. La función objetivo, parámetros de entrada, etc. representan el medio ambiente
- 3.- Procedimientos de adaptación que gobiernen la dinámica del sistema, esto es, cómo cambia el comportamiento en el tiempo.

CAPÍTULO IV.- SISTEMA INMUNE ARTIFICIAL

De acuerdo a lo anteriormente mencionado de Castro y Timmis proponen el siguiente marco de trabajo

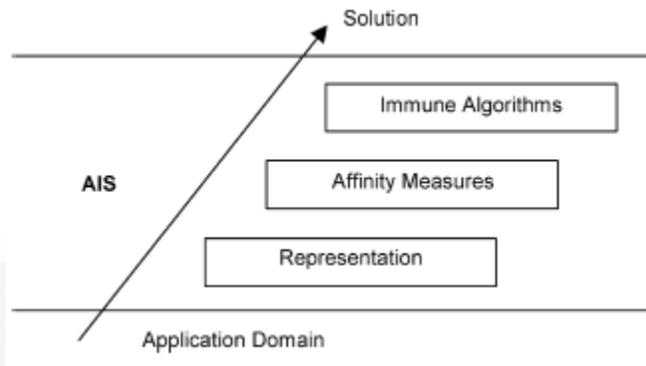


Figura 10 Marco de Trabajo para el SIA propuesto por Leandro Nunes de Castro y Jonathan Timmis (De Castro et al., 2002; De Castro et al., 2003).

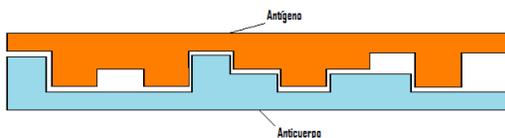
Este marco de trabajo puede ser usado para determinar los elementos del sistema y su representación dependiendo del área de aplicación, ya que es diferente la forma en que se van a representar los elementos que va a contener un sistema diseñado para resolver un problema de optimización que uno aplicado a la seguridad de redes de computadoras.

De esta manera las etapas del marco de trabajo se definen así:

Dominio de Aplicación: Problema a resolver

Representación: ¿Cómo se va a representar los componentes del sistema?

Medidas de Afinidad: Medir las interacciones entre los componentes del sistema



Distancia Ej. Euclideana

$$D = \sqrt{\sum_{i=1}^L (Ab_i - Ag_i)^2}$$

Ejemplo:

$$Ab = \langle ab_1, ab_2, ab_3, ab_4, ab_5 \rangle \quad Ag = \langle ag_1, ag_2, ag_3, ag_4, ag_5 \rangle$$

$$Ab = \langle 10.5, 2.5, 7.6, 11.8, 14.8 \rangle \quad Ag = \langle 12.5, 17.2, 11, 2.8, 18 \rangle$$

Aplicando la fórmula

$D = 17.98$, es la distancia que hay entre el antígeno y el anticuerpo. Así entre más grande la distancia, menor es la afinidad, y entre más corta, es más alta la afinidad.

Aunque dependiendo del problema se puede usar otra distancia como Hamming o alguna otra medida de distancia.

En algunos problemas de optimización, la forma de medir la afinidad es a través de la función objetivo, ya que no hay una población de antígenos contra quien medirse (de Castro, 2002).

Algoritmos inmunes: ¿Cómo van a interactuar los componentes del sistema para determinar la dinámica del sistema?

Ej. Selección Clonal, Red inmune, Teoría del Peligro, Selección Positiva y Negativa, etc. (Timmis et al., 2002)

Guía para desarrollar un SIA descrita por Leandro Nunes de Castro (De Castro, 2001)

1. Definición del problema.
2. Mapear el problema dentro del dominio de SIA.
 - a. Definir los tipos de células y moléculas inmunitarias que serán usadas.
 - b. Definir los principios inmunitarios que serán utilizados.
 - c. Definir la representación matemática de los elementos del SIA.
 - d. Evaluar las interacciones entre los elementos del SIA.

- e. Controlar la dinámica del SIA.
3. Regresar del dominio del SIA al dominio del problema real.

4.4 Modelos del SIA

Existen diversos algoritmos y modelos basados en mecanismos del SI. Estos mecanismos han sido estudiados para trasladarlos al área computacional y poder así aplicarlos a la solución de problemas en diversos campos.

Ya que existen gran variedad de algoritmos y modelos, De Castro y Timmis hacen una clasificación de estos (De Castro et al., 2003):

- Modelos de timo y de médula ósea
- Algoritmo de selección clonal
- Teoría de red inmune

Los **modelos de timo** son aquellos basados en el proceso del sistema inmune mediante el cual se crea la población de células detectoras.

Los **modelos de médula** ósea se refieren a la creación y maduración de los detectores del sistema inmune. Una de las características más importantes del los linfocitos maduros es su capacidad para poder distinguir entre los elementos que pertenecen al sistema de los agentes externos. La mayoría de estos modelos se basan en el principio de selección negativa y/o positiva.

Los **modelos de selección clonal** emulan el proceso mediante el cual el sistema inmune, ante la presencia de un antígeno específico, estimula únicamente a aquellos linfocitos que sean más afines, para después ser clonados y mutados.

Los modelo **de red inmune**, emulan las interrelaciones que existen entre anticuerpos, aún en una ausencia de antígenos, es decir, la estimulación y supresión entre anticuerpos.

Otra clasificación dada por Jason Brownlee (Brownlee, 2006) considera la siguiente clasificación:

- Paradigma de la Selección Negativa
- Paradigma de la Selección Clonal
- Paradigma de la Red Inmune
- Paradigma de la Señal de Peligro
- Sistemas Inmunes Artificiales Distribuidos
- Otros Trabajos

Un nuevo modelo que ha surgido es el **de la teoría del peligro**, el cuál se basa en las señales de peligro que envían las células que están sufriendo porque están muriendo y esta muerte no estaba planeada. Esta teoría no distingue lo propio y lo no propio, sino lo peligroso (Matzinger, 2001) y está basado en el trabajo de Polly Matzinger (Matzinger, 1994), (Matzinger, 2002).

4.4.1 Generación del repertorio inicial de anticuerpos

Este algoritmo se basa en la generación de linfocitos B en la médula ósea y de la generación de clonas en el mecanismo de la selección clonal. En esta generación participan los segmentos génicos V, D y J en cadenas ligeras y pesadas de inmunoglobulina (Ig) así como las proteínas de cadenas H y L.

El genoma de un individuo corresponde a 4 bibliotecas

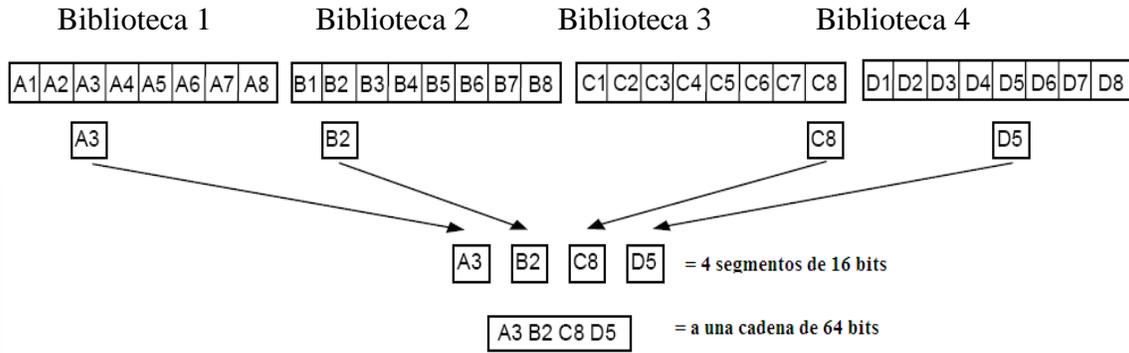


Figura 11: Modelos de la generación del repertorio inicial de anticuerpos (De Castro et al., 2000)

4.4.2 Algoritmo de Selección Negativa

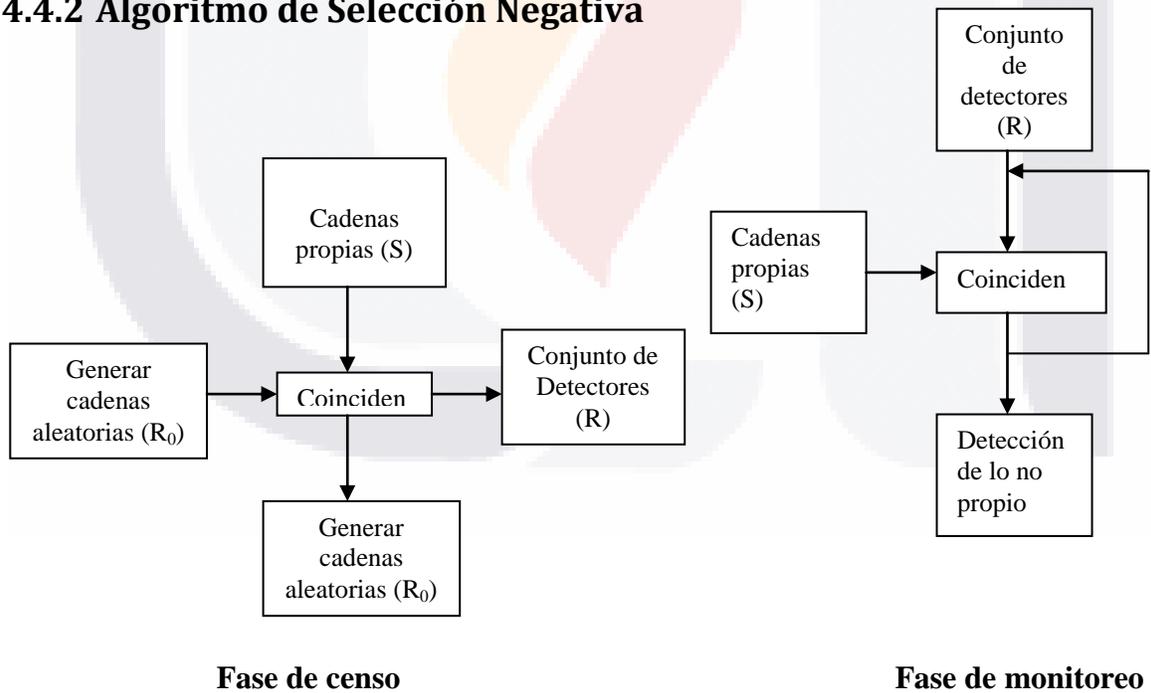


Figura 12: Modelo del Algoritmo de Selección Negativa, Forrest et al. 1994

Este algoritmo se basa en el proceso que sigue el SIN para detectar y eliminar las células que reaccionan contra el organismo y que pueden tener un efecto negativo en la buen funcionamiento del mismo.

A continuación se muestra una tabla con algunos de los trabajos que han utilizado este modelo para la solución de los problemas propuestos.

Tabla 4 Resumen de trabajos relevantes en los que se utiliza el modelo de Selección Negativa

Año	Autores	Descripción
1993,	Robert E. Smith et al.	Reconocimiento temprano usando Algoritmos Genéticos
1993	Forrest et al.	Reconocimiento temprano usando Algoritmos Genéticos, trabajo en el que se definen claramente la propuesta y resultados.

4.4.3 Algoritmo de Selección Clonal

✓ CLONALG

Figura 13: Modelo del Algoritmo de Selección Clonal utilizado en optimización, aprendizaje de máquina y reconocimiento de patrones, De Castro et al. 2002.

En este algoritmo primero se genera el repertorio inicial de anticuerpos que luego son evaluados y se selecciona un subconjunto de los anticuerpos. Estos son clonados generando un nuevo conjunto C el cuál pasa a la etapa de maduración donde se aplica la hipermutación somática para generar individuos más afines, este conjunto se le llama C*. Esto nuevos individuos son evaluados y se selecciona un subconjunto de C*. El proceso se repite hasta que alcance el criterio de parada.

✓ Algoritmo Clonal Mejorado propuesto por Yunyi Zhu

(Zhu et al., 2007)

```
Inicialización ();
Desde i=1 Hasta NG
Desde j=1 Hasta NCS
Hipermutación somática ()
Evaluación ();
Si (D < cDis)
cDis=D;
Actualizar();
Else
Edición de Receptor();
CC++
FinSi
Si (CC > CMAX)
Cruzamiento propio();
CC=0;
FinSi
FinDesde
FinDesde
```

Los parámetros del algoritmo son los siguientes

NG= Número de generaciones

NCS=El tamaño de la población de clones

Dopt=Distancia más corta de la solución óptima.

D, Dm, Db= Distancia, distancia promedio, y la mejor distancia generada por el algoritmo (realizado 10 veces). cDis= Distancia actual.

PDM, PDB= El porcentaje de desviación (diferencia) de la Dopt de la Dm y la Db

$PDM = 100 * (Dm - Dopt) / Dopt$, $PDB = 100 * (Db - Dopt) / Dopt$

CC= El contador de inactivación.

CMAX= El criterio de cruzamiento de si mismo

Ppm, Pre = probabilidad de punto de mutación (hipermutación somática) y edición de receptor, $Ppm + Pre = 1$.

T= tiempo promedio de operación

✓ Algoritmo propuesto por Shangce Gao

(Gao et al., 2007)

```
PoblaciónInicial();
CalcularFunciónEnergía();
Seleccionar(); // seleccionar los n mejores anticuerpos
Desde i=1 Hasta i=gen
    Desde j=1 Hasta j=n
        Contenedori = Anticuerpoi;
        ClonarAnticuerpoi (pi)
        Hipermutación();
        EdicióndeReceptor();
    FinDesde
CruzamientoGeeedy();
Reselección();
RecocidoSimulado();
Actualización();
NuevosAnticuerpos(nc);
FinDesde
```

✓ **Algoritmo propuesto por M. Bakhouya**

(Bakhouya et al., 2007)

```

Inicialización();
Crear agente móvil A;
A.ListaCiudades=Ciudades;// conjunto de ciudades
A.CiudadInicial=Random (Ciudades); // se posiciona al agente en una ciudad inicial
A.ListaCiudadesVisitadas={ }; //Conjunto de ciudades visitadas
A.mAfinidad=dist; // afinidad máxima generada aleatoriamente
A.CiudadActual=Nula; // La ciudad en la cuál el agente está posicionado
A.ÚltimaCiudad=A.CiudadActual // La última ciudad visitada
A.cAfinidad=0; // Afinidad actual de el recorrido actual
// El agente termina si todas las ciudades son visitadas
Mientras (A.ListaCiudades <> Nula) Hacer
    A.cAfinidad= A.cAfinidad+ $\delta$ (ÚltimaCiudad.CiudadActual)
    Si (A.CiudadActual nopertencea A.ListaCiudadesVisitadas y A.cAfinidad <
        A.mAfinidad)
        A.ListaCiudadesVisitadas(CiudadActual)
        A.ListaCiudades.Borrar(CiudadActual) // El agente se clona a si mismo y se mueve
        con su clon
        A.ListaCiudades=CiudadActual
        B=A.Clonar(); // Si hay al menos dos vecinos
        // n1 es seleccionada aleatoriamente de los vecinos siempre y cuando n1 no pertenezca
        a la lista de ciudades visitadas.
        A.CiudadActual=n1; B.Mover(); // n2 es seleccionada aleatoriamente de los vecinos
        siempre y cuando no pertenezca a la lista de ciudades visitadas y n1
        B.CiudadActual=n2; A.Mover();
        SiNo
        // Por medio de la selección positiva el agente se elimina el mismo pues no hace un r
        ecorrido válido.
        A.Morir(); // Solución menos útil
    FinSi
FinMientras
// Por medio de la selección negativa se elimina a si mismo si otro agente tiene un mejor
recorrido
A.Morir(); // Solución mala
    
```

✓ Hipermutación Somática

Este es un elemento muy importante en la maduración de los anticuerpos, ya que va haciéndolos más afines al antígeno. La hipermutación somática es inversamente proporcional a la afinidad (Cruz, 2004), es decir, cuanto más afín es el anticuerpo al antígeno, menor es el porcentaje de mutación.

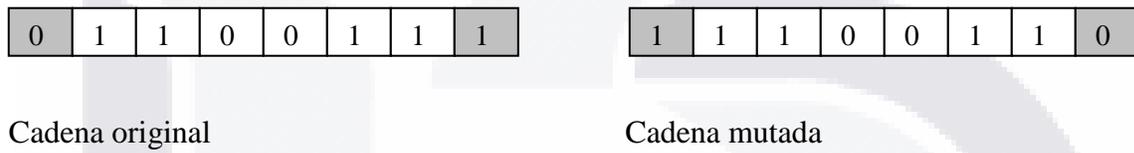


Figura 14 Ejemplo de cadenas modificadas por la Hipermutación Somática

4.4.4 Teoría del peligro

Los siguientes puntos deben considerarse para decidir utilizar esta teoría dentro de un modelo (Aickelin et al., 2002)

- La selección negativa puede ser imperfecta y se pueden dar falsos positivos.
- Las fronteras de lo propio y lo no propio no son claras ya que elementos propios y no propios comparten regiones comunes.
- Lo propio durante el tiempo. De esta forma, se puede esperar problemas con las células de memoria, las cuales se pueden volver inexactas y ser autoreactivas.

4.5 Mapeo del SIN y del AIS

Tabla 5: Mapeo de los elementos que conforman el SI y los que corresponden el AIS, Timmis et. al 2002.

SISTEMA INMUNE	SISTEMA INMUNE ARTIFICIAL
Receptores	Cadenas de bits
Linfocito (B y T)	Detectores
Células de memoria	Detector de memoria
Patógeno	Cadena no propia
Zona de unión	Aproximación a la afinidad de la cadena
Circulación	Detectores móviles
Complejo de Histocompatibilidad	Parámetros de representación
Citoquinas	Nivel de sensibilidad
Tolerancia	Selección negativa distribuida
Co-estimulación- Señal 1	Afinidad que excede el umbral de activación
Co-estimulación-Señal 2	Operador o factor humano
Clones de lo linfocitos	Detector replicado

4.6 Comparación del Sistema Inmune Artificial con otras Heurísticas

4.6.1 Algoritmos Genéticos

El AIS tiene relación con la Computación evolutiva ya que está basado en poblaciones y tiene operadores como la selección, la reproducción, etc.

CAPÍTULO IV.- SISTEMA INMUNE ARTIFICIAL

En la Tabla 6 se mencionan las principales diferencias y similitudes entre el SIA y los AGs (Aickelin et al., 2004)

Tabla 6: Comparación entre el AG y el SIA.

	Algoritmo Genético	SIA
Componentes	Cromosoma	Cadenas de atributos
Localización de los componentes	Dinámica	Dinámica
Estructura	Componentes discretos	Componentes discretos
Almacenamiento del conocimiento	Cromosomas	Concentración de los componentes
Dinámica	Evolución	Evolución y aprendizaje
Metadinámica	Selección y eliminación de componentes	Selección y eliminación de componentes
Interacción de los componentes	Cruzamiento	Reconocimiento, afinidad
Interacción con el ambiente	Función a optimizar Fitness	Función de similitud

4.6.2 Redes Neuronales

El SIA está relacionado con las Redes Neuronales (RN) porque tiene un mecanismo en el que se forma una red para detectar un antígeno y mantener la memoria (Aickelin et al., 2004). En la

Tabla 7 se mencionan los elementos que son similares entre estas dos técnicas.

Tabla 7: Comparación entre el AG y RN.

	Redes Neuronales	SIA
Componentes	Neuronas	Cadenas de atributos
Localización de los componentes	Pre-definidas	Dinámica
Estructura	Componentes en red	Componentes discretos
Almacenamiento del conocimiento	Fuerza de la conexión	Concentración de los componentes
Dinámica	Aprendizaje	Evolución y aprendizaje
Metadinámica	Construcción y poda o eliminación de conexiones	Selección y eliminación de componentes
Interacción de los componentes	Conexiones de la red	Reconocimiento, afinidad
Interacción con el ambiente	Estímulos externos	Función de similitud

4.7 Aplicación de algoritmos del SIA en Optimización

4.7.1 Optimización MultiModal

Fukuda *et al.* propusieron un algoritmo inmune que modelaba matemáticamente la diversidad inmune, teoría de la red inmune y la selección clonal como una función multimodal de un problema de optimización. La diversidad y los vectores de múltiples soluciones introducidos fueron mantenidos como la memoria del sistema.

Las principales características de este método es la diversidad entre las distintas soluciones candidatas y una eficiente búsqueda paralela.

De Castro y Von Zuben (De Castro, 1999), propusieron un algoritmo de selección clonal, el cuál toma en cuenta la *maduración de la afinidad* de la respuesta inmune, esto para resolver problemas complejos, como aprendizaje y optimización multimodal. En este algoritmo no considera ninguna sofisticación matemática para mejorar su funcionamiento en tareas particulares.

De Castro y Timmis (De Castro et al., 2002) propusieron un SIA para problemas de optimización multimodal. Esta propuesta se basa en el principio de selección clonal y en la teoría de red inmune.

4.7.2 Optimización Combinatoria

De Castro y Von Zuben (De Castro et al., 2002) propusieron un algoritmo basado en la teoría de la selección clonal, llamado CLONALG, el cuál toma en cuenta la *maduración de la afinidad* de la respuesta inmune, esto para resolver problemas complejos, como aprendizaje y optimización multimodal y combinatoria.

4.7.3 Optimización con Restricciones y Multiobjetivo

Hajela, Lee y Hajela, Yoo (Hajela et al., 1996), proponen un algoritmo en el cuál separan a los individuos infactibles y se busca reducir la distancia genotípica entre estos dos grupos intentando con ello reducir la cantidad de violación a las restricciones.

Yoo y Hajela (Yoo et al., 1999) propusieron un algoritmo como la primera propuesta de un SIA para optimización multiobjetivo en el que utilizaron una función agregativa que combina la función objetivo y la cantidad de violación de restricciones, la cual se utiliza como aptitud. Este esquema consiste realmente de un algoritmo genético simple al que se le acopla un SIA cuya búsqueda se basa en similitudes entre cadenas binarias.

