

# LOS ALGORITMOS GENÉTICOS Y EL MÉTODO DE GENERACIÓN Y PRUEBA

*Luis Roberto Ojeda Ch.  
M.S. in Computer Science U.C.L.A.  
Profesor Asociado  
Departamento de Ingeniería de Sistemas  
Universidad Nacional de Colombia, Santafé de Bogotá*

## INTRODUCCIÓN

Este trabajo pretende mostrar dos paradigmas de solución de problemas en inteligencia artificial. Los algoritmos genéticos han tomado una posición destacada en los últimos tiempos y el método de generación y prueba, cuyo perfil se aproxima al fundamento de los algoritmos genéticos constituye la forma de diseño de DENDRAL, un producto ampliamente reconocido en IA. Se desea mostrar que ambos paradigmas se ubican naturalmente dentro de las expectativas de la inteligencia artificial.

## I. INTELIGENCIA ARTIFICIAL E INTELIGENCIA HUMANA.

Es bien difícil aproximarse directamente a una concepción de lo que es inteligencia en el ser humano. Nos ceñiremos a la definición aportada por un autor reconocido (Philip Jackson, *Introduction to Artificial Intelligence*) quien después de explorar el diccionario Webster establece que inteligencia es la habilidad para actuar correctamente en una situación dada.

Si aceptamos lo anterior como la expectativa fundamental de la inteligencia artificial, debemos pensar que en el computador:

- Debe existir un conocimiento de la situación en términos de una descripción formal.
- Deben estar las leyes que modelan adecuadamente lo que puede ser una actuación correcta, referidas al formalismo de representación.
- Debe existir un mecanismo aplicativo de tales leyes.

## II. ALGORITMOS GENÉTICOS.

Un *algoritmo genético* es un proceso iterativo que mantiene una *población* de estructuras candidatas a ser solución para un problema.

En cada momento, que se denomina una *generación*, las estructuras en la población actual son calificadas de acuerdo con su acercamiento a la solución y con base en estas evaluaciones se forma una nueva población de candidatas usando *operadores genéticos* específicos, tales como *reproducción*, *cruzamiento* y *mutación*. En los algoritmos genéticos las leyes que rigen la situación y el mecanismo aplicativo para ellas es copiado de la naturaleza.

Los operadores primarios se definen así:

### A. Reproducción.

Los algoritmos genéticos producen nuevas generaciones mejoradas seleccionando padres con más ajuste a la solución, o dándoles a tales padres eventuales más probabilidad de efectivamente llegar a serlo.

### B. Cruzamiento.

Muchos algoritmos genéticos usan cadenas binarias para representar candidatas. El cruzamiento escoge al azar una posición en la cadena (por ejemplo después de dos dígitos) e intercambia segmentos ya sea a la

derecha o a la izquierda de este punto con otra cadena igualmente seccionada.

### C. Mutación

La mutación es un cambio arbitrario y permite sacar al algoritmo de un callejón sin salida. El procedimiento cambia el uno por el cero o un cero por un uno en lugar de duplicar. La mutación tiene una probabilidad muy baja (por ejemplo, uno en mil).