Anchor y Lamont (Anchor et al., 2002) propusieron un SIA multiobjetivo que usa selección basada en ordenamiento lexicográfico y optimalidad de Pareto en un algoritmo evolutivo para detectar intrusos. El algoritmo está enfocado hacia la aplicación y no a la solución de problemas multiobjetivo en general.

G. C. Luch, C. H. Chuen (Luch et al., 2003), proponen un algoritmo para optimización multiobjetivo que aplica el principio de selección clonal, recombinación de segmentos de ADN y diversificación de anticuerpos. En esta propuesta los objetivos del problema toman el papel de los antígenos, mientras que las posibles soluciones representan a los anticuerpos. Se aplica clonación y mutación a los anticuerpos no dominados, con parámetros definidos por el usuario. Se define una medida de avidez que es proporcional a la jerarquía del anticuerpo (de acuerdo a dominancia de Pareto) y a su distancia Euclidiana con respecto a los demás anticuerpos. Esta medida de avidez se utiliza para aplicar por torneo. A continuación se crea una biblioteca utilizando fragmentos (de tamaño definido por el usuario) de los anticuerpos ganadores del torneo, la cual es usada para formar la siguiente generación de anticuerpos. Una desventaja clara de este algoritmo es la gran cantidad de parámetros que introduce. Además, no se especifica como manejar los problemas con restricciones. Los autores argumentan que su algoritmo produce resultados competitivos con respecto a SPEA (Knowles et al., 2000).

Cruz Cortés Nareli (Cruz, 2004) propone un modelo, basado en el algoritmo desarrollado por Hajela y Lee, para resolver problemas de optimización multiobjetivo mediante el frente de Pareto. El algoritmo desarrollado utiliza la selección clonal dentro de un algoritmo genético del cuál se reportan muy buenos resultados.

4.8. Afinidad

Los conceptos de afinidad y avidéz son de suma importancia en lo que al SIN se refiere. Es a través de estos que se puede explicar como se da la selección de un determinado linfocito y la fuerza con la que se une el antígeno a los receptores del linfocito.

La afinidad en términos del SIN se mide con la fórmula descrita en la sección 3.11 que explica como es la interacción entre un antígeno y un anticuerpo.

También se han descrito sistemas de ecuaciones diferenciales que modelan la afinidad ().

Escribir el conjunto de ecuaciones.

Como se describe en el capítulo 3, la afinidad es un concepto clave en lo que a la inmunidad adaptativa se refiere y tiene ese mismo sentido en el AIS.

En cada Sistema Inmune Artificial que se desarrolla se debe definir el concepto de afinidad y la forma en como se va a medir o calcular, lo cual va a variar dependiendo del problema a resolver.

En relación con los problemas de optimización comúnmente se la afinidad es igual a la calidad de la solución o población de soluciones Pac que se tiene en determinado tiempo t . La afinidad de cada elemento en el tiempo t ($Pac(t)$) va a ser calculado evaluándolos con la función objetivo.

Sin embargo en algunos casos se hace uso de otros mecanismos para evaluar la afinidad de los elementos que conforman la población.

A continuación se dan algunos ejemplos de cómo se ha definido y medido la afinidad.

CAPÍTULO IV.- SISTEMA INMUNE ARTIFICIAL

Cruz Cortés Nareli (Cruz, 2004), propone un algoritmo de AIS, basado en lo descrito en (Hajela, 1996; Hajela, 1999). Primero genera una población que es evaluada de la cual se toma el 3% mejor como antígenos, y la población restante se consideran anticuerpos. Para evaluar su afinidad (similitud) define una función que llama de aptitud. Esta función es una función de distancia genotípica (Hamming) entre un anticuerpo y un antígeno tomado aleatoriamente. A los anticuerpos que son más afines se les aplican los operadores de cruce y mutación para entrar nuevamente a los pasos descritos anteriormente. En este caso en la tesis se mencionan que se usaron otras medidas de distancia, sin embargo la que mejores resultados dio fue la Hamming.

Se hace notar en la tesis que similitudes a nivel genotipo resultarán en similitudes a nivel fenotipo, lo cual es sustentado por (Hajela, 1996; Hajela, 1999) y los obtenidos en la tesis.

Backouya (Backouya, 2005), propone que la afinidad de cada vector v a estar dado por el costo de cada recorrido.

Dado $C=\{a, \dots, z\}$ conjunto de ciudades, $A=\{(x,y), x, y \in C\}$ el conjunto de aristas, y $\partial(x,y)$ es el costo de la medida asociada con la arista.

$\Sigma \partial(x,y) = \partial(1,2) + \partial(2,3) + \partial(3,4)$ medida de afinidad

En el trabajo de Zhu (Zhu et al., 2007) no se menciona claramente pero considerando el algoritmo que se propone, la función de evaluación da como resultado el valor de afinidad de cada elemento de la población. Es pues la sumatoria del costo del recorrido lo que determina la afinidad.

Otro trabajo es el propuesto por Gao (Gao et al., 2007), en el cual parte del algoritmo está basado en recocido simulado, por lo cual la función de energía que proponen es la siguiente:

$$E = \sum D(v, v_{i+1})$$

donde

$$D(v, v_{i+1}) = d_{ij} \text{ distancia del ir del vértice } i \text{ al } i+1$$

Cuando se sigue el algoritmo del AIS, para determinar la afinidad de cada anticuerpo se usa la función $D(\cdot)$, es decir, la función del costo del recorrido.

Leandro Nunes de Castro (De Castro, 2002b), en el trabajo en el que trabaja con el algoritmo CLONALG, basado en la teoría de la selección clonal, propone el uso de este para resolver problemas de optimización combinatoria, en particular el problema del TSP (Travelling Salesman Problem, por sus siglas en inglés). El algoritmo tiene un procedimiento por medio del cual se mide la afinidad, que en este caso viene dada por la evaluación de la función objetivo en cada anticuerpo.

En esta misma línea, en que la afinidad es igualada con la calidad de la solución, están los trabajos de otros investigadores que usan el AIS para resolver problemas de Optimización como el TSP.

Otra aplicación del AIS es la detección de intrusos o malware en la cual también definen la afinidad o similitud, de acuerdo a ciertas características.

En el trabajo de Dan Taylor (Taylor et al., 2002) comenta que para utilizar modelos del AIS basados en la selección negativa es necesario convertir los datos de series de tiempo en subcadenas de bits de cierto tamaño. Por ejemplo para una serie de tiempo $T = \{t_1, t_2, t_3, \dots, t_4\}$ donde el tamaño del patrón es 4, las subcadenas quedan de la siguiente manera $P_0 = \{t_0, t_1, t_2, t_3\}$, $P_1 = \{t_4, t_5, t_6, t_7\}$, los cuales tomarán valores entre $[0,1]$. Lo propio se define con datos de comportamientos normales, donde los patrones representan los comportamientos normales del sistema. Después se genera el conjunto de detectores y usando la selección negativa se comparan con los patrones con los cuales no deben

empatar. Así cada nueva patrón de comportamiento del sistema se compara con los detectores para buscar anomalías.

4.9 Mecanismos de Exploración y Explotación

Los algoritmos heurísticos requieren de hacer un búsqueda eficiente del espacio de búsqueda sin necesidad de hacer un barrido completo del mismo.

Para esto es necesario detectar zonas que son promisorias en la cuales buscar una buena solución.

Los términos exploración y explotación, también conocidos como diversificación e intensificación, son usados frecuentemente en el diseño de algoritmos basados en el alguna técnica heurística y en los cuales se desea tener un equilibrio. Esto con la finalidad de tener un algoritmo eficiente que obtenga buenas soluciones y que su costo computacional no sea muy alto.

El Sistema Inmune Artificial, como otras heurísticas, también integra mecanismos de exploración y explotación, esto basado en las características y modelos tomados del Sistema Inmune Natural.

Como se menciona en la sección 4.2, una de las características que tiene el AIS y que toma del SIN, es la diversidad. Esta se logra a través de la generación celular, en la médula ósea, y que constantemente esta generando nuevas células para mantener un equilibrio en la población de dichas células.

Como se puede notar, el AIS es de manera natural poblacional, ya que el SIN trabaja con poblaciones de células, que elimina o mantiene vivas dependiendo si han sido activadas o no.

El mecanismo de explotación es propiamente la hipermutación somática y la recombinación genética, mecanismos que son usados en el SIN para aumentar la

CAPÍTULO IV.- SISTEMA INMUNE ARTIFICIAL

afinidad del linfocito que ha sido activado y que tiene la función de eliminar al antígeno. Este proceso se ha llamado maduración de la afinidad.

Otros mecanismos que existen en el SIN son la recirculación linfocitaria, la migración y el asentamiento linfocitario ver secciones 3.9, 3.10.

A grandes rasgos la recirculación y la migración tienen la función de hacer posible que los linfocitos recorran el organismo en busca de algún foco de infección a causa de un patógeno que ha invadido al huésped.

El asentamiento linfocitario se refiere al mecanismo que tiene la función de hacer que los linfocitos permanezcan en la zona donde se tiene la infección y que hacen falta para lograr eliminar al antígeno extraño dañino.

Tabla 8 Clasificación de los mecanismos de exploración y explotación del SIN y que son integrados al AIS.

Exploración	Generación celular Recirculación Linfocitaria Migración Linfocitaria
Explotación	Maduración de la afinidad Hipermutación somática Recombinación

Capítulo 5

Metodología propuesta

En este capítulo se hace un análisis de las metodologías y marcos de trabajo más relevantes desarrollados para la modelación y diseño de Sistemas Inmunes Artificiales.

Además del análisis se hace una descripción de la metodología propuesta, describiendo las fases de las que se compone y los elementos que se debe considerar en cada una de ellas.

5.1. Introducción

Es este capítulo se hace un análisis de las metodologías existentes relacionadas con el diseño de heurísticas para optimización y también de las metodologías diseñadas específicamente para el desarrollo de algoritmos del SIA. Las metodologías seleccionadas son las consideradas más relevantes de acuerdo a la relación con el tema y su impacto en el estado del arte.

A partir de las conclusiones respecto de las metodologías analizadas se hace una comparación contra la metodología que se propone en esta tesis, para lo cual primeramente definen claramente las fases que la componen y sus elementos, así como la simulación del SIN tomando en cuenta las células, mecanismos, teorías, etc. que se consideran importantes para la solución de problemas.

5.2. Análisis de Metodologías y Marcos de trabajo

A continuación se mencionan los trabajos relacionados con marcos de trabajo y metodologías realizados y que de diferentes formas influyen este trabajo.

Como se ha mencionado, algunos modelos desarrollados del SIA tienen una característica en común, son algoritmos simplistas, como lo dice (Stepney et al. 2004) "reasoning by metaphor".

Esta situación ya había sido vislumbrada en (Timmis et al, 2002), ya que se hace mención de que algunos conceptos no están del todo claros, creando así una mala interpretación de los mismos al momento de ser simulados en los algoritmos del SIA. Esto da como resultado algoritmos que no están del todo apegados al funcionamiento real del SIN y que probablemente su desempeño no sea tan bueno como se espera.

Otro efecto es que no se puede diferenciar claramente los algoritmos del SIA con respecto a otras heurísticas existentes. Por ejemplo el algoritmo de la selección clonal es muy similar a los AGs, esto es por un lado, porque la Teoría de la Selección Clonal se puede considerar como un micromundo de la evolución, por lo que lo acerca en general a los algoritmos evolutivos (AE), y por otro lado, porque no se ha clarificado de manera contundente la diferencia de los operadores y la forma en que trabajan para hacer funcionar el SIA.

En los trabajos (Stepney et al. 2004) y su continuación (Stepney et al. 2005), en el cual se hace un análisis matemático y un caso de estudio, los autores logran definir un marco de trabajo conceptual para el SIA, aunque ellos mencionan que puede ser aplicado a cualquier algoritmo bioinspirado Ver Figura 15. Este marco de trabajo

CAPÍTULO V.- METODOLOGÍA PROPUESTA

consta de 4 Aspectos importantes que son tomar pruebas, observaciones, etc. de los sistemas biológicos, seguido por la representación y el marco de trabajo o principio hasta llegar al Algoritmo bioinspirado. En este trabajo se hace una clara diferenciación de los algoritmos del SIA, los basados en población y los basados en redes, pues esto hará una diferencia en el algoritmo que se va a generar.

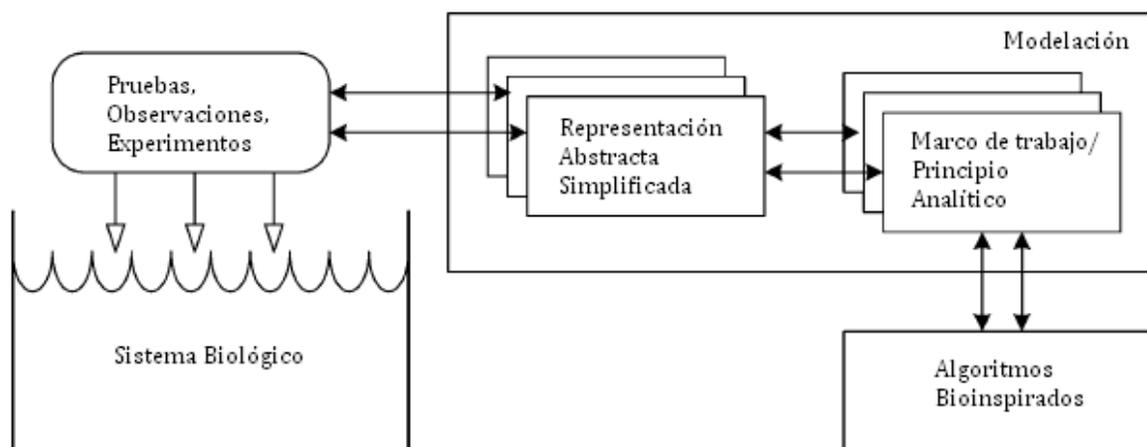


Figura 15. Marco de trabajo Conceptual para Sistema Inmune Artificiales propuesto por Stepney et al., 2004

Para lograr tener un buen modelo del algoritmo, se proponen 6 aspectos que se deben cuestionar, los cuales se citan textualmente a continuación:

Openness (continuidad, claridad, abertura): No queremos que nuestro cómputo se detenga; nosotros queremos evolución continua, crecimiento continuo, adición continua de recursos: esto es, abierto, far-from-equilibrium system. ¿Cuanto (continuo, claro) es necesario?, ¿Como es controlada la openness por estructura e interacción?, ¿Como se mantiene la unidad del sistema en la presencia de (abertura, claridad, continuidad)?

Diversidad: heterogeneidad esta presente en todos los sistemas biológicos complejos, y ocurre en la estructura, comportamiento e interacciones. ¿Cuando podemos hablar de un promedio de agentes? ¿Cuanta diversidad es necesaria

dentro de un nivel de una estructura? ¿Entre niveles? Cuánto cuesta esto? ¿Cómo esto combate la fragilidad?.

Interacción: Los agentes interactúan con su ambiente y con cada uno de los otros. ¿Cuáles son las características de interacción dentro de niveles estructurales? ¿Entre niveles? ¿Cuál es el balance entre computación y comunicación?

Estructura: Sistemas biológicos tienen una estructura en una variedad de niveles, aún cuando los niveles no estén claramente definidos. ¿Son los niveles que discernimos artefactos de nuestro modelo del marco de trabajo? ¿Como podemos reconocer los niveles? ¿Cuándo es la jerarquía un modelo estructural adecuado? ¿Cómo afecta la estructura la interacción? ¿Qué son las relaciones entre estructuras físicas y estructuras de información? ¿Qué o Cuál es la relación entre la función de especialización? ¿Con localización de función?.

Escala: Sistemas biológicos tienen una basta escala, a vasto número de componentes. ¿Cuándo y como “más” se convierte “diferente”? ¿Qué son los puntos críticos/ fases de transición? ¿Cuánto pequeño puede ser un sistema, y aún ser emergente? Cuándo un sistema es demasiado grande? ¿Que tan importante es una modelación multiescala? ¿Cuáles son las relaciones entre escala y diversidad?

Preguntas genéricas: aplican a cada pregunta de la meta-prueba del área X: ¿Cuál es el rol de X en un sistema?. ¿Cómo y cuándo X emerge?, ¿Cómo evoluciona X?, ¿Cómo la representación o personificación física afecta X? ¿Como puedo explotar X?.

Estas preguntas resultan útiles para analizar en que punto se encuentra el algoritmo y en que aspectos puede ser mejorado. En este trabajo no se realizó implementación alguna del algoritmo, solamente se hizo un caso de estudio en el cual se mostraba

como aplicar las 6 preguntas anteriormente mencionadas a un sistema artificial que se interrelacionaba con otros sistemas.

En el trabajo de (Newborough, et al. 2005) proponen un marco de trabajo general para algoritmos bioinspirados basados en población. Esta investigación está influenciada también por las ideas de (Stepney et al., 2004), en el cual buscan analizar un grupo de diferentes algoritmos bioinspirados como Algoritmos Genéticos genéricos, Selección Negativa, Selección Clonal, Optimización de Cúmulo de Partículas. Se consideran ciertos componentes y comportamientos e interacciones que hay en ellos para determinar una generalización a través de ellos.

Como parte del estudio, determinaron 6 conceptos o elementos que forman parte de cada uno de los algoritmos, los cuales se mencionan a continuación

1. **Crear:** Crear individuos para la población
2. **Evaluar:** Cada individuo se evalúa por su afinidad a la solución
3. **Probar:** Probar si se ha alcanzado el criterio de termino.
4. **Seleccionar:** Seleccionar ciertos individuos de la generación actual, basados en su afinidad, para ser usados en la creación de la siguiente generación.
5. **Proliferación, Expansión:** Crear nuevos individuos para la siguiente generación.
6. **Mutar:** Cambiar los individuos seleccionados

Una vez que se sitúan en este marco de trabajo, se pueden tener características de los diferentes algoritmos y agregarlos a alguno de los otros. Este trabajo fue aplicado en FPGAs para probar el marco de trabajo. Este es uno de los trabajos pioneros en tomar ideas del trabajo (Stepney et al., 2004), aunque no sigue completamente las ideas propuestas, y se considera el más completo en lo que implementación se refiere (Andrews, 2008).

Otro trabajo que está basado o sigue ideas de (Stepney et al., 2004), es el realizado por (Twycross & Aickelin 2005) en el que incluye elementos de la inmunidad innata la cuales son introducidas a partir de un marco de trabajo para desarrollar modelos. Los

autores presentan una tabla con las características de la inmunidad innata y mencionan que es importante que a los algoritmos del SIA se les agreguen estas características para que realmente tenga las funciones de un SIN.

(Twycross & Aickelin 2005) utilizan las preguntas de (Stepney et al., 2004) para analizar los modelos biológicos. Las 6 preguntas del modelo conceptual son usadas para comparar los mecanismos de la inmunidad innata y la inmunidad adaptativa. El marco de trabajo presentado no queda claro y el modelo conceptual parece ser usado sólo para estructurar el análisis de la inmunidad innata y los beneficios que puede dar al SIA.

El trabajo de (Guzella et al. 2007a) propone usar el tipo de células reguladoras T para incorporar al SIA. Ellos sugieren que al utilizar estas células en los algoritmos del SIA, les darán a los modelos un aspecto más cercano a la biología y un mejor desempeño en la solución de problemas reales. Los modelos que se presentan en este artículo, podrían empezar en algún punto de algún enfoque similar al modelo conceptual de Stepney (Andrews, 2008).

5.3. Desarrollo de la Metodología

Toda Metodología consta de pasos o fases para llegar a un punto u objetivo. Estos pasos se realizan de manera sistemática y repetidamente hasta obtener los resultados que esperados, como la solución de determinado problema mediante una técnica particular.

Fases de la Metodología

1 Planteamiento del Problema

Esta es la fase inicial de la metodología, en la cual lo que se pretende es plantear y modelar el problema de manera tal que pueda ser trasladado al SIA para ser resuelto.

Los puntos más relevantes en esta fase son tanto la modelación propia de problema así como la clara definición de los elementos que adelante se mencionan.

Dentro de esta fase se deben definir los siguientes puntos del problema:

- Modelar/ Plantear el problema
- Entradas (Variables).
- Salidas (Resultado).
- Tipos de variables (discretas, continuas).
- Restricciones del problema
- Alfabeto con el que se van a representar las soluciones.
- Representación de las soluciones factibles (cadenas)
- Función Objetivo (hacer los ajustes necesarios).
- Función para evaluar las soluciones
- Medida de afinidad.
- Otros elementos

2 Análisis

En esta fase se eligen los conceptos, modelos y elementos que se van a tomar del SIN y el nivel de simulación que se requiera de acuerdo al tamaño, tipos de variables, entre otras características del problema.

De acuerdo a lo definido en la fase anterior, se elegirá el método de optimización local, si se considera necesario o alguna otra técnica, en caso contrario estos no serán parte del pseudocódigo.

Los puntos más relevantes de esta fase son los siguientes:

- Desarrollar el pseudocódigo que represente el nivel de simulación que se ha elegido. Este pseudocódigo debe contener las funciones y variables que representen los modelos que se toman del SIN.
- Analizar el algoritmo e identificar posibles riesgos.

A continuación se muestra como el sistema inmune lleva a cabo en lo general sus funciones; esto de acuerdo a los elementos, funciones y conceptos de interés para esta investigación.

Este diagrama permite identificar elementos clave que debieran tener los algoritmos del SIA para ser aplicado a la solución de problemas de OC o a otro tipo de problemas.

En este diagrama se hace una comparación de los elementos que se toman del SIN y como se representa en el SIA.

De acuerdo al tipo de problemas es posible que no se tome en cuenta alguno de estos elementos y que se enfoque sólo en uno o algunos de ellos.

El primer elemento de este marco conceptual es la generación de células que en el Sistema Inmune Artificial corresponde a la generación de soluciones del problema.

CAPÍTULO V.- METODOLOGÍA PROPUESTA

Para la generación de las soluciones se pueden considerar varios puntos importantes y considerar uno o más de ellos.

El primero es: la generación de soluciones ¿va a ser aleatoria o va a generarse a partir de una librería genética?

El segundo: ¿cuál será el estímulo para generar nuevas soluciones y cada cuanto se enviará esa señal?

El tercero: ¿cuántas tipos de células se van a generar y cuál va a ser su función?

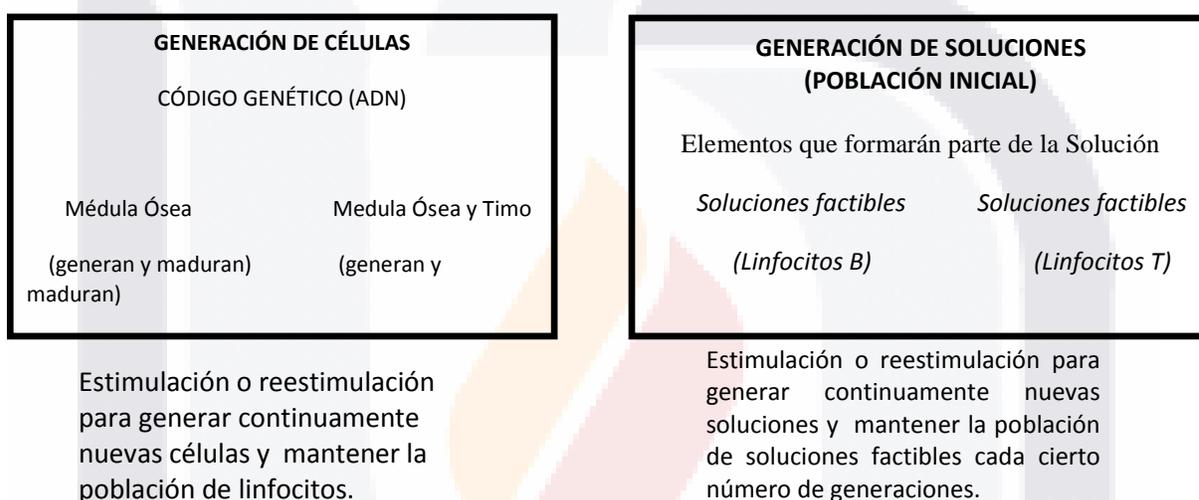


Figura 16: Correspondencia del funcionamiento del SIN hacia el SIA. Esta es la etapa inicial en la que se considera la formación de las células (Linfocitos T y B).

Mecanismos para la maduración de los linfocitos

En la fase de la generación celular, se llevan a cabo ciertos procesos que permiten que los linfocitos maduren, hasta el punto de poder reconocer a un antígeno, es decir que tienen una determinada afinidad.

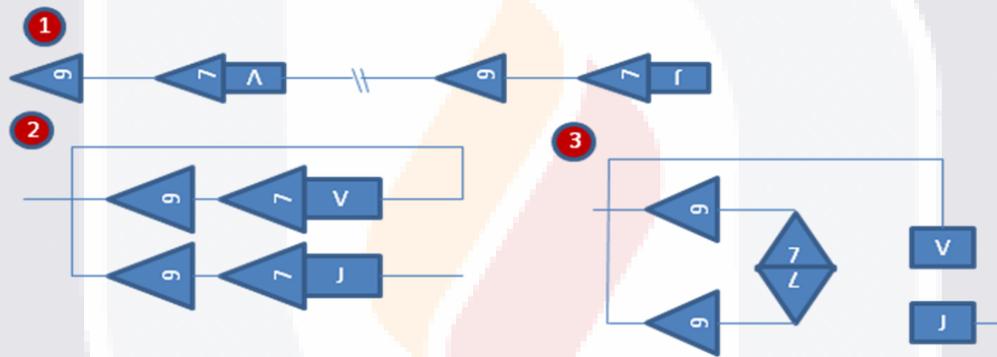
CAPÍTULO V.- METODOLOGÍA PROPUESTA

Podemos decir que el paso que con el que el SIN inicia es con la generación de linfocitos T y B en médula ósea y timo (entre otras células), de los cuales sus receptores tendrán afinidad para un sólo antígeno, teniendo en cuenta que un antígeno puede activar n linfocitos. Estas células tienen un proceso de maduración en el cual sufren cambios a nivel genético para poder formar parte del SIN.

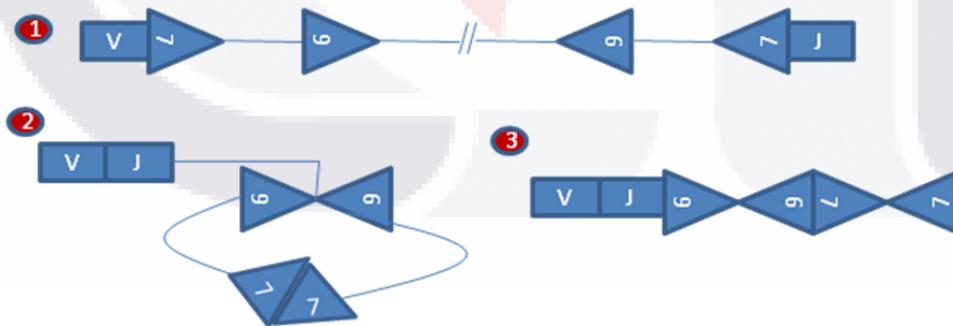
El proceso se llama “Recombinación” de los segmentos V(D)J y se lleva a cabo por medio de 2 mecanismos: Inversión y Borrado, que a continuación se mencionan.

Recombinación V(D)J

Inversión



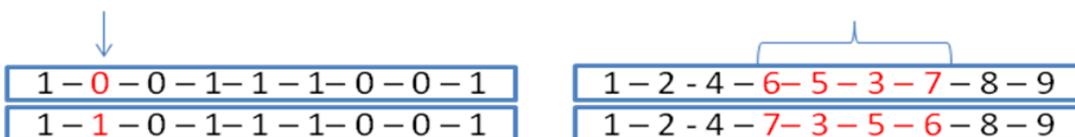
Borrado



CAPÍTULO V.- METODOLOGÍA PROPUESTA

A continuación se simula como funcionan los 2 operadores

Inversión



El siguiente es el pseudocódigo con el que se trata de modelar el funcionamiento de estos mecanismos.

1. Generar la población inicial de linfocitos
Generar cada uno de los linfocitos que formaran parte de la solución
Evaluar la afinidad
2. Definir que solución es la J, la V el 7 y el 9.
3. Fin Generación de la población
4. Inicio proceso de maduración
5. Aplicar el operador de Inversión
6. Evaluar su afinidad
7. Si es mejor ya no es necesario el operador de borrado
8. Si la mejora no es sustancial (menos del 10%) pasa al operador de borrado sólo la solución modificada.
9. Si no es mejor pasan al operador de Borrado la solución original y la modificada por el operador de Inversión
10. Aplicar el operador de Borrado
11. Evaluar su afinidad
12. Si las dos soluciones son mejoradas pasan a la población. Si sólo una es mejorada esa es la que pasa.
13. Si no es mejor se deja como estaba y pasa la solución original
14. Fin del proceso de maduración.

Representación de las soluciones

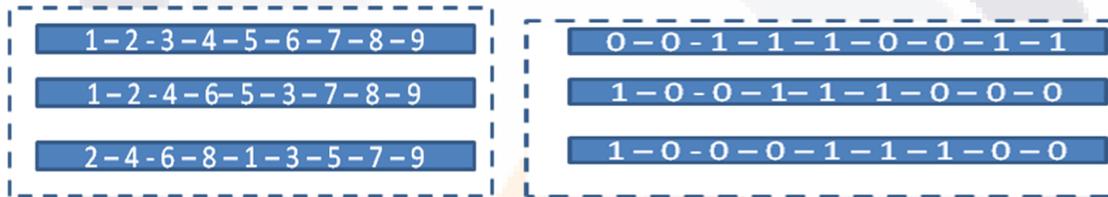
Las soluciones que se generan en un Sistema Inmune Artificial poblacional pueden ser consideradas como linfocitos o como los receptores (anticuerpos) o algún otro elemento que los conforme.

Si se considera una sola cadena, este puede ser considerado como un anticuerpo.

Si el linfocito es considerado como la solución, este se forma por más de una solución.

Cada cadena puede ser considerada una subsolución o solución del problema.

Ejemplo:



El segundo elemento que se tiene es el proceso de selección positiva y negativa procesos en los cuales el Sistema Inmune Natural logra identificar los linfocitos o células que podrían ser autoreactivas y dañar el organismo. Los linfocitos que se detectan como autoreactivos son eliminados (selección positiva) los que no continúan vivos (selección negativa).

Para usar modelos basados en estas teorías se deberá analizar si es necesario un proceso de este tipo para la solución del problema. Además se debe definir si, en el caso de que se utilice, cuales van a ser los criterios para llevar a cabo una selección positiva y una negativa.

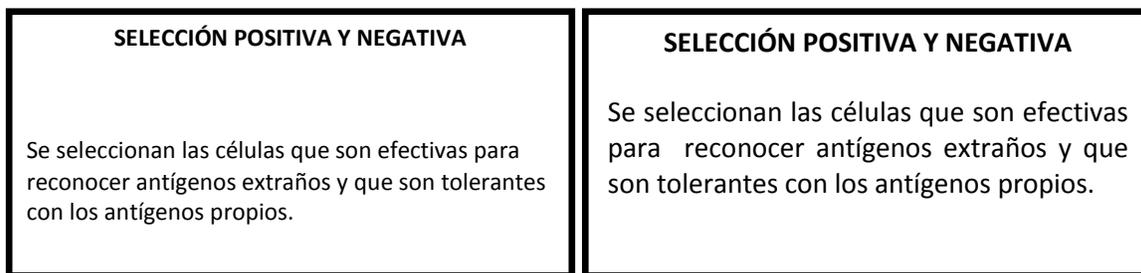


Figura 17: Correspondencia del funcionamiento del SIN hacia el SIA. Esta es la etapa en la que se consideran maduros los linfocitos (buenas soluciones) y se decide forman parte del SIN.

El tercer elemento es en el cual se lleva a cabo la respuesta inmunitaria, la cual consta de 5 fases y que internamente tiene un proceso de adaptación. En este proceso se busca mejorar o aumentar la eficacia de las células que están activas y que sean mejores para una segunda respuesta inmunitaria a un mismo antígeno.

La selección Clonal es el principio central de la inmunidad adaptativa (Janeway, 2001) y es el que principalmente se modela para aplicarlo a la solución de problemas.

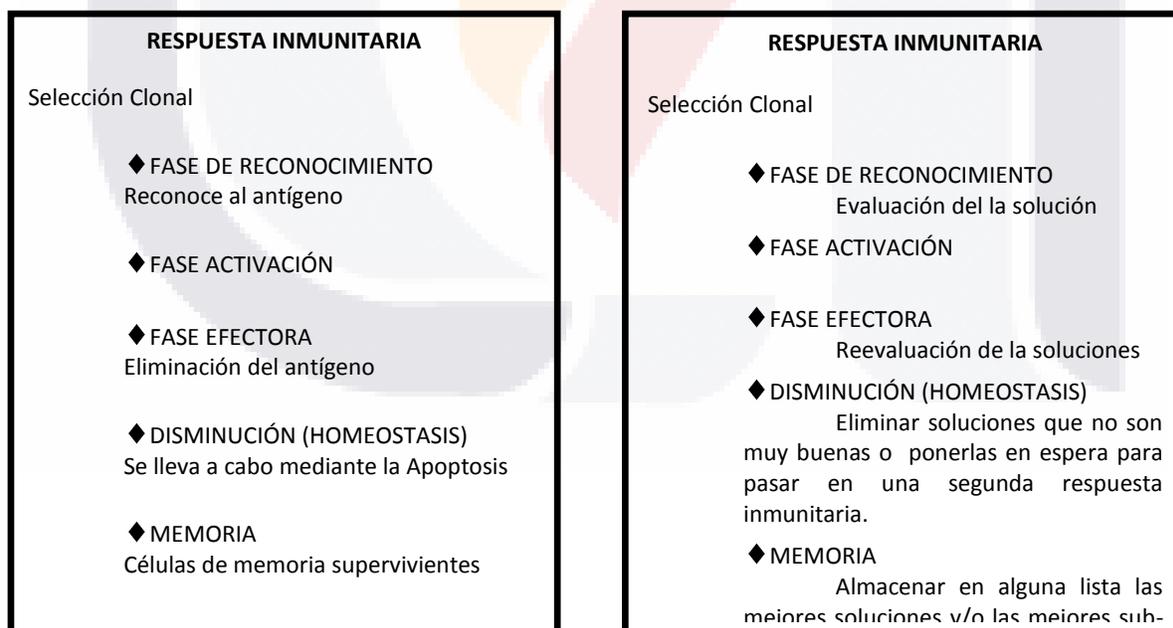


Figura 18: Correspondencia del funcionamiento del SIN hacia el SIA. Esta es la etapa en que se activan los diferentes mecanismos para atacar al antígeno.

Algoritmo Propuesto basado en la selección clonal (CLONALG)

1. Generar aleatoriamente población inicial de anticuerpos (**pobac**) de tamaño **N**.
2. Calcular o medir la afinidad de cada anticuerpo mediante una función.
3. Seleccionar **n** anticuerpos con el menor costo.
4. Generar **Nc** Clones de los mejores anticuerpos (**pobclon**).
5. Maduración de la Afinidad
 1. Dividir en **m** partes la solución. $P_{div}=0.5$ o 0.3 o 0.25 (2 partes, 3 partes o 4 partes)
 2. Evaluar la afinidad de cada parte. La menor es la parte en que se aplica la hipermutación somática.
 3. Hipermutación : algún método de búsqueda local o de manera aleatoria.
6. Generar nuevos **Nnac** anticuerpos aleatoriamente (**pobnac**).
7. Unir la población de anticuerpos, clones, y nuevos anticuerpos.
8. Regresar al paso 2 y repetir por el número de respuestas inmunitarias dadas por el usuario.

A continuación se muestra un diagrama del pseudocódigo propuesto:

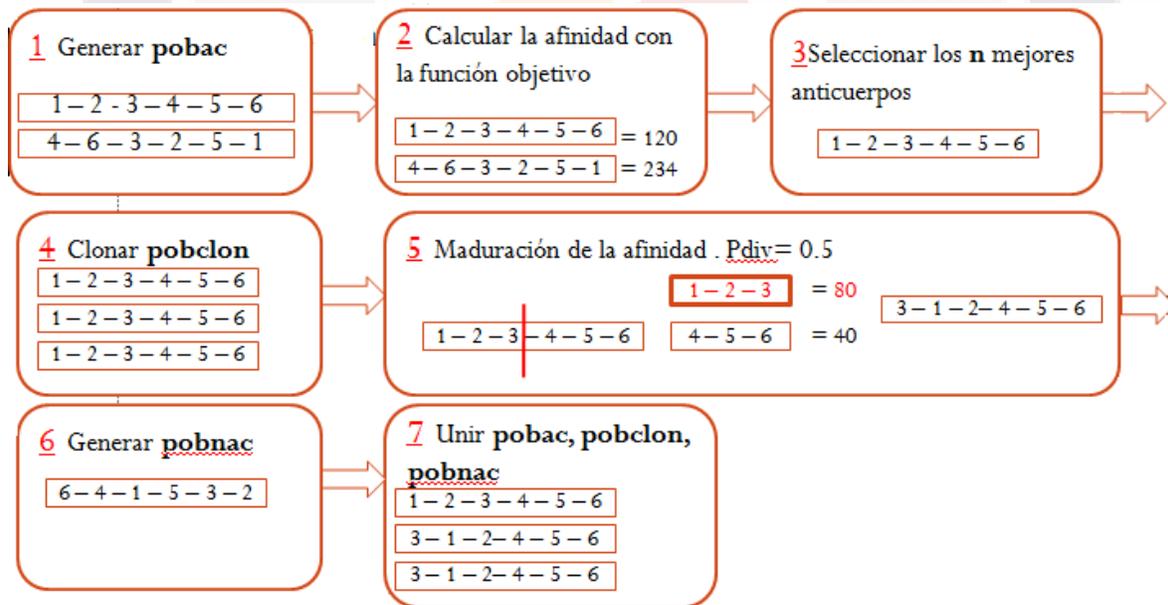


Figura 19. Diagrama del funcionamiento del algoritmo que se propone, basado en el CLONALG

Cada elemento del Sistema Inmune Natural tiene su correspondencia en el Sistema Inmune Artificial o al menos eso es lo que se intenta. La siguiente tabla muestra dicha correspondencia de acuerdo al algoritmo propuesto.

Tabla 9: Mapeo del Sistema Inmune Natural con el Sistema Inmune Artificial

Sistema Inmune Artificial	Sistema Inmune Natural
pobac, pobclon, pobnac (soluciones)	Conjunto de Anticuerpos con determinada afinidad
Solución Óptima	Antígeno
Costo de la solución (evaluada con la función objetivo)	Afinidad (Especificidad/similitud con determinado antígeno)
Cambio en el recorrido	Maduración de la Afinidad
Procedimiento par modificar el recorrido (optimizador local o aleatoriamente)	Hipermutación somática

El cuarto elemento encierra funciones que se lleva a cabo el Sistema Inmune, independientemente de que haya o no un antígeno externo dañino, esto para detectar que todo el organismo este en buen estado o de lo contrario reaccionar de manera expedita.

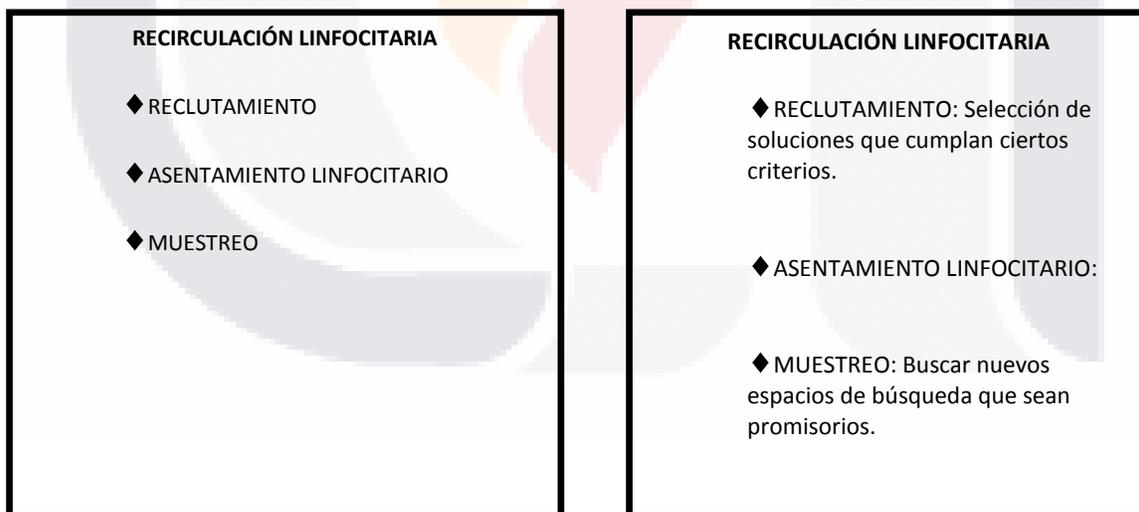


Figura 20: Correspondencia del funcionamiento del SIN hacia el SIA. Esta son funciones que se llevan a cabo constantemente independientemente de si se ha activado o no una respuesta inmunitaria.

3 Ingeniería

En esta fase

- Implementar en el lenguaje de programación deseado el algoritmo. Pasar del pseudocódigo a la programación en computadora.
- Probar el funcionamiento del programa.

4 Experimentación

En esta fase se debe definir claramente el tipo de experimento que se desea realizar, esto con la finalidad de calibrar los parámetros que inciden directamente en el funcionamiento del algoritmo, además de dar una medida de la confiabilidad

- Diseñar el experimento para determinar la confiabilidad y eficiencia del algoritmo para solucionar el problema.

5.3 Relación entre los Elementos del SIA, SIN y el Problema

En la Tabla 10 se muestran los elementos del SIN, que se han considerado relevantes para esta investigación, y de que modo se pueden representar o la equivalencia tanto en el problema como en el SIA.

CAPÍTULO V.- METODOLOGÍA PROPUESTA

Tabla 10. Correspondencias entre el AIS, el problema a resolver y el SIN

SIN	PROBLEMA	SIA
❖ Organismo (Individuo)	➤ Variables	<input type="checkbox"/> Soluciones Factibles
❖ Linfocitos B	➤ Función Objetivo	<input type="checkbox"/> Población Inicial
❖ Linfocitos T	➤ Espacio de búsqueda	<input type="checkbox"/> Selección clonal
❖ Células accesorias	➤ Soluciones factibles	<input type="checkbox"/> Función de similitud
❖ Anticuerpos	➤ Representación	<input type="checkbox"/> Representación
❖ Afinidad	➤ Restricciones	<input type="checkbox"/> Fitness (Afinidad)

Conclusiones

Se concluye que los trabajos anteriormente analizados influyen el desarrollo de esta tesis, aunque no de manera explícita, ya que los objetivos de esta tesis no son probar si funcionan o no los marcos de trabajo y las metodologías propuestas por otros investigadores. Sin embargo, si aportaron puntos clave en los que reforzar la metodología propuesta y que no habían sido profundamente estudiados y que son nichos propicios para la investigación, así como en la clarificación de la ubicación del tema. Además es posible hacer un comparativo de lo que en esta tesis se propone y los resultados obtenidos en los trabajos anteriormente realizados.