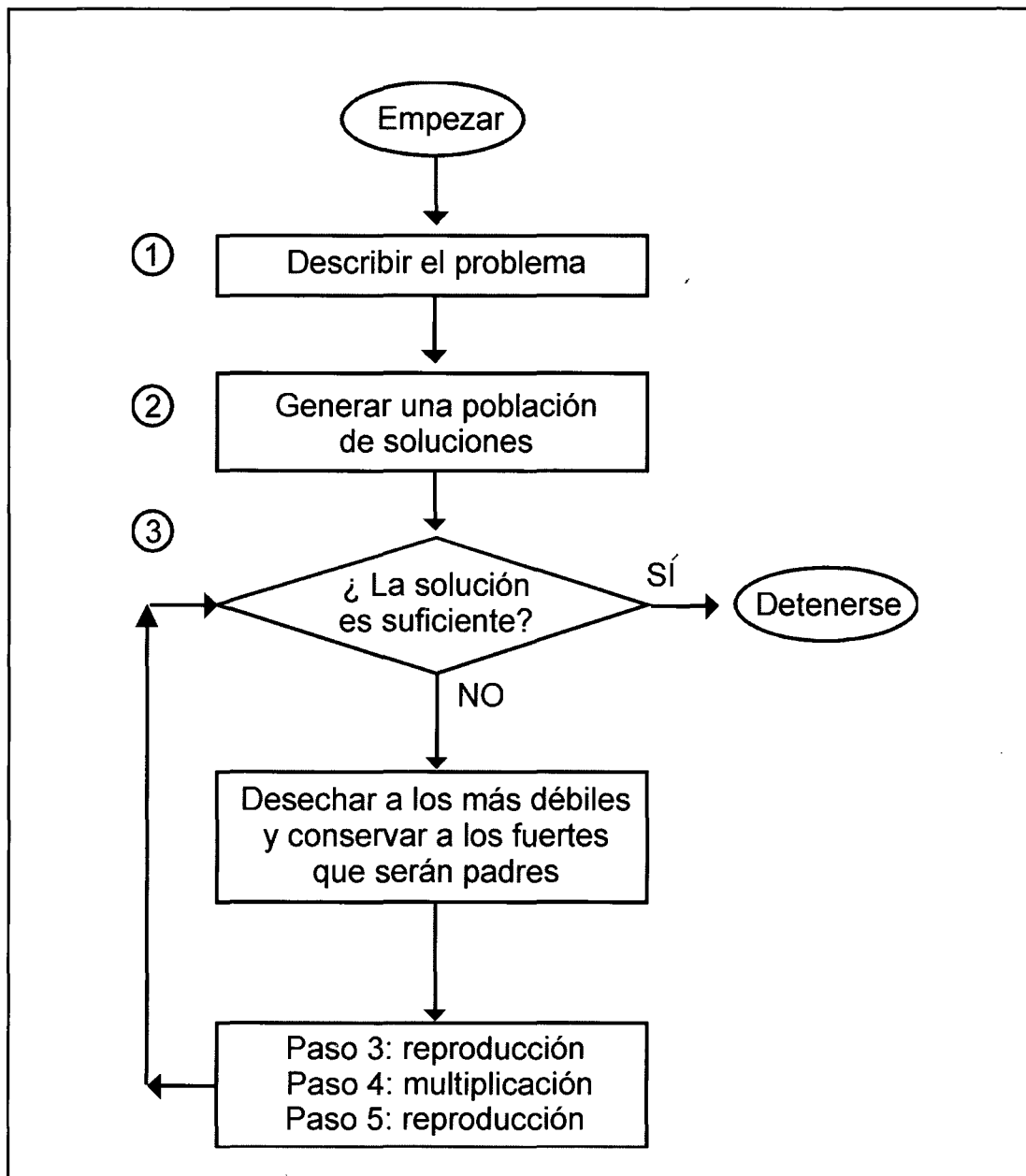
## III. APLICACIONES

Los algoritmos genéticos son una forma de aprendizaje de maquina. (Austin S. "Genetic Solution of XOR Problems", *AI Expert*, Dic de 1990) indica algunas áreas de aplicaciones:

- Control dinámico.
- Modelos biológicos.
- Diseño en ingeniería.
- Identificación de patrones.

Un algoritmo genético puede recibir información que le permite rechazar soluciones inferiores y acumular buenas soluciones.

Recolectamos en un diagrama de flujo el procedimiento seguido, que es la infraestructura básica de trabajo de los algoritmos genéticos.



**A. Un ejemplo simple de optimización.**

Se va a optimizar la estrategia de una empresa comercial que vende un sólo producto; la empresa está compuesta de cuatro sucursales.