Capítulo 6

Experimentación

En este capítulo se muestran los resultados obtenidos del diseño de experimento factorial realizado para probar el desempeño del algoritmo propuesto basado en el CLONALG, así como el desarrollo de la metodología definida en el capítulo anterior tomando como ejemplo el problema del agente viajero.

6.1. Desarrollo de la Metodología

De acuerdo a las fases definidas para la metodología, los elementos de la **FASE 1** para el problema del TSP se definen de la siguiente manera:

- Modelar/Plantear el problema

CAPÍTULO VI.- EXPERIMENTACIÓN

- Visitar n número de ciudades sin repetir ninguna, iniciando y terminando en la misma ciudad. Generar un ciclo hamiltoniano.
- Entradas
 - Número de ciudades o sitios a visitar (Tamaño). Ej. Ciudades=30.
 - Distancia entre ciudades
- Salidas
 - Solución: Recorrido
 - Costo de la solución: Costo/distancia del recorrido
 - Probable mente el tiempo.
- Tipos de variables: Discretas, Continuas
 - Para el TSP son discretas
- Restricciones
 - No visitar más de una vez cada ciudad
 - Iniciar y terminar el recorrido en la misma ciudad
- Alfabeto
 - Números reales
 - Cardinalidad= número de ciudades
- Representación de la solución
 - Cadena de número enteros que representan las ciudades.
- Medida de afinidad
 - Suma de las distancias del recorrido (función objetivo)

1 - 2 - 4 - 6 - 5 - 3 - 7 - 8 - 9

$$\min z(x) = \sum_{(i,j) \in A} c_{ij} x_{ij}$$

En la **FASE 2** lo que se busca es definir el modelo del SIN a utilizar. En este caso se toma el modelo de la Selección Clonal, del cuál se muestra el pseudocódigo.

1. Generar aleatoriamente población inicial de anticuerpos (**pobac**) de tamaño **N**.
2. Calcular o medir la afinidad de cada anticuerpo mediante una función.
3. Seleccionar **n** anticuerpos con el menor costo.
4. Generar **Nc** Clones de los mejores anticuerpos (**pobclon**).

5. Maduración de la Afinidad

1. Dividir en **m** partes la solución. Pdiv=0.5 o 0.3 o 0.25 (2 partes, 3 partes o 4 partes)
 2. Evaluar la afinidad de cada parte. La menor es la parte en que se aplica la hipermutación somática.
 3. Hipermutación : aleatoria.
6. Generar nuevos **Nnac** anticuerpos aleatoriamente (**pobnac**).
 7. Unir la población de anticuerpos, clones, y nuevos anticuerpos.
 8. Regresar al paso 2 y repetir por el número de respuestas inmunitarias dadas por el usuario.

Para la **FASE 3** se definen principalmente en que lenguaje se va a programar. Para este problema se eligió C++ Builder con programación orientada a objetos.

6.2. Análisis de los experimentos con el TSP Simétrico

Los experimentos (**FASE 4**) se realizaron con los Benchmarks de la TSPLIB (<http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/>). Para este caso se utilizaron los benchmarks del TSP simétrico.

En cada experimento realizado se definieron 3 niveles y 10 réplicas para cada combinación de parámetros. Se obtuvieron un total de 270 observaciones (Niveles parámetros * Num Réplicas).

Los parámetros que tomaron 3 diferentes valores fueron: Generaciones, Población Inicial y Punto de división. A continuación se muestra la tabla con los diferentes valores de cada parámetro que se utilizaron para las corridas.

CAPÍTULO VI.- EXPERIMENTACIÓN

Tabla 11: Valores de cada parámetro para el experimento con el benchmark Burman 14.

Burman 14						
Generaciones	Pob. Inicial	Clones	Nuevos	% Selección	P. Divi	B
40	20	1	1	50	25	0
80	40	1	1	50	30	0
120	60	1	1	50	50	0

El benchmark Burman 14, es un problema en el que se consideran 14 ciudades, cada una con sus coordenadas (latitud y longitud). Las distancias entre cada ciudad se calcula con la función de distancia geográfica, para así generar la matriz de costos.

En la **¡Error! No se encuentra el origen de la referencia.** se muestran los resultados del experimento.

Estos resultados fueron analizados estadísticamente con la ayuda del programa SPSS 10. Se hizo un análisis el cuál, en este primer experimento, el análisis arrojó los siguientes resultados.

Dependent Variable: COSTO

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	7740958.928 ^a	26	297729.190	4.470	.000
Intercept	4944775729	1	4944775729	74234.618	.000
P_DIVI	299990.614	2	149995.307	2.252	.107
POB_INI	5624301.196	2	2812150.598	42.218	.000
GENERACI	936397.658	2	468198.829	7.029	.001
P_DIVI * POB_INI	202455.349	4	50613.837	.760	.552
P_DIVI * GENERACI	75356.593	4	18839.148	.283	.889
POB_INI * GENERACI	210473.549	4	52618.387	.790	.533
P_DIVI * POB_INI * GENERACI	391983.967	8	48997.996	.736	.660
Error	16186255.7	243	66610.106		
Total	4968702943	270			
Corrected Total	23927214.6	269			

CAPÍTULO VI.- EXPERIMENTACIÓN

Para este análisis se tomo como variable dependiente el Costo. Y los parámetros que tomaron diferentes valores fueron Punto de división, Población Inicial y Generaciones. Dado los valores de significancia obtenidos, se concluye que tanto las generaciones como la población inicial son significativos en el valor del Costo que se obtiene.

Las siguientes tablas nos muestran en que valor de cada parámetro es donde se encuentra un mejor costo.

Para el punto de división (P_DIVI) es 25, para la población inicial es 60 y para las generaciones es 80.

1. P_DIVI

Dependent Variable: COSTO

P_DIVI	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
25	4233.384	27.205	4179.796	4286.972
30	4311.060	27.205	4257.473	4364.648
50	4294.008	27.205	4240.420	4347.596

2. Pob. Inicial

Dependent Variable: COSTO

Pob. Inicial	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
20	4477.307	27.205	4423.719	4530.895
40	4224.112	27.205	4170.524	4277.700
60	4137.033	27.205	4083.446	4190.621

3. Generaciones

Dependent Variable: COSTO

Generaciones	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
40	4362.733	27.205	4309.145	4416.321
80	4235.760	27.205	4182.173	4289.348
120	4239.959	27.205	4186.371	4293.546

CAPÍTULO VI.- EXPERIMENTACIÓN

El benchmark Bayg 29, es un problema en el que se consideran 29 ciudades, cada una con sus coordenadas (latitud y longitud). Las distancias entre cada ciudad se calcula con la función de distancia geográfica (geodésica), para así generar la matriz de costos.

Tabla 12 Valores de cada parámetro para el experimento con el benchmark Baygs 29.

Bayg 29						
Generaciones	Pob. Inicial	Clones	Nuevos	% Selección	P. División	B
40	20	1	1	50	25	0
80	40	1	1	50	30	0
120	60	1	1	50	50	0

Dependent Variable: COSTO

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2.348E+10 ^a	26	903183306.0	10.884	.000
Intercept	1.029E+13	1	1.029E+13	124043.1	.000
PDIVI	1.106E+10	2	5529660753	66.634	.000
POB._INI	6247090620	2	3123545310	37.640	.000
GENERACI	4107588064	2	2053794032	24.749	.000
PDIVI * POB._INI	451677450	4	112919362.5	1.361	.248
PDIVI * GENERACI	833160493	4	208290123.2	2.510	.043
POB._INI * GENERACI	151794648	4	37948662.04	.457	.767
PDIVI * POB._INI * GENERACI	632133175	8	79016646.86	.952	.474
Error	2.017E+10	243	82985570.59		
Total	1.034E+13	270			
Corrected Total	4.365E+10	269			

a. R Squared = .538 (Adjusted R Squared = .489)

Para este análisis se tomo como variable dependiente el Costo. Y los parámetros que tomaron diferentes valores fueron Punto de división, Población Inicial y Generaciones.

CAPÍTULO VI.- EXPERIMENTACIÓN

Dado los valores de significancia obtenidos, se concluye que tanto las generaciones como la población inicial y el punto de división son significativos en el valor del Costo que se obtiene.

Las siguientes tablas nos muestran en que valor de cada parámetro es donde se encuentra un mejor costo.

Para el punto de división (P_DIVI) es 50, para la población inicial es 60 y para las generaciones es 120.

1. PDMI

Dependent Variable: COSTO

PDMI	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
25	200295.7	960.241	198404.233	202187.147
30	199248.3	960.241	197356.865	201139.780
50	186225.8	960.241	184334.358	188117.272

2. Pob. Inicial

Dependent Variable: COSTO

Pob. Inicial	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
20	202059.0	960.241	200167.551	203950.466
40	191895.1	960.241	190003.690	193786.604
60	191815.7	960.241	189924.215	193707.129

3. Generaciones

Dependent Variable: COSTO

Generaciones	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
40	200300.9	960.241	198409.410	202192.325
80	194667.5	960.241	192776.033	196558.948
120	190801.5	960.241	188910.013	192692.927

El benchmark Bayg 29, es un problema en el que se consideran 29 ciudades, cada una con sus coordenadas (latitud y longitud). Las distancias entre cada ciudad se calcula con la función de distancia geográfica (geodésica), para así generar la matriz de costos.



Capítulo 7

Caso de Estudio

Actualmente los ataques y las intrusiones a las redes de computadoras son eventos muy comunes que causan perjuicios a la integridad y seguridad de la información. Existen herramientas para evitar estos ataques pero no logran su objetivo en 100% ya que ninguna herramienta es perfecta además que constantemente se generan nuevos tipos de intrusiones para evadir la seguridad que exista en el momento.

7.1 Sistemas de detección de intrusos

Estos sistema son los encargados de detectar e identificar ataquen en una red de computadoras o en un sistema computacional. Los Sistemas de Detección de Intrusos

CAPÍTULO VII.- CASO DE ESTUDIO

(IDS, por sus sigla en inglés) realizan estas actividades recolectando información acerca del tráfico de la red, datos de un nodo en particular o ambos.

Hoy en día empresas, instituciones, universidades y particulares están presentes en Internet de forma que cualquier usuario del planeta puede acceder a información, compra-venta, transacción o entretenimiento a través de la red. Algunos datos que se manejan deben mantenerse a salvo de cualquier intrusión de modo que la seguridad en este tipo de actividades tiene una importancia crítica. Se puede definir intrusión como cualquier conjunto de acciones que tratan de comprometer la integridad, confidencialidad o disponibilidad de un recurso (Heady et al., 1990). El NIST (National Institute of Standards and Technology) define la detección de intrusiones como el proceso de monitorización de eventos que suceden en un sistema informático y el análisis de dichos eventos en busca de signos de intrusiones.

Las principales categorías de los SDI de acuerdo a la forma en como detectan un ataque son: basados en firmas, basados en anomalías, y los híbridos (Axelsson, 2000). Un sistema basado en firmas, usa conocimiento sobre el ataque, tal como el patrón o firma de un ataque, para determinar cuando está ocurriendo un ataque. Si el sistema no reconoce un patrón de ataque, entonces asume que el dato es aceptable (Axelsson, 2000). La ventaja principal de este tipo de sistemas es que un ataque que no está en la base de conocimiento no es detectado. La segunda técnica es conocida como basada en anomalías. Esta técnica usa un modelo de buenos comportamientos conocidos y entonces detecta desviaciones de este modelo. Cualquier comportamiento que no empata con el modelo se asume que es una intrusión (Axelsson, 2000). Así, el sistema puede detectar nuevos ataques porque no está dependiendo de una base de conocimiento de patrones de ataque conocidos; si no, en una base de conocimiento de buenos comportamientos (Axelsson, 2000). Para este tipo de sistemas, el modelo de buenos comportamientos conocidos debe ser exacto o el sistema genera demasiadas alertas falsas de detección.

CAPÍTULO VII.- CASO DE ESTUDIO

Algunos sistemas híbridos usan ambas, la base de conocimientos de ataques y la información basada en un modelo parcial de tráfico de la red “normal” o “bueno”.

Otra clasificación descrita por Zurutuza (Zurutuza et al., 2005) según tres aspectos principales: fuente de datos, tipo de análisis y mecanismo de respuesta.

Las fuentes de datos se refieren a los datos que maneja el sistema para analizar las posibles intrusiones. Se diferencian tres fuentes principales: basadas en los registros de auditoría que generan las máquinas (host-based), basadas en las aplicaciones informáticas (applicationbased), y basadas en la información que circula en las redes telemáticas (network-based). De este modo, los tipos de datos que deben procesar los IDS pueden ser paquetes de red, llamadas al sistema de procesos en ejecución, registros de auditoría, logs producidos por diferentes aplicaciones, o incluso comandos del teclado. Los tipos de análisis de los datos recogidos, cuyos paradigmas principales son el de la detección de usos indebidos (misuse detection) que compara la información recogida con descripciones (o firmas) de ataques conocidos, y la detección de anomalías (anomaly detection), que haciendo uso de datos históricos sobre la actividad de un sistema y/o de especificaciones sobre el comportamiento deseado de usuarios y aplicaciones, construyen un perfil que representa la operación normal del sistema monitorizado, e identifica patrones de actividades que se desvían del perfil definido. El mecanismo de respuesta. Se pueden diferenciar los sistemas de respuesta pasiva (passive response), que en lugar de tomar acciones, se limitan a generar la alerta correspondiente, y por otro lado, los sistemas de respuestas activas (active response), que además de generar las alertas correspondientes, reaccionan modificando el entorno (ya sea generando nuevas reglas en el cortafuegos, o cortando la conexión correspondiente, entre otras posibilidades). El primer modelo de detección de anomalías fue el llamado IDES (Intrusion Detection Expert System), propuesto por Dorothy Denning (Denning et al., 1987), con la idea básica de monitorizar las operaciones estándares de un sistema objetivo, observando desviaciones en su uso. Su artículo provee un enmarque metodológico que más tarde inspiraría a muchos investigadores. Este se considera el punto de partida de nuevas

investigaciones aplicando otras técnicas para solucionar el problema de la detección de intrusos.

A continuación se muestra un análisis de algunas técnicas que se han utilizado para crear IDSs

7.3. Técnicas de Data mining para detección de Intrusos

El data mining, también conocido por KDD (Knowledge Discovery in Databases), se refiere al proceso no trivial de identificar patrones o modelos válidos, nuevos, potencialmente útiles, y comprensibles a partir de datos (Fayyad et al., 1996). Los modelos de data mining se componen de dos factores relevantes: su función (como la de clasificar, agrupar, resumir,...), y el modo de representar el conocimiento (como una función lineal de múltiples variables, en modo de árbol, de reglas, una red,...).

7.3.1. Funciones de modelos

Las funciones de los modelos más comunes en data mining utilizados para la detección de intrusiones incluyen:

- Clasificación: clasifica un caso entre varias clases o categorías predefinidas (intrusión o no intrusión; o entre tipos de intrusión).
- Clustering: clasifica un caso en una de las clases o agrupaciones en las que las clases se deben determinar a partir de los propios datos. Los clusters o grupos se definen en base a medidas de similitud, densidad de probabilidad o distancia entre casos.
- Modelado de dependencias: describe las dependencias significativas entre variables. El análisis de relaciones (como reglas de asociación), que determina relaciones existentes entre elementos de una base de datos, podría considerarse un caso particular de modelado de dependencias.

7.3.2. Representación del modelo

La representación del modelo determina tanto la flexibilidad del modelo en cuanto a la representación de los datos, como la facilidad de interpretación del modelo en términos humanos. Modelos de representación incluyen reglas y árboles de decisión, modelos no-lineales (como redes neuronales), métodos basados en ejemplos (como el vecino más cercano y métodos de razonamiento basado en casos), modelos de dependencias de probabilidades (como redes Bayesianas) y modelos de atributos relacionales.

7.4. Data mining y los IDS

Son diversos los trabajos que han utilizado este tipo de técnicas en el desarrollo de IDSs. En la Figura 21 se muestra una clasificación de IDSs en función de la estrategia de análisis, entre los que se encuentran aquellos relacionados con el aprendizaje automático. Como se puede observar, algunos IDS derivan de otros (representado mediante una línea de puntos y flecha de dirección), y otros en cambio son fruto de la aplicación de diferentes técnicas y estrategias de análisis. Aspectos como la capacidad explicativa, complejidad, coste computacional y capacidad de aprendizaje incremental de los modelos son determinantes a la hora de seleccionar entre unos y otros.

Desde 1997, investigadores de la universidad de Columbia, liderados por el profesor Sal Stolfo, han realizado numerosos trabajos en detección de intrusiones basados en técnicas de data mining. Lee y sus colaboradores (Lee et al., 1999) proponen un marco general para la construcción de IDSs. Con este trabajo se dio pie a un gran número de investigaciones en lo que se refiere a la aplicación de data mining a sistemas de detección de intrusiones.

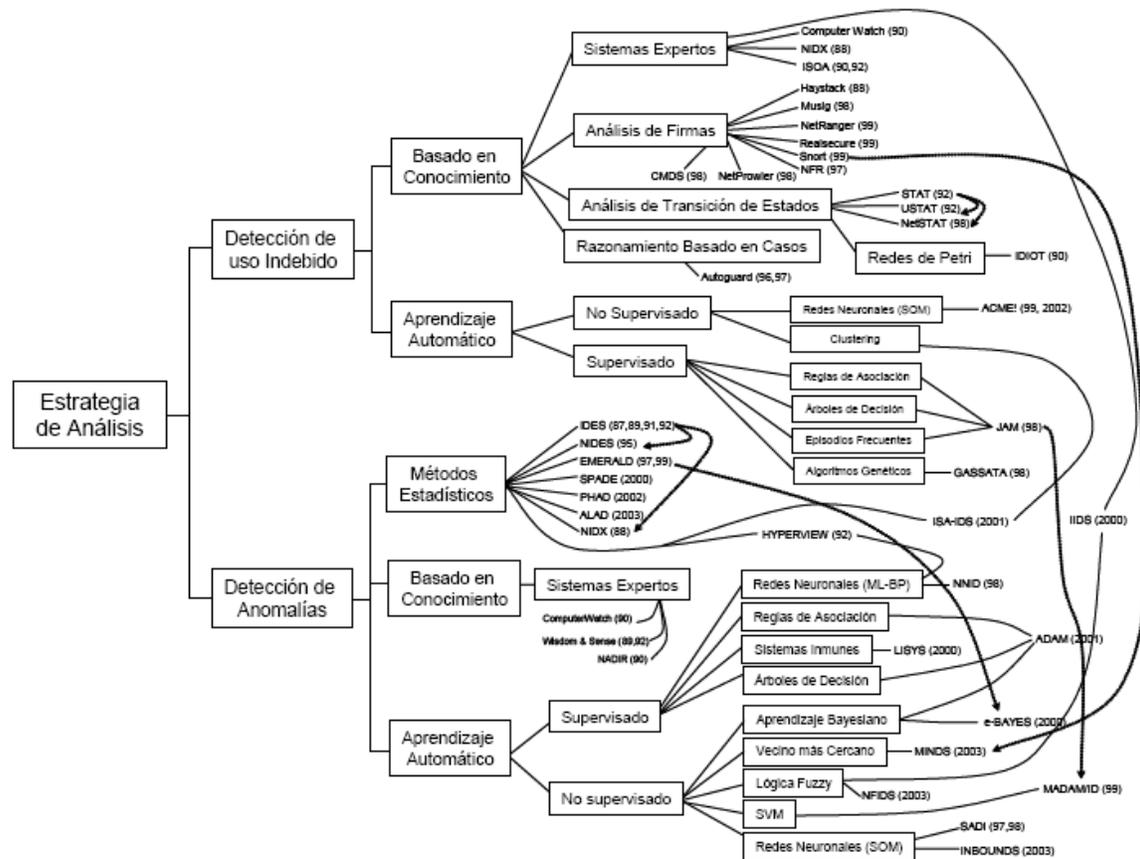


Figura 21 Estrategia de Análisis de los diferentes IDS, (Zurutuza et al., 2005)

A continuación se hace una breve revisión de los trabajos de esta área, clasificados por el paradigma de aprendizaje que se utiliza.

7.4.1. Reglas de asociación

Las reglas de asociación han sido ampliamente utilizadas para la detección de intrusiones. Su objetivo es el de derivar correlaciones de varias características a partir de datos de ejemplo, dada la evidencia de que frecuentemente los datos obtenidos de ejecuciones de programas y actividades de usuarios muestran correlaciones. Mediante las reglas de asociación se pueden encontrar relaciones internas entre datos de una misma conexión. Lee y Stolfo (Lee et al., 1998) presentan sus algoritmos de reglas de asociación y episodios frecuentes (para descubrir un conjunto de eventos que

ocurrían en una ventana temporal) utilizados para calcular un conjunto de patrones a partir de datos de auditoría. Daniel Barbará y colaboradores (Barbará et al., 2001) también hacen uso de esta técnica en su proyecto ADAM pero como método de detección de anomalías. ADAM utiliza esta técnica para encontrar eventos anómalos en conexiones de tráfico de red para luego clasificar dichos eventos o reglas de asociación inesperadas en instancias normales o anómalas. MINDS de la Universidad de Minnesota, en el cuál trabajó Ertoz y su equipo de trabajo (Ertoz et al., 2004), también utiliza reglas de asociación para caracterizar las anomalías detectadas, y ayudar en la creación de nuevas firmas y modelos de ataques emergentes.

7.4.2. Inducción de reglas

Teng y colaboradores (Teng et al., 1990) propusieron un método para inducir patrones secuenciales temporales en una secuencia de eventos. De este modo se podían aprender patrones secuenciales para detectar intrusiones. La herramienta RIPPER, ha sido utilizada para el aprendizaje o inducción de conjuntos de reglas sobre un conjunto de datos previamente etiquetados (normal/intrusión). El sistema se compone de un conjunto de reglas de asociación y patrones de episodios frecuentes que pueden ser aplicados tanto a eventos de seguridad como a conjuntos de datos de tráfico de red. Lee y Stolfo (Lee et al., 1998) fueron los que propusieron el uso de RIPPER para la detección de intrusiones con éxito en su proyecto JAM.

7.4.3. Árboles de decisión

B. Pfahringer (Pfahringer, 2000), utilizó como solución boosted trees basado en el algoritmo C5 para construir conjuntos de clasificadores y combinarlos. I. Levin (Levin, 2000) utilizó la herramienta Kernel Miner creó un conjunto de árboles de decisión localmente óptimos (llamados bosque de decisión), de donde se seleccionaban aquellos óptimos para predecir nuevas clases (subbosque). Nong Ye (Nong et al., 2000), de Arizona State University, realizaron una propuesta de árboles de decisión

como método de aprendizaje de firmas de ataques. En el trabajo de Kruegel (Kruegel et al., 2002) se hace uso de árboles de decisión como método de optimización de IDS basados en análisis de firmas. Construyen una variante del algoritmo ID3 a partir de las firmas de Snort 2.0 y demuestran que mejora el proceso de detección.

7.4.4. Clasificadores Bayesianos

El resultado de la aplicación de sistemas Bayesianos en IDS se presenta en forma de relaciones de probabilidad condicional en lugar de reglas o firmas. Las redes Bayesianas son herramientas poderosas como modelos de decisión y razonamiento bajo incertidumbre. S. Axelsson (Axelsson, 1999) publicó un artículo que utilizaba la regla Bayesiana de probabilidad condicional para apuntar las implicaciones de la falsedad de la tasa básica (base-rate fallacy) en detección de intrusiones. El grupo de investigación, en el cuál participó A. Valdés (Valdés et al., 2000), del SRI desarrolló un módulo para su IDS EMERALD llamado eBayes TCP que utilizaba tecnología de redes Bayesianas para analizar, lo que ellos llaman, “explosiones de tráfico”.

En el trabajo realizado por Barbará y su grupo de investigación (Barbará et al., 2001) hacen uso de la teoría Bayesiana para su sistema ADAM (Audit Data Analysis and Mining). En dicho trabajo, proponen el uso de estimadores pseudo-Bayes para afinar la capacidad detectar anomalías, reduciendo a su vez el número de falsas alarmas. Tras aplicar la técnica de pseudo-Bayes, construyen un clasificador naive Bayes para clasificar los ejemplos en normales, ataques conocidos y ataques nuevos.

7.4.5. Redes Neuronales

Se han realizado numerosos trabajos con redes neuronales artificiales en detección de intrusiones gracias a su flexibilidad y adaptación a los cambios naturales que se pueden dar en el entorno y, sobre todo, a la capacidad de detectar instancias de los ataques conocidos. La mayor deficiencia que presentan las redes neuronales es que

son un modelo no descriptivo, sin que se pueda conocer la razón de la decisión tomada. El primer modelo de detección de intrusiones basado en redes neuronales fue realizado por Fox (Fox et al., 1990) como método para crear perfiles de comportamiento de usuarios. Al igual que Debar (Debar et al., 1992), ellos utilizan redes neuronales para predecir el siguiente comando basado en una secuencia de comandos previos ejecutados por un usuario. El aprendizaje lo realizan mediante redes neuronales recurrentes por lo que la red está continuamente observando y tiene la capacidad de “olvidar” comportamientos antiguos. Más tarde, investigadores de Brazil (Cansian et al., 1997), desarrollan el sistema SADI (Sistema Adaptativo de Detecção de Intrusão), que usan la red neuronal para identificar el comportamiento intrusivo de patrones capturados de una red. En el trabajo de Höglund (Höglund et al., 1998) del centro de investigación de Nokia, en Finlandia, propusieron el uso de dos mapas autoorganizativos. Uno de ellos se utilizó como método de visualización del comportamiento de los usuarios, y el otro compara el comportamiento de un usuario con el perfil generado en la red anterior ayudado con medidas estadísticas. En la Universidad de Ohio se ha desarrollado un IDS de red (Ramadas et al., 2003) llamado INBOUNDS (Integrated Network-Based Ohio University Network Detective Service) donde un módulo de detección de anomalías basado en análisis estadístico se ha sustituido por otro que utiliza mapas autoorganizativos.

7.4.6. Support Vector Machines (SVM)

SVM es una técnica que se ha utilizado como aprendizaje supervisado, pero tras alguna modificación ha sido utilizado en datos sin un conocimiento a priori. (Eskin et al., 2002) utilizaron una SVM como complemento a sus métodos de clustering para el aprendizaje no supervisado.

(Mukkamala et al., 2002) utilizaron cinco SVM, uno para identificar tráfico normal, y el resto para identificar cada uno de los cuatro tipos de ataques representados en el

conjunto de datos de KDD Cup 99. Los compararon con redes neuronales llegando a la conclusión de que los SVM demostraban mejor desempeño.

7.4.7. Modelos de Markov

Los modelos de Markov son un tipo de aprendizaje basado en secuencias. Se pueden distinguir, entre otros, las cadenas de Markov y los modelos ocultos de Markov (HMM ó Hidden Markov Models). Una cadena de Markov es una secuencia de eventos, donde la probabilidad del resultado de un evento depende sólo del resultado del evento anterior. De esa misma manera, los HMM son una técnica probabilística para el estudio de series en el tiempo. (T. Lane et al., 1999) investigó el uso de modelos ocultos de Markov para crear perfiles de usuarios y medidas de similaridad. En esa misma fecha, C. Warrender (Warrender et al., 1999) utilizó HMM como el modelo subyacente para la detección utilizando llamadas del sistema.

7.4.8. Algoritmos Genéticos

Los algoritmos genéticos son métodos sistemáticos para la resolución de problemas de búsqueda y optimización que aplican los mismos métodos de la evolución. La utilización de algoritmos genéticos para la detección de intrusiones se ha llevado a cabo principalmente con el fin de mejorar la eficiencia seleccionando subconjuntos de características para reducir el número de características observadas manteniendo, o incluso mejorando, la precisión del aprendizaje. Los algoritmos genéticos han sido aplicados a la seguridad informática desde principios de los noventa. El investigador Ludovic Mé (Mé, 1996) planteó el uso de dichos algoritmos como método para analizar los rastros de auditorías de seguridad. En la tesis de Ludovic Mé (Mé, 1998) se presenta el proyecto GASSATA (Genetic Algorithm as an Alternative Tool for Security Audit Trail Analysis) que utiliza un algoritmo genético para buscar la combinación de los ataques conocidos que mejor se correspondan con el evento (o

registro de auditoría) observado. El último trabajo conocido hasta la fecha relativo a los algoritmos genéticos se ha realizado en la universidad del estado de Mississippi (Li, 2004). Inicialmente se clasifica un conjunto de conexiones de red entre normal o intrusiva de forma manual. El algoritmo genético se inicia con un pequeño conjunto de reglas generadas aleatoriamente, y dichas reglas evolucionan hasta generar un conjunto de datos mayor que contiene las nuevas reglas del IDS.

7.4.9.k-NN - Vecino más Cercano

La técnica del vecino más cercano es una técnica que utiliza el aprendizaje basado en instancias o IBL (Instance Based Learning). En este tipo de aprendizaje, se almacenan los ejemplos de entrenamiento y, cuando se quiere clasificar un nuevo objeto, se extraen los objetos más parecidos y se usa su clasificación para catalogar al nuevo objeto. Lo que principalmente diferencia a unos y otros algoritmos es la medida utilizada para definir la similitud entre los objetos. De ese modo, se pueden diferenciar medidas basadas en distancias o medidas basadas en densidad. T. Lane realizó un IDS basado en host para la detección de anomalías basándose en IBL (Lane et al., 1999). Tomó como entradas comandos del shell de UNIX con el fin de mapear datos temporales sobre el espacio, y basó la medida de similaridad en la regla de clasificación 1-NN. Ertoz y Steinbach utilizaron la técnica del vecino más cercano compartido (SNN: Shared Nearest Neighbour) que es particularmente apropiado para encontrar clusters o agrupaciones de diferentes tamaños, densidades y formas en los datos con gran cantidad de ruido (Ertoz et al., 2003).

7.4.10. K-means

K-means es una técnica de clustering basado en particiones: se construyen k particiones de los datos donde cada partición representa un grupo o cluster. Estos métodos, crean una partición inicial e iteran hasta satisfacer el criterio de parada. Los

trabajos más significativos son los realizados por Portnoy, Eskin y Stolfo (Portnoy et al., 2001) y Honig (Honig et al., 2002). Estos últimos realizaron clustering en instancias de datos que contenían tanto comportamientos normales como ataques. Para ello utilizaron un algoritmo modificado de kmeans. Tras realizar la agrupación, utilizaron heurísticos para etiquetar de forma automática cada grupo en normal o ataques.

7.4.11. Lógica Fuzzy

La lógica fuzzy resulta adecuada en el problema de la detección de intrusiones por dos razones principales. Por un lado, están involucradas una gran cantidad de características cuantitativas, y por el otro, la seguridad en sí misma incluye la confusión, es un hecho borroso. Dada una característica cuantitativa, se puede usar un intervalo para indicar un valor normal. En el año 2000 Dickerson y el grupo de investigación comienzan el desarrollo del IDS FIRE (Fuzzy Intrusion Recognition Engine) (Dickerson et al., 2001). En el mismo, varios agentes fuzzy monitorizan una red en busca de signos de intrusión a bajo nivel. La relación entre agentes de bajo nivel y los agentes de mayor nivel se modela mediante Mapas Cognitivos Borrosos (Fuzzy Cognitive Maps) que sirven a modo de base de conocimiento para diferentes tipos de intrusión. También en el 2000, investigadores de la universidad de Mississippi desarrollan el prototipo de un IDS llamado IIDS (Intelligent Intrusion Detection System). El sistema combina tanto la detección de anomalías usando técnicas fuzzy data mining, como detección de uso indebido usando técnicas tradicionales de sistemas expertos basados en reglas (Bridges et al., 2000). El laboratorio de investigación de Sistemas de Seguridad Inteligentes (Intelligent Systems Group) de la universidad de Memphis, con D. Dasgupta a la cabeza, en colaboración con la Universidad Nacional de Colombia, lleva trabajando en este ámbito los últimos años. En su primer trabajo proponen el uso de un algoritmo genético para generar reglas fuzzy (en lugar de realizarlas manualmente) para la detección de anomalías e intrusiones específicas (Gómez et al., 2002). Tratando de

solventar el problema que tienen actualmente los IDS en cuanto a su excesivo número de falsos positivos, en la universidad Carlos III de Madrid utilizan umbrales fuzzy para mejorar la predicción cuando se trabaja con diferentes IDS (Orfilia et al., 2003).

7.4.12. Sistemas Inmunes Artificiales

En esta tesis nos enfocaremos y se muestran más ejemplos relacionados con el SIA, ya que es de nuestro interés.

La idea de utilizar principios inmunológicos en la seguridad informática se inició en 1994. Forrest y su grupo de la universidad de Nuevo México llevan trabajando desde entonces con el propósito de crear un sistema inmune artificial para los ordenadores. El grupo de Forrest realizó su primer experimento de detección de intrusiones a partir de llamadas al sistema de procesos UNIX (Forrest et al., 1996). El sistema recoge información de secuencias de comandos del agente de correo sendmail de UNIX, y la utiliza en el periodo de entrenamiento para definir lo que es propio. En otro trabajo relacionado presentan un IDS basado en sistemas inmunes llamado LISYS (Hofmeyr et al., 2000), y demuestran que el sistema es efectivo a la hora de detectar intrusiones manteniendo un nivel bajo de falsos positivos.

Otros trabajos relacionados son los realizados por Anchor (Anchor, 2005). Este trabajo tiene la finalidad de detectar ataques los cuales son cada vez más sofisticados. Proponen un SIA para detección de virus y de intrusos. Con ello extienden el Sistema Inmune de Defensa Computacional a través del concepto de “vacuna”, creando un nuevo tipo de linfocito, que implementados como transductores de estados finitos, para lo cual proponen una nueva forma de representar linfocitos. Para ello utilizan el concepto “Vacuna” y evolucionan los anticuerpos para detectar los ataques al sistema.

CAPÍTULO VII.- CASO DE ESTUDIO

El modelo central en CDIS es un Sistema Inmune Computacional (CIS) o Sistema Inmune Artificial (ASI), el cual es modelado bajo el Sistema Inmune Biológico[8,9]. Los algoritmos basados en CUS no son modelos exactos del sistema inmune biológico; sin embargo, ellos son abstracciones de ideas del sistema inmune aplicables al problema para ser resuelto, y ellos son implementados en algún grupo computacional o área de la computación. Sólo los conceptos requeridos para la aplicación en particular son mapeadas dentro del CIS, y procesos continuos naturales deben necesariamente ser digitalizados para trabajar en una computadora digital [10]. CDIS, el cual está basado en el modelo de selección negativa, usa el concepto de propio no propio. La definición de lo propio y lo no propio es importante para el funcionamiento apropiado del CIS, pero cada aplicación particular tiene su propio método para definir estos conceptos [10]. Idealmente, el entrenamiento del conjunto que representa lo propio está diseñado para representar lo más cercano posible la actividad normal que toma lugar en la parte que el SIC protege.

Existen trabajos previos en el desarrollo de sistemas inmunes artificiales y sistemas inmune computacionales, y varios ejes de investigación están en marcha en el área de defensa computacional.

Estas técnicas se han mostrado como una promesa en la detección de intrusos. Este trabajo, de cualquier modo, se ha inspirado en el trabajo de los grupos de investigación de personas Dasgupta, Forrest y Lamont. Dasgupta y Gonzalez trabajaron la detección de intrusos en la red de trabajo y desarrollaron una herramienta para la toma (soporte) de decisiones basadas en un AG y un clasificador (GA-based, Classifier-based) para asistir en la respuesta a un intruso. Forrest, Hofmeyr, y otros han hecho trabajo extensivo con el uso y diseño de CISs para aplicaciones de seguridad computacional. Ellos han construido CIS para host-based ID y para network-based ID. Sus host-based IDS definen lo propio como secuencia de llamadas del sistema hechas por programas con privilegios, así este detecta llamadas del sistema anormales o no propias. Sus network-based IDS, llamado LISYS, usa 3 características para definir lo propio : la dirección IP de la fuente, la dirección IP del destino y el puerto. Sólo paquetes TCP SYN, del cual la señal de inicio de una

CAPÍTULO VII.- CASO DE ESTUDIO

conexión, son monitoreados por el IDS. Las conexiones que ocurren con frecuencia son consideradas como propias [19]. Esta investigación difiere de los trabajos mencionados anteriormente en muchos aspectos. Por ejemplo, CDI usa 28 características, que incluyen las 3 que se usan en LISYS. También, CDIS examina todos los paquetes TCP, UDP, y ICMP, y no sólo monitorea los paquetes TCP SYN.



Capítulo 8

Conclusiones y Trabajo Futuro

De acuerdo al objetivo de crear una metodología para resolver problemas de Optimización Combinatoria mediante el SIA, es todavía un punto en el cuál trabajar ya que sólo se eligieron 2 problemas tipos. Es entonces tema de futuras investigaciones el probar la metodología para otros problemas de Optimización Combinatoria.

Como parte de la investigación, se recabo información de literatura del área de Inmunología así como la asesoría de expertos en el tema para tomar elementos del SIN, que ya se conocen en SIA, para que quedaran claros buscando una correcta interpretación, así como elementos que se consideran novedosos y que pudieran hacer del SIA una mejor técnica de solución de problemas. Es por esto que se considera como trabajo futuro el explotar nuevas teorías, conceptos, etc. del SIN, especialmente de la inmunidad innata, que hagan del SIA

ANEXO

una metaheurística más robusta, eficiente y que puedan ser incorporados para que la simulación sea más cercana a la realidad.

La metodología incorpora un diagrama en el cual se muestra etapas que importantes y que se consideran está siempre presentes en el SIN. Cuando se trata de crear un algoritmo del SIA este siempre va de la mano con el problema a resolver, es decir, cuando ya se ha definido de manera clara el problema, se puede elegir el algoritmo del SIA. De este modo se logra no violar el “no free lunch”.

A cada fase se puede regresar cuantas veces sea necesario para mejorar alguna punto en particular, como en una espiral, con lo cual se asegura que el algoritmo sea lo más apegado a lo que se necesita para resolver el problema y también que simule lo mejor posible al SIN. No consideramos que con la metodología se van a resolver todos los problema pero si que se pueda identificar si el problema que se desee resolver se puede con el SIA y de ser así se pueda crear un algoritmo que tome los modelos del SIN que mejor ajusten para el propósito de resolver un problema.

Los algoritmos del SIA tienen características que los hacen similar a otras heurísticas y a través de la metodología también se logra incorporar características que los hace más del SIA y que lo diferencié más claramente de alguna otra técnica. Con los resultado obtenidos consideramos se le da un punto a favor del SIA para consolidarse como una técnica que verdaderamente resuelve problemas.

Como trabajo futuro se considera el análisis de la metodología y verificar en que fase esta puede ser fortalecida para lograr crear algoritmos más eficientes y con mejor desempeño.

Anexo 1

Libros

1. Francisco Javier Álvarez Rodríguez, Jaime Muñoz Arteaga, Juan Pedro Cardona Salas, Mónica Brizuela Sandoval, Fatima Sayuri Quezada Aguilera, Julio Cesar Ponce Gallegos. Interpretación del Modelo de Madurez de Capacidades (CMM) para pequeñas industrias de software, Textos Universitarios, ISBN 978-970-728-020-5, 200 Ejemplares, Primera Edición, Septiembre 2008.

Capítulos en Libros

2. Julio Ponce, Felipe Padilla, Alberto Ochoa, Alejandro Padilla, Eunice Ponce de León, Fátima Quezada, Ant Colony Algorithm for Clustering through of Cliques, Artificial Intelligence & Applications, A. Gelbukh(Ed.), ISBN 978-607-95367-0-1 Noviembre del 2009.
3. Methodology to solve Combinatorial Optimization Problems with the Artificial Immune System. Fatima Quezada, Felipe Padilla, Eunice Ponce, Alberto Ochoa and Christian Correa. MICAI 2009, Workshop Artificial Intelligence Applications, p.p. 5-10. ISBN978-607-95367-0-1.Ed.Gelbukh.

Congresos Internacionales

4. Alberto Ochoa, Julio Ponce, Alejandro Padilla, Fatima Sayuri Quezada y Miguel Meza. Six degrees of separation in a graph to a Social Networking. XXII Congreso Nacional y VIII Internacional de Informática y Computación ANIEI 2009, 21-23 de Octubre del 2009, Ensenada, Baja California



ANEXO



Anexo 2

Artículos

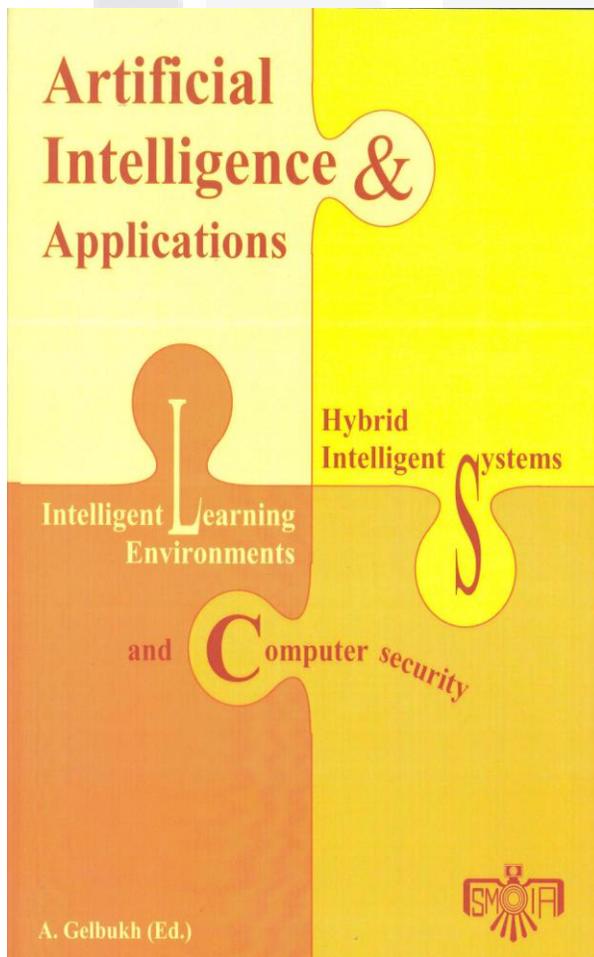


Table of Contents		Page
PART 1. Hybrid Intelligent Systems		
<i>Evolutionary Computation</i>		
Methodology to Solve Combinatorial Optimization Problems with the Artificial Immune System		5
<i>Fatima Quezada, Felipe Padilla, Eunice Ponce, Alberto Ochoa and Christian Correa</i>		
New Algorithm Transgenic with Homology to solve the problem of OneMax		11
<i>Christian José Correa Villalón, Felipe Padilla Díaz, Eugenio Pérez Morphe, Fátima Sayuri Quezada Aguilera and Ochoa Ortiz-Zezatti</i>		
Genetic Algorithm using Migration and Modified GSX as Support in the Diversification of Populations		21
<i>Francisco Ornelas, Alejandro Padilla, Alberto Ochoa, Eunice Ponce de León and Felipe Padilla</i>		
Ant Colony Algorithm for Clustering through of Cliques		29
<i>Julio Cesar Ponce Gallegos, Felipe Padilla Díaz, Carlos Alberto Ochoa Ortiz Zezatti, Alejandro Padilla Díaz, Eunice Esther Ponce de León and Fatima Sayuri Quezada Aguilera</i>		
Fuzzy Characterization of Qualitative Variables for the Development of an Overall Efficiency Index of Competitiveness: Using a Modified Triangular Membership Function		35
<i>Darwin Young, Pedro Pérez, Mario Cantú and Mauricio Garza</i>		
A Hybrid Simulation based on Multi-Objective Algorithm for Manufacturing Cells Optimization		49
<i>Corram-Siller Elías, Torres-Treviño Luis and Pérez-Villanueva Pedro</i>		
<i>Logistics and Simulation Modeling</i>		
Adaptation of Particle Swarm Optimization Technique to Optimize the Choice of Routes between two Points in the Transport System of the City of Leon, Guanajuato		59
<i>Oscar Alberto García Ortiz, Alberto Ochoa Zezatti and Sodel Vázquez Reyes</i>		

ISBN: 978-607-95367-0-1

Methodology to Solve Combinatorial Optimization Problems with the Artificial Immune System

Fatima Quezada¹, Felipe Padilla¹, Eunice Ponce¹, Alberto Ochoa²,
Christian Correa¹,

¹ Universidad Autónoma de Ciudad Juárez

sayu_172002@yahoo.com.mx fpadilla2007@yahoo.ca, eponce@correo.uaa.mx, crisjc_2000@yahoo.com.mx

² U megamax8@hotmail.com

Abstract. This paper presents an approximation to the Methodology to solve Combinatorial Optimization Problems (CO) using the Artificial Immune System (AIS). This Methodology is based principally in the **performance** of the Natural Immune System (NIS), which one consist in the cell creation until the elimination of the antigen and the recirculation of the different cells trough the organism, this with the objective to find the best solution to a OC problem.