La estrategia implica la toma de tres decisiones que pueden codificarse binariamente así:

Código binario	Variable		
	Descuento	Transporte	Forma de pago
0	No se hace	Por la empresa	Contado
1	5%	Por el cliente	1 a 30 días

El cuadro anterior debe entenderse así:

Una cadena 000 indica: no se hace descuento; el transporte es por la empresa y la forma de pago es de contado.

Una cadena 101 indica: el descuento es de 5%; el transporte es por la empresa y la forma de pago es 1 a 30 días.

Cada variable: descuento, transporte y forma de pago reciben individualmente un código de 0 o de 1 y la cadena de variables constituye una estrategia.

Se trata de combinar las tres decisiones de modo que el beneficio resulte máximo.

Hay  $(2)^3$ , o sea 8 posibles estrategias.

La única información disponible para cada estrategia es el beneficio mensual obtenido por cada sucursal.

No se sabe cuál de las tres variables, descuento, transporte o forma de pago es la más importante, ni la forma como las variables pueden relacionarse para construir la función beneficio.

Por otra parte, las condiciones de mercadeo en el medio no garantizan que la situación sea estable, de modo que una estrategia que funciona razonablemente bien en un momento dado puede dejar de funcionar en un futuro no lejano.

El gerente debe tomar la decisión de cómo deben funcionar las sucursales a partir del día siguiente. Él decide probar aleatoriamente una estrategia inicial para cada una de las cuatro sucursales así:

Sucursal	Estrategia
1	011
2	001
3	110
4	010

El primer dígito binario por la izquierda representa a la variable descuento, el siguiente a la variable transporte y el de más a la derecha forma de pago.

El gerente está procediendo como lo haría un algoritmo genético: se ha generado una población aleatoriamente, de tamaño 4.

Para cada generación en la que se ejecutaría el algoritmo genético, cada individuo se prueba para establecer su idoneidad.

La idoneidad de un individuo en este caso estará dada por el beneficio reportado a la empresa.

Como no disponemos de datos reales de beneficio, puesto que la compañía es ficticia, ilustramos con el beneficio que en cada caso tomamos (ingenuamente) como el valor decimal de la cadena binaria.

Sucursal	Estrategia	Beneficio
1	011	3
2	001	1
3	110	6
4	010	2
		Total 12
		Peor 1
		Promedio 3
		Mejor 6

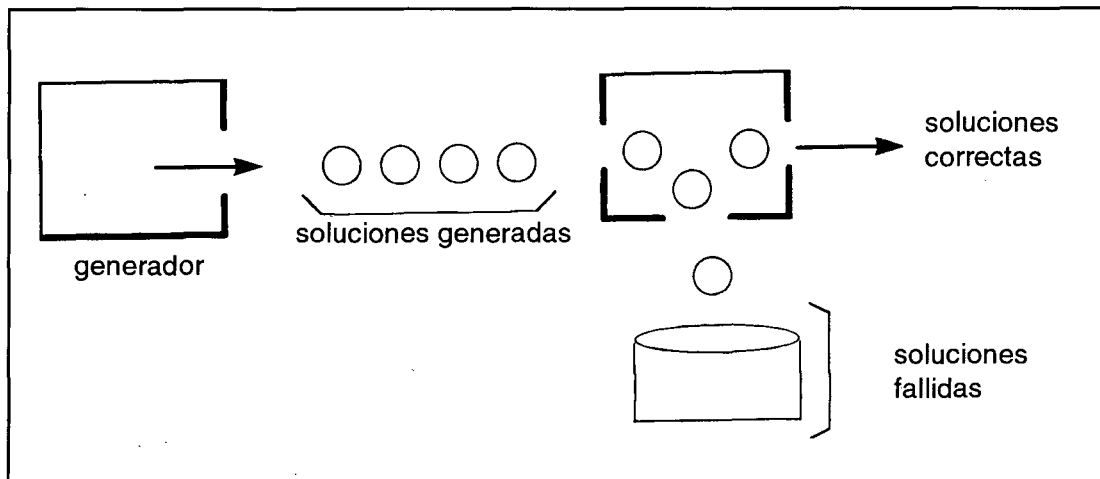
La única información en el proceso genético está constituida por los valores directamente observados en los rendimientos (beneficios) de los individuos actualmente presentes en la población.