Keywords: Natural Immune System, Artificial Immune System, Combinatorial Optimization Problems

1 Introduction

The Natural Immune System is the responsible for maintain is a health state the organism, it means, to save from illness, infections, etc. caused for the presence of harmful agents like microorganism, called antigens. The NIS is divided by the innate immunity and the adaptive immunity. The last one have been an object of deeper study in the AIS because it has characteristics that are desirables in any intelligent or optimization system. The work of NIS follows a methodology, in which the active and passive cells are defined, the interactions between each component of the system and even so the defense mechanisms that will be used, this will depend of the antigen that attack the organism. This NIS is active even when there is not a detected threat, or the threat is unknown, and for this reason the NIS is in alert and in vigilance of the different parts of the organism that can be affected by the strange harmful antigens. This methodology can be translated to the AIS to be simulated and used to solve diverse problems like the NIS do.

2 Natural Immune System

The Word immune means protection to face the illness and, more specifically, to face the infectious illness. The cells an molecules responsible of the immunity constitute the Immune System, and the global coordinated response to the introduction of stranger substances is the immune response. [1]

2.1 Innate Immunity

The defense face to the microorganism is realized by the first reactions of the innate immunity and for the later reactions of the adaptive immunity. The innate immunity is constituted by mechanisms that exist before the infection was developed, which are capable to establish fast responses to the microorganism and react of the same way to repeated infections. [1]

2.2 Adaptive Immunity

It is conformed by mechanisms of defense more evolved that are stimulated after de exposition to the infectious agents y of which intensity and defense capacity grow after the subsequent exposition to a determined microorganism. This is called Adaptive Immunity because is developed like a response to an infection.

Some of its characteristics are the specificity to different macromolecules and the capacity to remember and response with greater intensity after repeated expositions to the same microorganism. Due to this capacity to differentiate between the microorganisms and macromolecules, even they are closely related; the adaptive immunity is called too Specific Immunity, or Acquired Immunity to highlight the fact that is based in strong protector responses acquired with the experience.

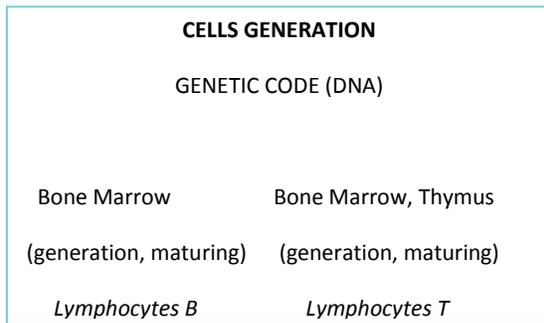
The components are the lymphocytes B and T and the products of them. [1]

2.3 Estimated Model of the performance of the Natural Immune System

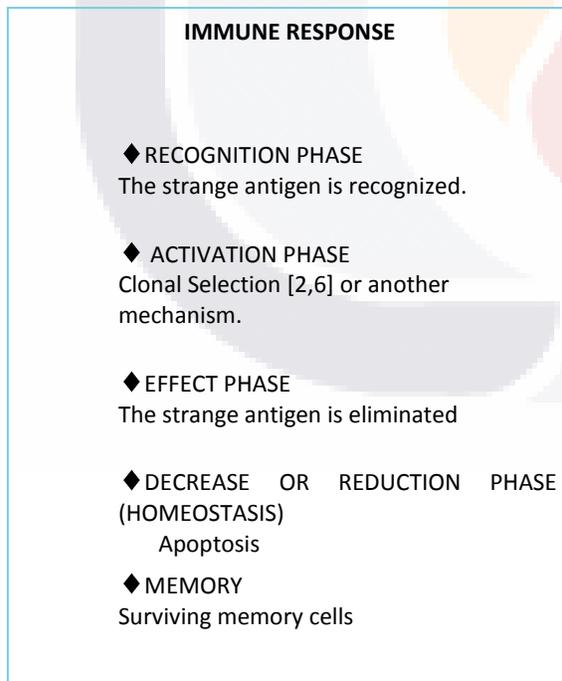
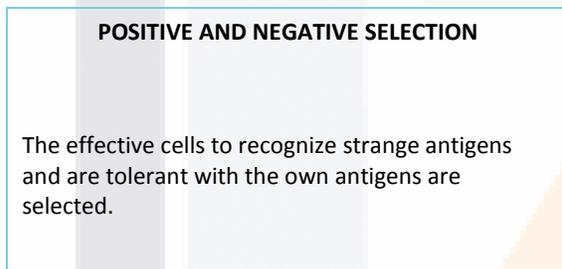
The next is the model in which is described briefly the work of the NIS, in which is considered since the NIS cells generation phase until the recirculation.

In the Immune Response is only considered the mechanisms that work in the adaptive immunity.

ANEXO



There is a stimulation to generate continually new cells and maintain the lymphocytes population.



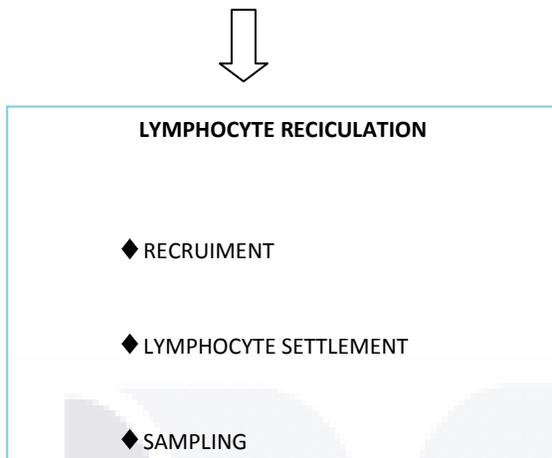


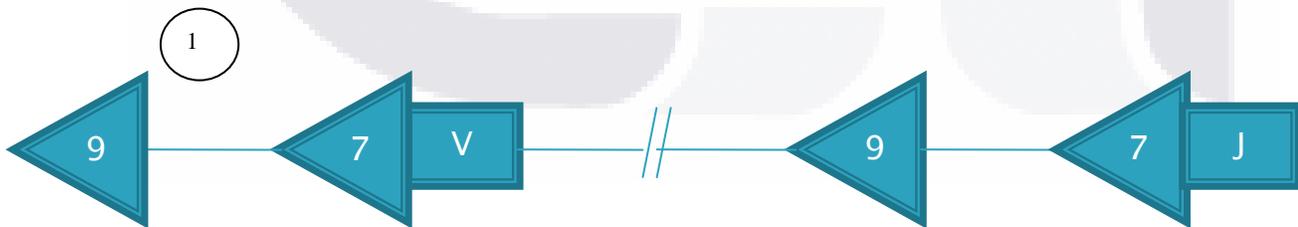
Fig. 1. Estimated Model of the performance of the Natural Immune System

2.3 AIS Methodology

Considering the model of the work of the NIS, the simulation of the generation cells phase is shown, that in the AIS is the feasible solution generation for OC problem.

Even though the NIS begins with the generation of lymphocytes T and B (bone marrow and thymus), whose receptors will have affinity to one antigen. This cells have a maturation process in which have changes a genetic level to be able of take part of the NIS. Some of this mechanism that are used to achieve the maturation are shown.

REVERSAL



ANEXO

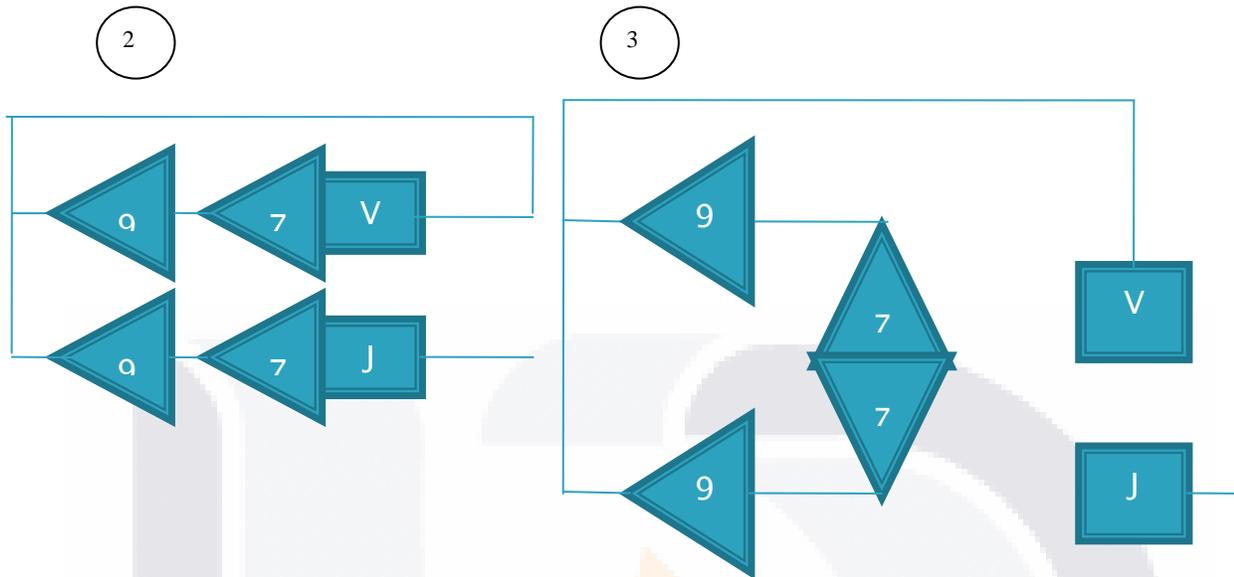
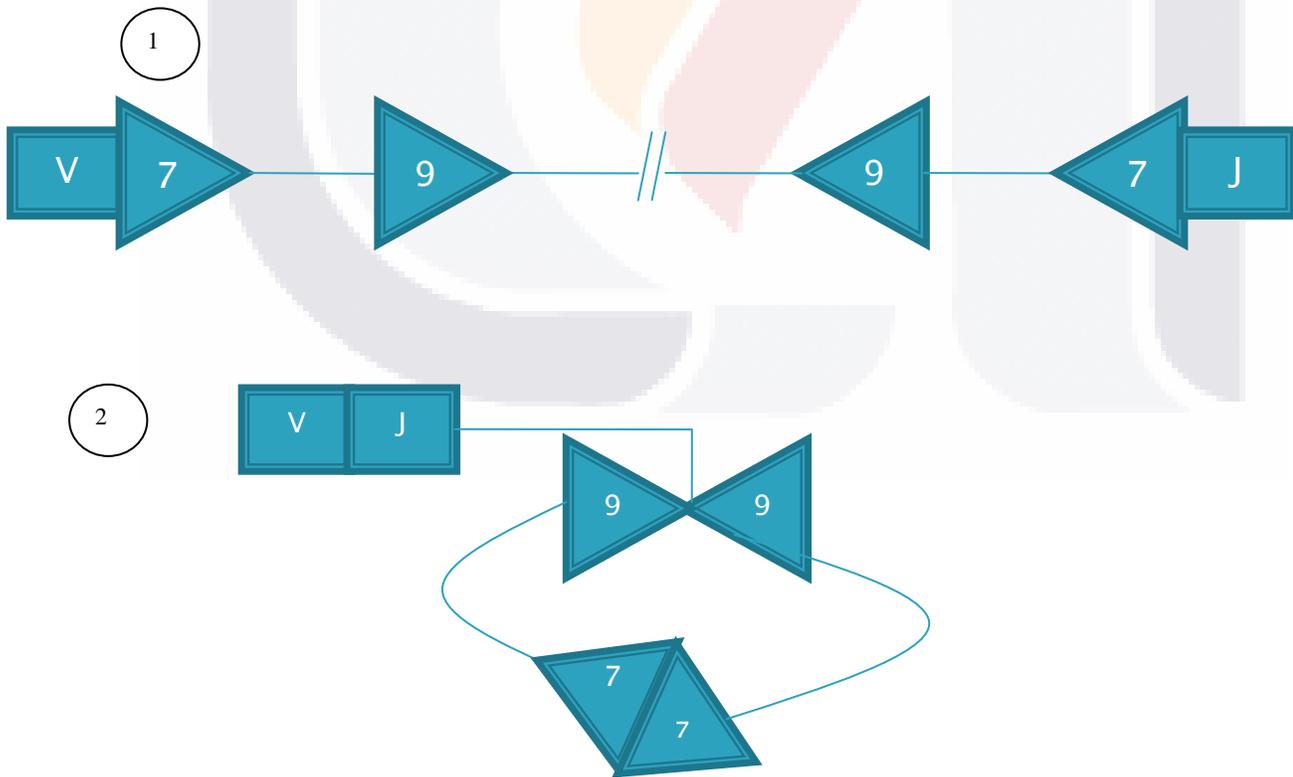


Fig. 2. Reversal Operator

DELETION



ANEXO

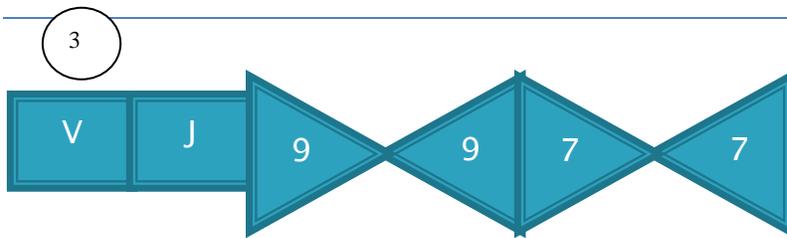


Fig. 3. Deletion Operator

The generated feasible solutions are equivalents to the lymphocytes. The next pseudocode shows how the simulation of the first phase (generation cells) and how the solutions are modified in the maturation process.

Generate population

 Generate the initial population of solutions (lymphocytes)

 Generate each one of the sub solutions that will be part of the solution

 Evaluate the affinity

Define the J, V, 7 and 9 solutions.

End of Generate population

Maturation process

 Apply Reversal Operator

 Evaluate affinity

 If the solution improve

 If the improvement is low (less than 10%) go to the Deletion operator only
 the modified solution

 Else

 End maturation process

 Else

 Go to Deletion operator the original and modified solution

 Only the best solution go to the population

End Maturation process

Fin del proceso de maduración.

Referencias

1. Abbas Abul K., Lichtman Andrew H., Pober Jordan S. Inmunología Celular y Molecular. Cuarta Edición. McGraw-Hill.
2. Burnet FM. A modification of Jerne's theory of antibody production using the concept of clonal selection. Australian Journal of Science 20:67-69, 1957.
3. De Castro Leandro Nunes y Timmis Jonathan. An introduction to Artificial Immune Systems: A New Computational Paradigm. Springer-Verlag, 2002.
4. De Castro Leandro N. and Von Zuben F. J.. Learning and Optimization Using the Clonal Selection Principle. IEEE Transactions on Evolutionary Computation, 6(3):239_251, 2002.

ANEXO

5. De Castro Leandro Nunes and Timmis Jon. An artificial immune network for multimodal function optimization. In Proceedings of the special sessions on artificial immune systems in the 2002 Congress on Evolutionary Computation, 2002 IEEE World Congress on Computational Intelligence, volume I, pages 669_674, Honolulu, Hawaii, May 2002.
6. Farmer J. D., Packard N. H., and Perelson A. S.. The Immune System, Adaptation, and Machine Learning. Physica D, 22:187_204, 1986.
7. Jerne NK. The natural-selection theory of antibody formation. Proceedings of the National Academy of Sciences USA 41:849-857,1955.



ANEXO



Anexo 3

Tablas de Resultados de los experimentos realizados con los benchmarks de la TSPLIB.

Tabla 13 Resultados de los experimentos realizados con el benchmark Burman 14

Réplicas	Ciudades	Generaciones	Pob. Inicial	Clones	Nuevos	% Selección	P.Divi	Costo
1	14	40	20	1	1	50	25	4487.43883
2	14	40	20	1	1	50	25	5017.26226
3	14	40	20	1	1	50	25	4524.34076
4	14	40	20	1	1	50	25	4213.10852
5	14	40	20	1	1	50	25	4471.57669
6	14	40	20	1	1	50	25	4967.36309
7	14	40	20	1	1	50	25	4130.02067
8	14	40	20	1	1	50	25	4196.98232
9	14	40	20	1	1	50	25	4743.8199
10	14	40	20	1	1	50	25	3963.27921
1	14	80	20	1	1	50	25	4263.02624
2	14	80	20	1	1	50	25	3916.45609
3	14	80	20	1	1	50	25	4107.97571
4	14	80	20	1	1	50	25	4344.40399
5	14	80	20	1	1	50	25	3993.01456
6	14	80	20	1	1	50	25	4753.5256
7	14	80	20	1	1	50	25	4453.37212
8	14	80	20	1	1	50	25	4362.7937
9	14	80	20	1	1	50	25	4233.22576
10	14	80	20	1	1	50	25	4421.71678
1	14	120	20	1	1	50	25	4579.89382
2	14	120	20	1	1	50	25	4277.38275
3	14	120	20	1	1	50	25	4176.81324
4	14	120	20	1	1	50	25	4560.65769
5	14	120	20	1	1	50	25	4192.40616
6	14	120	20	1	1	50	25	4360.87007
7	14	120	20	1	1	50	25	4620.99059
8	14	120	20	1	1	50	25	4326.59399
9	14	120	20	1	1	50	25	4840.94625

ANEXO

10	14	120	20	1	1	50	25	4519.41139
1	14	40	40	1	1	50	25	4190.4464
2	14	40	40	1	1	50	25	4471.47923
3	14	40	40	1	1	50	25	4760.78464
4	14	40	40	1	1	50	25	4216.33366
5	14	40	40	1	1	50	25	4070.05859
6	14	40	40	1	1	50	25	4209.59303
7	14	40	40	1	1	50	25	4621.16373
8	14	40	40	1	1	50	25	3950.90368
9	14	40	40	1	1	50	25	4322.71473
10	14	40	40	1	1	50	25	4590.75555
1	14	80	40	1	1	50	25	4002.59813
2	14	80	40	1	1	50	25	4473.56071
3	14	80	40	1	1	50	25	3999.20262
4	14	80	40	1	1	50	25	4525.64544
5	14	80	40	1	1	50	25	4357.25402
6	14	80	40	1	1	50	25	4053.88689
7	14	80	40	1	1	50	25	4212.12966
8	14	80	40	1	1	50	25	4316.69478
9	14	80	40	1	1	50	25	3952.14295
10	14	80	40	1	1	50	25	3883.89172
1	14	120	40	1	1	50	25	4099.57375
2	14	120	40	1	1	50	25	3805.75323
3	14	120	40	1	1	50	25	4044.70048
4	14	120	40	1	1	50	25	4168.8259
5	14	120	40	1	1	50	25	4476.9948
6	14	120	40	1	1	50	25	4148.86941
7	14	120	40	1	1	50	25	3789.13603
8	14	120	40	1	1	50	25	3852.02432
9	14	120	40	1	1	50	25	4112.2994
10	14	120	40	1	1	50	25	4319.05011
1	14	40	60	1	1	50	25	4109.33632
2	14	40	60	1	1	50	25	4226.95481
3	14	40	60	1	1	50	25	4022.44355
4	14	40	60	1	1	50	25	4435.88637
5	14	40	60	1	1	50	25	4454.69657
6	14	40	60	1	1	50	25	4288.96307
7	14	40	60	1	1	50	25	3746.99018
8	14	40	60	1	1	50	25	4179.66006
9	14	40	60	1	1	50	25	4172.61211
10	14	40	60	1	1	50	25	4132.5273
1	14	80	60	1	1	50	25	4259.86273
2	14	80	60	1	1	50	25	4157.67109
3	14	80	60	1	1	50	25	3957.48661
4	14	80	60	1	1	50	25	4301.30183
5	14	80	60	1	1	50	25	3954.7706
6	14	80	60	1	1	50	25	4080.44648
7	14	80	60	1	1	50	25	4042.69432
8	14	80	60	1	1	50	25	4003.51054
9	14	80	60	1	1	50	25	3958.67337
10	14	80	60	1	1	50	25	4257.62654
1	14	120	60	1	1	50	25	3800.62369