El proceso genético transforma una población en otra nueva mediante operaciones basadas en el principio de Darwin de reproducción y supervivencia del más idóneo.

*El proceso genético es identificable como un proceso de aprendizaje en el que el computador va perfeccionando su comportamiento (estrategia) frente al medio (mercado), utilizando para el efecto los valores de beneficio ligados con cada estrategia realizada. Esto constituye un proceso de aprendizaje.*

### B. El método de generación y prueba

El método de generación y prueba obedece a un paradigma que puede representarse gráficamente así:



El *generador* enumera todas las soluciones posibles.

El *probador* rechaza las soluciones incorrectas y separa las soluciones aceptables.

El funcionamiento más común del sistema establece la prueba de cada solución cuando ella acaba de ser generada. Un algoritmo esquemático de generación y prueba es así :

Para realizar generación y prueba

- Hasta que no se puedan generar más soluciones.
- Generar una solución.
- Probar la solución generada.
- Si la solución es correcta, coleccionarla.

En este método, las leyes que enmarcan la situación están dadas por el conjunto de limitaciones que regulan la generación y por el mecanismo de detección de soluciones erróneas.

El modelo de generación y prueba es utilizado más frecuentemente para resolver problemas de identificación. Deben cumplirse siempre tres condiciones en todo método de generación y prueba:

- *Debe ser completo:* debe estar en capacidad de generar todas las soluciones posibles.

- *No debe ser redundante:* de ninguna manera puede generar una misma solución dos veces.

- *Debe estar informado:* la información se encamina a limitar el número de soluciones propuestas para prueba, sin excluir la posibilidad de generar soluciones factibles.

**Un ejemplo de generación y prueba**

Boronoff, Pavlov, Revitzky y Sukarek son cuatro talentosos artistas, uno es bailarín, otro pintor, otro cantante y otro escritor, pero no necesariamente en ese orden.

**Condiciones**

1. Boronoff y Revitzky estaban en el auditorio cuando el cantante hizo su estreno.
2. Tanto Pavlov como el escritor han sido pintados por el pintor.
3. El escritor, cuya biografía de Sukarek fue un éxito, planea escribir una biografía de Boronoff.
4. Boronoff nunca ha oído hablar de Revitzky.

¿Quién es cada uno?

**Notas :**

- a. Cuando un pintor hace un retrato conoce al personaje (sabe su nombre).
- b. No hay autorretrato ni hay autobiografía.

Se trata de asignar los nombres :

B: Boronoff            S: Sukarek  
P: Pavlov              R: Revitzky

a un vector de profesiones con cuatro elementos:

1: Cantante	3: Escritor	1	2	3	4
2: Pintor	4: Bailarín	C	P	E	B

Hay cuatro condiciones de prueba (y asignación):

1.  $\text{prof}[1] \llcorner 'B'$
2.  $\text{prof}[1] \llcorner 'R'$

Lo anterior en cumplimiento de la primera condición. Esto constituye información para la fase de generación: no se debe intentar ninguna solución con tales características.

La segunda condición establece:

3.  $\text{prof}[2] \llcorner 'P'$
4.  $\text{prof}[3] \llcorner 'P'$

Pavlov no es escritor y Pavlov no es pintor. Se adiciona ésta información a la fase de generación.

La condición 3 establece:

5.  $\text{prof}[3] \llcorner 'S'$
6.  $\text{prof}[3] \llcorner 'B'$

Se adicionan estos elementos a la información en la fase de generación.