ANEXO

2	14	120	60	1	1	50	25	3975.45168
3	14	120	60	1	1	50	25	4109.80148
4	14	120	60	1	1	50	25	3950.03421
5	14	120	60	1	1	50	25	4294.74809
6	14	120	60	1	1	50	25	4085.23996
7	14	120	60	1	1	50	25	3961.97586
8	14	120	60	1	1	50	25	4287.49429
9	14	120	60	1	1	50	25	3795.16556
10	14	120	60	1	1	50	25	3980.76871
1	14	40	20	1	1	50	30	5041.58266
2	14	40	20	1	1	50	30	4310.60772
3	14	40	20	1	1	50	30	4930.15643
4	14	40	20	1	1	50	30	4444.767
5	14	40	20	1	1	50	30	4781.66988
6	14	40	20	1	1	50	30	4755.94117
7	14	40	20	1	1	50	30	4552.83602
8	14	40	20	1	1	50	30	3754.17256
9	14	40	20	1	1	50	30	4798.48075
10	14	40	20	1	1	50	30	4437.97795
1	14	80	20	1	1	50	30	4377.92679
2	14	80	20	1	1	50	30	4894.33589
3	14	80	20	1	1	50	30	4463.12219
4	14	80	20	1	1	50	30	3947.23065
5	14	80	20	1	1	50	30	4587.22777
6	14	80	20	1	1	50	30	4732.65222
7	14	80	20	1	1	50	30	4859.32223
8	14	80	20	1	1	50	30	4180.80895
9	14	80	20	1	1	50	30	4269.29419
10	14	80	20	1	1	50	30	3957.75859
1	14	120	20	1	1	50	30	4327.32933
2	14	120	20	1	1	50	30	4546.12979
3	14	120	20	1	1	50	30	4362.44589
4	14	120	20	1	1	50	30	4588.91835
5	14	120	20	1	1	50	30	3935.67145
6	14	120	20	1	1	50	30	4562.0355
7	14	120	20	1	1	50	30	4309.37082
8	14	120	20	1	1	50	30	4704.29648
9	14	120	20	1	1	50	30	4384.22392
10	14	120	20	1	1	50	30	4784.83162
1	14	40	40	1	1	50	30	4288.32925
2	14	40	40	1	1	50	30	4354.05085
3	14	40	40	1	1	50	30	4388.14799
4	14	40	40	1	1	50	30	4018.7497
5	14	40	40	1	1	50	30	4106.14008
6	14	40	40	1	1	50	30	4383.7186
7	14	40	40	1	1	50	30	4525.96971
8	14	40	40	1	1	50	30	4617.18508
9	14	40	40	1	1	50	30	4376.48195
10	14	40	40	1	1	50	30	4123.69228
1	14	80	40	1	1	50	30	4160.46761
2	14	80	40	1	1	50	30	4281.61668

ANEXO

3	14	80	40	1	1	50	30	3898.01417
4	14	80	40	1	1	50	30	4774.32758
5	14	80	40	1	1	50	30	4341.17491
6	14	80	40	1	1	50	30	4121.86367
7	14	80	40	1	1	50	30	4337.4083
8	14	80	40	1	1	50	30	4194.26815
9	14	80	40	1	1	50	30	4195.4231
10	14	80	40	1	1	50	30	3937.43977
1	14	120	40	1	1	50	30	4151.74311
2	14	120	40	1	1	50	30	4013.03465
3	14	120	40	1	1	50	30	4484.81424
4	14	120	40	1	1	50	30	4323.75552
5	14	120	40	1	1	50	30	4253.9055
6	14	120	40	1	1	50	30	4287.07559
7	14	120	40	1	1	50	30	4053.40135
8	14	120	40	1	1	50	30	4432.81253
9	14	120	40	1	1	50	30	4546.80661
10	14	120	40	1	1	50	30	3739.94708
1	14	40	60	1	1	50	30	4063.53182
2	14	40	60	1	1	50	30	3940.91425
3	14	40	60	1	1	50	30	4136.40773
4	14	40	60	1	1	50	30	4449.40778
5	14	40	60	1	1	50	30	3964.65903
6	14	40	60	1	1	50	30	4323.46275
7	14	40	60	1	1	50	30	4167.43255
8	14	40	60	1	1	50	30	4375.40931
9	14	40	60	1	1	50	30	4345.47588
10	14	40	60	1	1	50	30	4378.79652
1	14	80	60	1	1	50	30	4378.65378
2	14	80	60	1	1	50	30	4263.06733
3	14	80	60	1	1	50	30	4046.4736
4	14	80	60	1	1	50	30	4332.1032
5	14	80	60	1	1	50	30	4099.88691
6	14	80	60	1	1	50	30	4155.55311
7	14	80	60	1	1	50	30	4205.41117
8	14	80	60	1	1	50	30	4257.14527
9	14	80	60	1	1	50	30	4598.16958
10	14	80	60	1	1	50	30	3999.42296
1	14	120	60	1	1	50	30	4272.38502
2	14	120	60	1	1	50	30	4089.21846
3	14	120	60	1	1	50	30	3793.68781
4	14	120	60	1	1	50	30	4271.5111
5	14	120	60	1	1	50	30	4394.91938
6	14	120	60	1	1	50	30	4013.87031
7	14	120	60	1	1	50	30	4040.84113
8	14	120	60	1	1	50	30	4039.91052
9	14	120	60	1	1	50	30	4067.17371
10	14	120	60	1	1	50	30	4235.63969
1	14	40	20	1	1	50	50	4930.17818
2	14	40	20	1	1	50	50	4500.22854
3	14	40	20	1	1	50	50	4976.43273
4	14	40	20	1	1	50	50	4743.46245

ANEXO

5	14	40	20	1	1	50	50	4453.92686
6	14	40	20	1	1	50	50	4064.59725
7	14	40	20	1	1	50	50	4084.31118
8	14	40	20	1	1	50	50	4355.19089
9	14	40	20	1	1	50	50	4658.52254
10	14	40	20	1	1	50	50	5016.51059
1	14	80	20	1	1	50	50	4645.50111
2	14	80	20	1	1	50	50	4911.55727
3	14	80	20	1	1	50	50	4882.47448
4	14	80	20	1	1	50	50	4834.24319
5	14	80	20	1	1	50	50	3989.84295
6	14	80	20	1	1	50	50	4724.817
7	14	80	20	1	1	50	50	4236.03884
8	14	80	20	1	1	50	50	4109.25112
9	14	80	20	1	1	50	50	4851.80021
10	14	80	20	1	1	50	50	4309.05146
1	14	120	20	1	1	50	50	4612.91024
2	14	120	20	1	1	50	50	4942.86275
3	14	120	20	1	1	50	50	4253.38534
4	14	120	20	1	1	50	50	3977.05721
5	14	120	20	1	1	50	50	4635.86445
6	14	120	20	1	1	50	50	4635.7624
7	14	120	20	1	1	50	50	4686.07568
8	14	120	20	1	1	50	50	4448.58738
9	14	120	20	1	1	50	50	4964.47215
10	14	120	20	1	1	50	50	3918.92911
1	14	40	40	1	1	50	50	4697.52689
2	14	40	40	1	1	50	50	4163.21949
3	14	40	40	1	1	50	50	4311.2897
4	14	40	40	1	1	50	50	4479.39411
5	14	40	40	1	1	50	50	4461.35974
6	14	40	40	1	1	50	50	4193.57309
7	14	40	40	1	1	50	50	4561.73134
8	14	40	40	1	1	50	50	4284.91692
9	14	40	40	1	1	50	50	4621.13133
10	14	40	40	1	1	50	50	4237.77994
1	14	80	40	1	1	50	50	3893.23346
2	14	80	40	1	1	50	50	4061.49687
3	14	80	40	1	1	50	50	3954.11339
4	14	80	40	1	1	50	50	4130.66042
5	14	80	40	1	1	50	50	3976.86061
6	14	80	40	1	1	50	50	4006.63065
7	14	80	40	1	1	50	50	3975.50686
8	14	80	40	1	1	50	50	4291.83705
9	14	80	40	1	1	50	50	4222.021
10	14	80	40	1	1	50	50	4102.0696
1	14	120	40	1	1	50	50	3848.43528
2	14	120	40	1	1	50	50	3887.03516
3	14	120	40	1	1	50	50	4107.5169
4	14	120	40	1	1	50	50	4910.42815
5	14	120	40	1	1	50	50	3973.38806

ANEXO

6	14	120	40	1	1	50	50	4201.1859
7	14	120	40	1	1	50	50	4522.05924
8	14	120	40	1	1	50	50	3740.92458
9	14	120	40	1	1	50	50	4173.53784
10	14	120	40	1	1	50	50	4468.96742
1	14	40	60	1	1	50	50	4032.81832
2	14	40	60	1	1	50	50	4242.77952
3	14	40	60	1	1	50	50	4263.30843
4	14	40	60	1	1	50	50	4076.27844
5	14	40	60	1	1	50	50	4579.13974
6	14	40	60	1	1	50	50	4333.12791
7	14	40	60	1	1	50	50	3714.10656
8	14	40	60	1	1	50	50	4356.65729
9	14	40	60	1	1	50	50	4044.87521
10	14	40	60	1	1	50	50	4181.95717
1	14	80	60	1	1	50	50	3902.92861
2	14	80	60	1	1	50	50	4022.04131
3	14	80	60	1	1	50	50	3990.74557
4	14	80	60	1	1	50	50	3981.16921
5	14	80	60	1	1	50	50	4224.8127
6	14	80	60	1	1	50	50	4512.79399
7	14	80	60	1	1	50	50	4341.06692
8	14	80	60	1	1	50	50	3973.23978
9	14	80	60	1	1	50	50	3724.66389
10	14	80	60	1	1	50	50	3987.83814
1	14	120	60	1	1	50	50	4104.40318
2	14	120	60	1	1	50	50	3853.51465
3	14	120	60	1	1	50	50	4212.98231
4	14	120	60	1	1	50	50	4175.85323
5	14	120	60	1	1	50	50	4121.97999
6	14	120	60	1	1	50	50	4099.86581
7	14	120	60	1	1	50	50	3627.03934
8	14	120	60	1	1	50	50	3934.36691
9	14	120	60	1	1	50	50	4425.74341
10	14	120	60	1	1	50	50	4604.94717

Tabla 14 Resultados de los experimentos realizados con el benchmark Baygs 29

Réplicas	Ciudades	Generaciones	Pob. Inicial	Clones	Nuevos	% Selección	P.Divi	Costo
1	29	40	20	1	1	50	25	226326.3503
2	29	40	20	1	1	50	25	215120.1307
3	29	40	20	1	1	50	25	214328.9157
4	29	40	20	1	1	50	25	208950.5363
5	29	40	20	1	1	50	25	199780.4834
6	29	40	20	1	1	50	25	198677.5357
7	29	40	20	1	1	50	25	208405.4203
8	29	40	20	1	1	50	25	206160.8046

ANEXO

9	29	40	20	1	1	50	25	224939.7717
10	29	40	20	1	1	50	25	213228.5212
1	29	80	20	1	1	50	25	210527.2281
2	29	80	20	1	1	50	25	214041.1563
3	29	80	20	1	1	50	25	196415.3233
4	29	80	20	1	1	50	25	216009.5067
5	29	80	20	1	1	50	25	194796.1934
6	29	80	20	1	1	50	25	209635.3909
7	29	80	20	1	1	50	25	193303.1517
8	29	80	20	1	1	50	25	198602.9046
9	29	80	20	1	1	50	25	202786.0704
10	29	80	20	1	1	50	25	211029.0127
1	29	120	20	1	1	50	25	217627.4331
2	29	120	20	1	1	50	25	189151.5075
3	29	120	20	1	1	50	25	188250.7753
4	29	120	20	1	1	50	25	216486.05
5	29	120	20	1	1	50	25	215400.1365
6	29	120	20	1	1	50	25	195107.9622
7	29	120	20	1	1	50	25	207782.0568
8	29	120	20	1	1	50	25	195003.7459
9	29	120	20	1	1	50	25	207866.0838
10	29	120	20	1	1	50	25	195237.4714
1	29	40	40	1	1	50	25	213951.8943
2	29	40	40	1	1	50	25	193720.8274
3	29	40	40	1	1	50	25	196854.852
4	29	40	40	1	1	50	25	204305.3219
5	29	40	40	1	1	50	25	205338.153
6	29	40	40	1	1	50	25	188050.075
7	29	40	40	1	1	50	25	198926.6705
8	29	40	40	1	1	50	25	212154.0948
9	29	40	40	1	1	50	25	203895.6794
10	29	40	40	1	1	50	25	208505.1811
1	29	80	40	1	1	50	25	208290.2366
2	29	80	40	1	1	50	25	194378.7528
3	29	80	40	1	1	50	25	194505.5936
4	29	80	40	1	1	50	25	196892.2011
5	29	80	40	1	1	50	25	196042.5968
6	29	80	40	1	1	50	25	201202.9664
7	29	80	40	1	1	50	25	200740.1221
8	29	80	40	1	1	50	25	209294.6647
9	29	80	40	1	1	50	25	200440.3664
10	29	80	40	1	1	50	25	203010.1763
1	29	120	40	1	1	50	25	196794.6532
2	29	120	40	1	1	50	25	195592.8222

ANEXO

3	29	120	40	1	1	50	25	206915.0139
4	29	120	40	1	1	50	25	201725.4714
5	29	120	40	1	1	50	25	183042.2823
6	29	120	40	1	1	50	25	194354.4622
7	29	120	40	1	1	50	25	192615.5813
8	29	120	40	1	1	50	25	199923.4199
9	29	120	40	1	1	50	25	173693.4786
10	29	120	40	1	1	50	25	194851.5984
1	29	40	60	1	1	50	25	213291.9924
2	29	40	60	1	1	50	25	205005.6555
3	29	40	60	1	1	50	25	185519.6159
4	29	40	60	1	1	50	25	172484.788
5	29	40	60	1	1	50	25	197595.2764
6	29	40	60	1	1	50	25	187730.8944
7	29	40	60	1	1	50	25	199665.7684
8	29	40	60	1	1	50	25	213484.8079
9	29	40	60	1	1	50	25	196781.3776
10	29	40	60	1	1	50	25	195169.1407
1	29	80	60	1	1	50	25	203527.1243
2	29	80	60	1	1	50	25	194624.8231
3	29	80	60	1	1	50	25	208103.4041
4	29	80	60	1	1	50	25	207218.2514
5	29	80	60	1	1	50	25	188873.2972
6	29	80	60	1	1	50	25	192168.0803
7	29	80	60	1	1	50	25	181397.9234
8	29	80	60	1	1	50	25	203400.6279
9	29	80	60	1	1	50	25	198968.7075
10	29	80	60	1	1	50	25	197708.2536
1	29	120	60	1	1	50	25	193602.5263
2	29	120	60	1	1	50	25	183532.5987
3	29	120	60	1	1	50	25	194700.7516
4	29	120	60	1	1	50	25	201150.1942
5	29	120	60	1	1	50	25	189874.3682
6	29	120	60	1	1	50	25	180462.5178
7	29	120	60	1	1	50	25	201127.7697
8	29	120	60	1	1	50	25	188315.5424
9	29	120	60	1	1	50	25	192873.8524
10	29	120	60	1	1	50	25	197265.3186
1	29	40	20	1	1	50	30	203699.4645
2	29	40	20	1	1	50	30	223457.9785
3	29	40	20	1	1	50	30	208035.179
4	29	40	20	1	1	50	30	219251.0426
5	29	40	20	1	1	50	30	204398.7194
6	29	40	20	1	1	50	30	205633.2081
7	29	40	20	1	1	50	30	218319.8356

ANEXO

8	29	40	20	1	1	50	30	216751.275
9	29	40	20	1	1	50	30	183765.8235
10	29	40	20	1	1	50	30	215136.5904
1	29	80	20	1	1	50	30	203622.8086
2	29	80	20	1	1	50	30	210031.8008
3	29	80	20	1	1	50	30	203400.4962
4	29	80	20	1	1	50	30	202781.3072
5	29	80	20	1	1	50	30	222899.059
6	29	80	20	1	1	50	30	216772.5343
7	29	80	20	1	1	50	30	195779.8088
8	29	80	20	1	1	50	30	197293.7875
9	29	80	20	1	1	50	30	197044.88
10	29	80	20	1	1	50	30	197203.9884
1	29	120	20	1	1	50	30	229345.6185
2	29	120	20	1	1	50	30	184098.3133
3	29	120	20	1	1	50	30	206964.0418
4	29	120	20	1	1	50	30	205863.3326
5	29	120	20	1	1	50	30	199642.84
6	29	120	20	1	1	50	30	209320.5847
7	29	120	20	1	1	50	30	205398.1635
8	29	120	20	1	1	50	30	192101.743
9	29	120	20	1	1	50	30	210000.7855
10	29	120	20	1	1	50	30	217441.4056
1	29	40	40	1	1	50	30	206866.4086
2	29	40	40	1	1	50	30	199983.9424
3	29	40	40	1	1	50	30	199737.1471
4	29	40	40	1	1	50	30	192422.037
5	29	40	40	1	1	50	30	205523.1883
6	29	40	40	1	1	50	30	194423.0627
7	29	40	40	1	1	50	30	202953.86
8	29	40	40	1	1	50	30	184594.2437
9	29	40	40	1	1	50	30	209230.9061
10	29	40	40	1	1	50	30	198813.1189
1	29	80	40	1	1	50	30	191585.2576
2	29	80	40	1	1	50	30	193555.0903
3	29	80	40	1	1	50	30	195211.407
4	29	80	40	1	1	50	30	193583.5997
5	29	80	40	1	1	50	30	193148.8733
6	29	80	40	1	1	50	30	212112.8844
7	29	80	40	1	1	50	30	192173.1654
8	29	80	40	1	1	50	30	198440.8303
9	29	80	40	1	1	50	30	198233.3475
10	29	80	40	1	1	50	30	188972.9987
1	29	120	40	1	1	50	30	199938.0221

ANEXO

2	29	120	40	1	1	50	30	194248.8605
3	29	120	40	1	1	50	30	183439.7664
4	29	120	40	1	1	50	30	185588.3223
5	29	120	40	1	1	50	30	188363.8535
6	29	120	40	1	1	50	30	182296.7599
7	29	120	40	1	1	50	30	200459.3042
8	29	120	40	1	1	50	30	189562.7641
9	29	120	40	1	1	50	30	198386.9982
10	29	120	40	1	1	50	30	197531.858
1	29	40	60	1	1	50	30	203313.7881
2	29	40	60	1	1	50	30	193154.5181
3	29	40	60	1	1	50	30	205938.2855
4	29	40	60	1	1	50	30	194010.9795
5	29	40	60	1	1	50	30	211868.5242
6	29	40	60	1	1	50	30	210815.2311
7	29	40	60	1	1	50	30	197614.2997
8	29	40	60	1	1	50	30	202972.0833
9	29	40	60	1	1	50	30	206702.3342
10	29	40	60	1	1	50	30	186250.8674
1	29	80	60	1	1	50	30	183382.7361
2	29	80	60	1	1	50	30	187244.7872
3	29	80	60	1	1	50	30	195694.158
4	29	80	60	1	1	50	30	186939.3235
5	29	80	60	1	1	50	30	178944.5606
6	29	80	60	1	1	50	30	194281.6915
7	29	80	60	1	1	50	30	190666.779
8	29	80	60	1	1	50	30	197886.226
9	29	80	60	1	1	50	30	201708.1109
10	29	80	60	1	1	50	30	192018.7206
1	29	120	60	1	1	50	30	206739.5149
2	29	120	60	1	1	50	30	191842.6344
3	29	120	60	1	1	50	30	193634.1419
4	29	120	60	1	1	50	30	175212.2812
5	29	120	60	1	1	50	30	210539.3798
6	29	120	60	1	1	50	30	184896.8634
7	29	120	60	1	1	50	30	191325.6637
8	29	120	60	1	1	50	30	182450.2677
9	29	120	60	1	1	50	30	200002.8282
10	29	120	60	1	1	50	30	197459.1419
1	29	40	20	1	1	50	50	193643.3598
2	29	40	20	1	1	50	50	209521.8855
3	29	40	20	1	1	50	50	198502.9397
4	29	40	20	1	1	50	50	205219.6572
5	29	40	20	1	1	50	50	200155.8878
6	29	40	20	1	1	50	50	211443.6313

ANEXO

7	29	40	20	1	1	50	50	186603.1472
8	29	40	20	1	1	50	50	214890.2041
9	29	40	20	1	1	50	50	176035.6758
10	29	40	20	1	1	50	50	201353.1199
1	29	80	20	1	1	50	50	201126.2448
2	29	80	20	1	1	50	50	192237.1818
3	29	80	20	1	1	50	50	204285.1366
4	29	80	20	1	1	50	50	186919.9333
5	29	80	20	1	1	50	50	188446.8693
6	29	80	20	1	1	50	50	176659.6128
7	29	80	20	1	1	50	50	205251.3068
8	29	80	20	1	1	50	50	205730.3695
9	29	80	20	1	1	50	50	190335.2341
10	29	80	20	1	1	50	50	203813.8151
1	29	120	20	1	1	50	50	165489.2981
2	29	120	20	1	1	50	50	196360.617
3	29	120	20	1	1	50	50	183043.7348
4	29	120	20	1	1	50	50	173803.6936
5	29	120	20	1	1	50	50	179877.8518
6	29	120	20	1	1	50	50	181654.2127
7	29	120	20	1	1	50	50	186736.9111
8	29	120	20	1	1	50	50	192603.8041
9	29	120	20	1	1	50	50	183313.5372
10	29	120	20	1	1	50	50	193817.8643
1	29	40	40	1	1	50	50	185989.7549
2	29	40	40	1	1	50	50	189942.0423
3	29	40	40	1	1	50	50	192367.9495
4	29	40	40	1	1	50	50	191527.8354
5	29	40	40	1	1	50	50	186545.1434
6	29	40	40	1	1	50	50	188141.4292
7	29	40	40	1	1	50	50	195207.3919
8	29	40	40	1	1	50	50	194582.7624
9	29	40	40	1	1	50	50	208127.2824
10	29	40	40	1	1	50	50	183725.3467
1	29	80	40	1	1	50	50	177845.0929
2	29	80	40	1	1	50	50	179492.7292
3	29	80	40	1	1	50	50	188802.971
4	29	80	40	1	1	50	50	185254.3432
5	29	80	40	1	1	50	50	178450.8221
6	29	80	40	1	1	50	50	181003.7481
7	29	80	40	1	1	50	50	167025.0321
8	29	80	40	1	1	50	50	184956.4259
9	29	80	40	1	1	50	50	171860.5555
10	29	80	40	1	1	50	50	170145.7029

ANEXO

1	29	120	40	1	1	50	50	143340.4934
2	29	120	40	1	1	50	50	173658.2511
3	29	120	40	1	1	50	50	171795.1904
4	29	120	40	1	1	50	50	177692.3414
5	29	120	40	1	1	50	50	175532.2285
6	29	120	40	1	1	50	50	190475.3005
7	29	120	40	1	1	50	50	187878.5522
8	29	120	40	1	1	50	50	175225.2615
9	29	120	40	1	1	50	50	155873.4217
10	29	120	40	1	1	50	50	176706.7196
1	29	40	60	1	1	50	50	192868.1612
2	29	40	60	1	1	50	50	191232.637
3	29	40	60	1	1	50	50	193780.316
4	29	40	60	1	1	50	50	192806.7681
5	29	40	60	1	1	50	50	195870.9583
6	29	40	60	1	1	50	50	180628.8606
7	29	40	60	1	1	50	50	190320.8876
8	29	40	60	1	1	50	50	188667.8568
9	29	40	60	1	1	50	50	189175.0263
10	29	40	60	1	1	50	50	184211.6583
1	29	80	60	1	1	50	50	183518.0555
2	29	80	60	1	1	50	50	190964.4866
3	29	80	60	1	1	50	50	191089.5674
4	29	80	60	1	1	50	50	179946.1223
5	29	80	60	1	1	50	50	188515.9691
6	29	80	60	1	1	50	50	184737.6592
7	29	80	60	1	1	50	50	174946.5792
8	29	80	60	1	1	50	50	185774.4904
9	29	80	60	1	1	50	50	181292.898
10	29	80	60	1	1	50	50	179096.0441
1	29	120	60	1	1	50	50	182117.0955
2	29	120	60	1	1	50	50	182195.6834
3	29	120	60	1	1	50	50	163614.2308
4	29	120	60	1	1	50	50	188950.6612
5	29	120	60	1	1	50	50	177366.8054
6	29	120	60	1	1	50	50	186161.5867
7	29	120	60	1	1	50	50	177424.2301
8	29	120	60	1	1	50	50	191323.9316
9	29	120	60	1	1	50	50	174799.6956
10	29	120	60	1	1	50	50	178875.5802

ANEXO



Glosario

Células blancas sanguíneas: Son la células que pertenecen al Sistema Inmune Natural.