La condición 4 va a relacionarse con la condición 2 así:

Pavlov y el pintor se conocen entre sí.  
El escritor y el pintor se conocen entre sí.

Esto debido a la condición 2.

Por tanto, debido a la condición 4:

<u>escritor</u>	<u>pintor</u>
Boronoff	Revitzky
Revitzky	Boronoff

no es posible.

En otros términos:

- (i)  $\text{prof}[3] = 'B' \wedge \text{prof}[2] = 'R'$   
no deberán ser simultáneos.

Tampoco deben serlo:

- (ii)  $\text{prof}[3] = 'R' \wedge \text{prof}[2] = 'B'$

Tomaremos (i) y (ii) para construir el mecanismo de prueba.

La generación debe ahora ceñirse a:

	1	2	3	4
	C	P	E	B
prof				
	S			
	P			
	1	2	3	4
	C	P	E	B
prof				
		S	S	
		B	B	
		R	R	

Al existir una sola posibilidad para R (Revitzky), antes de generar sabemos que Revitzky es escritor. Esto limita la generación en definitiva a las posibilidades siguientes :

	1	2	3	4
	C	P		B
prof			X	
	S	S		S
	P	B		P
				B

Generamos ahora todas las soluciones factibles :

		<u>C</u>	<u>P</u>	<u>B</u>
(a)	S	B	P	
(b)	P	S	B	
(c)	P	B	S	

Sucesivamente verifiquemos las soluciones

a.		1	2	3	4
		C	P	E	B
prof		S	B	R	P

(i)  $\text{prof}[3] = 'B' \wedge \text{prof}[2] = 'R'$

no deben ser simultáneos, y no lo son.

(ii)  $\text{prof}[3] = 'R' \wedge \text{prof}[2] = 'B'$

no deben ser simultáneos, pero lo son. Por lo tanto se rechaza la solución.

b.

	1	2	3	4
	C	P	E	B
prof	P	S	R	B

(i)  $\text{prof}[3] = 'B' \wedge \text{prof}[2] = 'R'$

no son simultáneos

(ii)  $\text{prof}[3] = 'R' \wedge \text{prof}[2] = 'B'$

no son simultáneos. Por lo tanto, *la solución es aceptable.*

c.

	1	2	3	4
	C	P	E	B
prof	P	B	R	S

(i)  $\text{prof}[3] = 'B' \wedge \text{prof}[2] = 'R'$

no son simultáneos

(ii)  $\text{prof}[3] = 'R' \wedge \text{prof}[2] = 'B'$

no deben ser simultáneos. Por lo tanto se rechaza la solución.

## CONCLUSIONES

Los algoritmos genéticos toman candidatos a solución sin normas restrictivas muy precisas. En el método de generación y prueba van planteándose una para unas formas eventuales de solución que se establecen a través de información suministrada para el efecto. Los algoritmos genéticos operan sobre los candidatos a solución, de forma que ellos son calificados por la función de aceptación y se establece un marco para producir nuevos candidatos, fundamentados en los anteriores. En el método de generación y prueba un candidato actual no depende de uno anterior, pues no existen los mecanismos de cruce, reproducción ni mutación. Es un punto de interés y de investigación delimitar fronteras de aplicación entre ambos métodos

## **BIBLIOGRAFÍA**

1. AUSTIN, S. "Genetic Solution of XOR problems" *AI Expert*, Dic. 1990.
2. JACKSON, Philip. *Introduction to Artificial Intelligence*. Second Edition, Dover Publications Inc. 1990.
3. SIERRA, Guillermo, *et al. Sistemas Expertos en contabilidad y administración de empresas*. Addison-Wesley Interamericana. EE.UU. 1995.
4. WINSTON, Patrick Henry. *Inteligencia artificial*. Tercera edición. Addison Wesley EE.UU. 1992.