Hematopoyesis: Generación de todas las células sanguíneas. Se produce principalmente en el esternón, las vértebras, los huesos ilíacos y las costillas.

Fagocitosis: En un proceso en el que se engloban moléculas grandes, para luego ser eliminadas.

Antígeno: Cualquier sustancia a la que se unen específicamente una molécula de anticuerpo o un receptor de la célula T. Estos pueden ser lípidos, hidratos de carbono complejos, los ácidos nucleicos, y las proteínas.

Afinidad: La afinidad de anticuerpo es una medida cuantitativa de la fuerza de unión. Es la fuerza combinada de las interacciones no covalentes entre un sitio de unión de antígeno único en un anticuerpo y un epítipo único. Los anticuerpos de afinidad baja se unen al antígeno de manera débil y tienden a disociarse con facilidad, en tanto que

GLOSARIO

los de afinidad alta unen con antígeno con más fuerza y la unión se mantiene por más tiempo

Avidez: Esta fuerza total de la unión

Sistema Inmune Artificial: Los sistemas inmunes artificiales son sistemas adaptativos, inspirados en la teoría, modelos y principios inmunológicos y las funciones inmunes observadas, que son aplicados a la solución de problemas"

GLOSARIO



Bibliografía

Abbas Abul K., Lichtman Andrew H., Pober Jordan S. *Inmunología Celular y Molecular*. Cuarta Edición. McGraw-Hill, ISBN 84-486-0405-9, 2002.

Aickelin Uwe and Cayzer Steve. *The Danger Theory and Its Application to Artificial Immune Systems*, Proceedings of the 1st Internat Conference on Artificial Immune Systems (ICARIS-2002), pp 141-148, Canterbury, UK, 2002.

Aickelin Uwe. *Artificial Immune Systems (AIS) – A New Paradigm for Heuristic Decision Making?*, Annual Operational Research Conference, Keynote Speech, 2004.

Akira Shizuo, Uematsu S, Takeuchi O. *Pathogen recognition and innate immunity*. Cell. Feb 24;124(4):783-801. Elsevier, 2006.

Anchor Kevin, Zydallis Jesse B., Gunsch Gregg H., and Lamont Gary B. *Extending the Computer Defense Immune System: Network Intrusion Detection with a*

BIBLIOGRAFÍA

Multiobjective Evolutionary Programming Approach. In Jonathan Timmis and Peter J. Bentley, editors, First International Conference on Artificial Immune Systems (ICARIS-2002), University of Kent at Catenbury, UK, September 2002.

Axelsson S. The Base-Rate Fallacy and its Implications for the Difficulty of Intrusion Detection. 6th ACM Conference on Computer and Communications Security, 1999.

Axelsson S. "Intrusion detection systems: A survey and taxonomy," Tech. Rep. 99-15, Department of Computer Engineering, Chalmers University, 2000.

Bakhouya M, Gaber J. An immune Inspired-based optimization Algorithm: application to traveling salesman problem. AMO- Advanced Modeling and Optimization, Volume 9 Number 1, pp 105-116, 2007.

Barbará, D., Wu, N., Jajodia, S. Detecting Novel Network Intrusions Using Bayes Estimators. First SIAM Conference on Data Mining, 2001.

Bector C. R., Chandra S., Dutta J., Principles of Optimization Theory, Ed. Alpha Science, pp. 1,4-5, ISBN 1-84265-166-8, 2005.

Berek, C. & Ziegner, M. The Maturation of the Immune Response, Imm. Today, 14(8), pp. 400-402. 1993.

Bridges, S. M. and Vaughn, R. B. Intrusion Detection via fuzzy Data Mining. 12th Annual Canadian Information Technology Security Symposium. 2000.

Burnet FM. A modification of Jerne's theory of antibody production using the concept of clonal selection. Australian Journal of Science 20:67-69, 1957.

BIBLIOGRAFÍA

Cansian, A. M., S. Moreira, E., Carvalho, A., and Bonifácio Jr, J. M. Network Intrusion Detection using Neural Networks. International Conference on Computational Intelligence and Multimedia Applications (ICCMA'97), 1997.

Cortés Rivera Daniel, Un Sistema Inmune Artificial para resolver el problema del Job Shop Sheduling, Tesis de Maestría, CINVESTAV, g Marzo 2004.

Cruz Cortés Nareli, Sistema inmune artificial para solucionar problemas de optimización. Tesis Doctoral CINVESTAV, Octubre 2004.

Dasgupta, D. An overview of artificial immune systems. Artificial Immune Systems and Their Applications. pp. 3-19. Pub: Springer-Verlag.. 1998

Dipankar Dasgupta, editor. Artificial Immune Systems and Their Applications. Springer-Verlag, Berlin, 1999.

De Castro Leandro, Von Zuben José Fernando. Artificial Immune Systems: Part I – Basic Theory Of Applications, 1999.

De Castro Leandro, Von Zuben José Fernando. Artificial Immune Systems: Part II – A Survey of Applications , 2000.

De Castro Leandro Nunes, An Introduction to the Artificial Immune Systems ICANNGA 2001, April, 2001.

De Castro Leandro Nunes and Timmis Jon. An artificial immune network for multimodal function optimization. In Proceedings of the special sessions on artificial

BIBLIOGRAFÍA

immune systems in the Congress on Evolutionary Computation, IEEE World Congress on Computational Intelligence, volume I, pages 669_674, Honolulu, Hawaii, May 2002.

De Castro, L.N and Timmis, J. Artificial Immune Systems: A New Computational Intelligence Approach. Springer-Verlag. London, September. ISBN 1-85233-594-7, 2002a.

De Castro, Von Zuben. Learning and Optimization Using the Clonal Selection Principle IEEE Transactions on Evolutionary Computation, Special Issue on Artificial Immune Systems, Vol. 6, n. 3, pp 239-251, 2002b.

De Castro L. N., Timmis J. I. Artificial immune systems as a novel soft computing paradigm. Soft Computing 7, 526–544 Springer-Verlag, 2003.

Debar, H., Becker, M., Siboni, D. A Neural Network Component for an Intrusion Detection System. IEEE Symposium on Research in Computer Security and Privacy, 1992.

Denning, D. E. An Intrusion-Detection Model. IEEE transaction on Software Engineering, 1987.

Dickerson, J.E., Juslin, J., Koukousoula, O., Dickerson, J.A. Fuzzy Intrusion Detection. North American Fuzzy Information Processing Society (NAFIPS) International Conference, 2001.

Dréo Johann, Pétrowsky Alain, Siarry Patrick, Taillard Eric. Metaheuristics for Hard Optimization, Methods and Case Studies. Springer-Verlag Berlin Heidelberg. ISBN: 13 978-3-540-23022-9, 2006.

BIBLIOGRAFÍA

Ertoz, L., Steinbach, M., Kumar, V. Finding Clusters of Different Sizes, Shapes, and Densities in Noisy, High Dimensional Data. SIAM International Conference on Data Mining (SDM '03), 2003.

Ertoz, L., Eilertson, E., Lazarevic, A., Tan, P., Srivastava, J., Kumar, V., Dokas, P. The MINDS - Minnesota Intrusion Detection System. Next Generation Data Mining, 2004.

Eskin, E., Arnold, A., Preraua, M., Portnoy, L., Stolfo, S. A Geometric Framework for Unsupervised Anomaly Detection: Detecting Intrusions in Unlabeled Data. Applications of Data Mining in Computer Security. Kluwer Academic Publishers, 2002.

Farmer J. D., Packard N. H. and Perelson A. S.. The immune System, Adaptation and Machine Learning. Physica D, 22:187_204, 1986.

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. The KDD Process for Extracting Useful Knowledge from Volumes of Data. Communications of the ACM, 1996.

Fernández Jeri Leoncio, Teoría Básica de Optimización, Universidad Agraria la Molina, Lima Perú, 1999.

Forrest Stephanie, Perelson Alan S., Lawrence Allen, Cherukuri Rajesh. Self- Nonself Discrimination in a Computer. In Proceedings of the IEEE Symposium on Research in Security and Privacy, Los Alamitos, CA, IEEE Computer Society Press. 1994.

Forrest, S., Hofmeyr, S., Somayaji, A., Longstaff T., A Sense of Self for UNIX Processes. IEEE Symposium on Computer Security and Privacy, 1996.

Fox, K., Henning, R., Reed, J., Simonian, R. A. Neural Network Approach Towards Intrusion Detection. 13th National Computer Security Conference, 1990.

BIBLIOGRAFÍA

Gómez, J., Dasgupta, J. Evolving Fuzzy Rules for Intrusion Detection IEEE Information Assurance Workshop, 2002.

Gowans J.L.. The Lymphocyte a disgraceful gap in medical knowledge Immunol. Today. 17: 288-291. (PubMed), 1996.

Guzella T. S, Mota-Santos T. A., and Caminhas W. M.. Regulatory T cells: Inspiration for artificial immune systems. In. Artificial Immune Systems, 6th International Conference, ICARIS, volume 4628 of Lecture Notes in Computer Science. Springer, pages 312–323, Editores Leandro Nunes de Castro, Fernando Jose Von Zuben, and Helder Knidel. 2007.

Hajela P. and Lee J. Constrained Genetic Search via Schema Adaptation. An Immune Adaptation. An Immune Network Solution. In Niels Olhoff and George I. N. Rozvany, editors, Proceedings of the First World Congress of Structural and Multidisciplinary Optimization, pages 915-920, Goslar, Germany, Pergamon, 1996.

Hajela Prabhat and Yoo Sun Jun. Immune network modeling in design optimization. In D. Corne, M. Dorigo, and F. Glover, editors, New Ideas in Optimization, pages 167-183. Mc Graw-Hill, 1999.

Heady, R., Luger, G., Maccabe, A., Servilla, M. The Architecture of a Network Level Intrusion Detection System. Technical report, Department of Computer Science, University of New Mexico, August 1990.

Hofmeyr, S., Forrest, S. Architecture for an Artificial Immune System. Evolutionary Computation, 2000.

BIBLIOGRAFÍA

Höglund, A., Hätönen, K., and Tuononen, T. A UNIX Anomaly Detection System Using Self-Organising Maps. Recent Advances in Intrusion Detection, RAID 1998.

Honig, A., Howard, A., Eskin, E., Stolfo, S. Adaptive Model Generation: An Architecture for the Deployment of Data Mining-based Intrusion Detection Systems. Applications of Data Mining in Computer Security. Kluwer Academic Publishers, 2002.

Hoos Holger H., Stützle Thomas, Stochastic Local Search. Foundations and Applications, Ed. Elsevier, Capítulo 1, ISBN 1-55860-872-9, 2005.

Jerne N K. The natural-selection theory of antibody formation. Proceedings of the National Academy of Sciences USA 41:849-857,1955

Janemay Charles A. Jr, Travers Paul, Walport Mark, Shlomchik Mark J. Immunobiology Garland Publisihing, 5a Edición. 2001.

Johnson David S., Garey Michael R.. Computers and Intractability: A Guide to the Theory of NP-Completeness (Series of Books in the Mathematical Sciences). W H Freeman & Co., ISBN 0-7167-1045-5, June 1979.

Kindt Thomas J., Goldsby Richard J., Osborne Barbara A. Inmunología de Kuby. Sexta edición. ISBN 13: 978-0-7167-8590-3. 2007.

Knowles J. and Corne D. Approximating the Nondominated Front Using the Pareto Archived Evolution Strategy. Evolutionary Computation, 8(2):149-172, 2000.

Kruegel, C., Toth, T., Kirda, E. Service Specific Anomaly Detection for Network Intrusion Detection. Symposium on Applied Computing (SAC), 2002.

BIBLIOGRAFÍA

Lane, T., Brodley, C.E. Temporal Sequence Learning and Data Reduction for Anomaly Detection. ACM Transactions on Information and System Security, 1999

Lee, W., Stolfo, S. Data Mining Approaches for Intrusion Detection. Seventh USENIX Security Symposium, 1998.

Lee, W., Stolfo, S., Mok, K. A Data Mining Framework for Building Intrusion Detection Models. IEEE Symposium on Security and Privacy, 1999.

Levin, I. KDD-99 Classifier Learning Contest LLSOFT's Results Overview. SIGKDD Explorations, ACM SIGKDD, 2000.

Li, W., Using Genetic Algorithm for Network Intrusion Detection. Proceedings of the United States Department of Energy Cyber Security Group, 2004.

Litman, G.W.. "Origen de la inmunidad de los vertebrados". Investigación y Ciencia, Enero. 1997.

<http://www.uco.es/grupos/inmunologia-molecular/inmunologia/inmunologia/tema16/etexto16.htm>

Luch G. C., Chieh C. H., and Liu W. W.. MOIA: Multi-Objective Immune Algorithm. Engineering Optimization, 35(2):143-164, 2003.

Matzinger P, The Danger Model in Its Historical Context, Scandinavian Journal of Immunology, 54: 4-9, 2001.

Matzinger P, Tolerance, Danger and the Extended Family, Annual Review of Immunology, 12:991 - 1045, 1994.

BIBLIOGRAFÍA

Matzinger Polly. The danger model: A renewed sense of self. *Science*, 296(1):301–305, 2002.

Mé, L. Genetic Algorithms, a Biologically Inspired Approach for Security Audit Trails Analysis. *IEEE Symposium on Security and Privacy (SSP)*, 1996.

Mé, L. GASSATA, A Genetic Algorithm as an Alternative Tool for Security Audit Trails Analysis. *Recent Advances in Intrusion Detection, RAID 1998*.

Mukkamala S., Janoski G., Sung A. H. Intrusion Detection Using Neural Networks and Support Vector Machines. *IEEE International Joint Conference on Neural Networks*, 2002.

Newborough John, Stepney Susan. A generic framework for population-based algorithms, implemented on multiple FPGAs. In. *Artificial Immune Systems, 4th International Conference, ICARIS*, volume 3627 of *Lecture Notes in Computer Science*. Springer, pages 43–55. Editores Christian Jacob, Marcin L. Pilat, Peter J. Bentley, y Jon Timmis. 2005.

NIST URL: <http://csrc.nist.gov>.

Orfila, A., Carbo, J., Ribagorda, A. Fuzzy Logic on Decision Model for IDS. *IEEE International Conference on Fuzzy Systems*, 2003.

Perelson, A. S., Mirmirani, M. & Oster, G. F.. Optimal Strategies in Immunology II. B Memory Cell Production, *J. Math. Biol.*, 5, pp. 213-256. 1978.

Pfahring, B. Winning the KDD99 Classification Cup: Bagged Boosting. *SIGKDD Explorations, ACM SIGKDD*, 2000.

BIBLIOGRAFÍA

Portnoy, L., Eskin, E., Stolfo, S. Intrusion Detection with Unlabeled Data Using Clustering. ACM Workshop on Data Mining Applied to Security (DMSA), 2001.

Ramadas, M., Ostermann, S., Tjaden, B. Detecting Anomalous Network Traffic with Self-Organizing Maps. Recent Advances in Intrusion Detection, RAID 2003.

Roitt, I. Essential Immunology: 9th Edition. Chap. Specific Acquired Immunity, pp. 22-39. Pub. Blackwell Science. 1997.

Rojas Espinoza Oscar. Inmunología (de memoria). Editorial médica panamericana. 2ª edición. ISBN 968-7988-28-2. 2001.

Salazar González Juan José, Optimización Combinatoria Poliédrica: problemas de rutas de localización, Tesis Doctoral, Universidad de la Laguna, 1996.

Shangce Gao, Zheng Tang, Hongwei Dai, Yang Gang. A Simulated Annealing PolyClonal Selection Algorithm and its Application to Traveling Salesman Problem. Neural Information Processing- Letters and Review, vol. 11, No. 7, pp 159-164, 2007.

Stepney Susan, Smith Robert E., Timmis Jon, and Tyrrell Andy M.. Towards a conceptual framework for artificial immune systems. In Artificial Immune Systems, Third International Conference, ICARIS, volume 3239 of Lectures Notes in Computer Science, Springer, pages 53-64, Editors Nicosia Giuseppe et al. 2004.

Stepney Susan, Smith Robert E., Timmis Jonathan, Tyrrell Andy M., Neil Mark J., Hone Andrew N. W. Conceptual frameworks for artificial immune systems. International Journal of Unconventional Computing, 1(3): 315-338, 2005.

BIBLIOGRAFÍA

Teng, H., Chen, K., Lu, S. Adaptive Real-Time Anomaly Detection Using Inductively Generated Sequential Patterns. IEEE Computer Society Symposium on Research in Security and Privacy. 1990.

Timmis J. and Neal. M. Investigating the evolution and stability of a resource limited artificial immune system. In Proc. of the GECCO, Special Workshop on Artificial Immune System and Their Applications, pages 40-41, AAAI press, 2000.

Timmis J., Knight T., de Castro L.N. and Hart E., An Overview of Artificial Immune Systems, 2002.

Timmis, J, de Lemos, R, Ayara, M and Duncan R. Towards Immune Inspired Fault Tolerance in Embedded Systems. To appear in the Proceedings of International Conference on Neural Information Processing. Singapore. November 2002a.

Timmis J., Knight T., De Castro L. N., and Hart E.. In R Paton, H Bolouri, M Holcombe, J H Parish, and R Tateson, editors, *"Computation in Cells and Tissues: Perspectives and Tools for Thought"*, Natural Computation Series, pages 51-86. Springer, November 2004.

Tizzard, I. (1988a). Immunology: An Introduction. 2nd edition. Chap. The response of B-cells to Antigen. Pp. 199-223. Pub. Saunders College.

TSPLIB. Ruprecht-Karls-Universität Heidelberg. Discrete and Combinatorial Optimization. <http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/>.

Twycross Jamie and Aickelin Uwe. Towards a conceptual framework for innate immunity. In Artificial Immune Systems, 4th International Conference, ICARIS, volume

BIBLIOGRAFÍA

3627 of Lecture Notes in Computer Science. Springer, pages 112–125. Editores Christian Jacob, Marcin L. Pilat, Peter J. Bentley, y Jon Timmis. 2005.

Valdés A., Skinner, K. Adaptive, Modelbased Monitoring for Cyber Attack Detection. Recent Advances in Intrusion Detection, RAID 2000.

Varela, F, Coutinho, A, Dupire, B and Vaz, N. Cognitive Networks : Immune and Neural and Otherwise. Theoretical Immunology: Part Two, SFI Studies in the Sciences of Complexity, 2, pp.359-371. 1988.

Warrender, C., Forrest, S., Pearlmutter, B. Detecting Intrusions Using System Calls: Alternative Data Models. IEEE Symposium on Security and Privacy, 1999.

Ye, N., Li, X., Emran, S. M. Decision Tree for Signature Recognition and State Classification. IEEE Systems, Man, and Cybernetics Information Assurance and Security Workshop, 2000.

Yoo J. and Hajela P. Immune network simulations in multicriterion design. Structural Optimization, 18:85_94, 1999.

Zhu Yunyi, Gao Shangce, Dai Hongwei, Li Fangjia and Tang Zheng. An Improved Clonal Algorithm and Its Application to Traveling Salesman Problem. IJCSNS International Journal of Computer Science and Network Security, Vol. 7 No. 8, pp 109-113. 2007.

Zurutuza Urko, Uribetxeberria Roberto. Revisión del estado actual de la investigación en el uso de data mining para la detección de intrusiones. In Actas del I Simposio sobre Seguridad Informática (SSI). I Congreso Español de Informática (CEDI), Granada, Spain. September 2005.