

Inteligencia Computacional

Teorema de esquemas II

Blanca A. Vargas-Govea

<http://blancavg.com>

vargasgovea@itesm.mx

21/Sep/2012

Índice

1. Antecedentes	1-1
2. Objetivo	1-2
3. Efectos de la selección	1-2
4. Efectos del cruce	1-7
5. Efectos de la mutación	1-11
6. Desventajas del teorema	1-13
7. Ejercicios	1-14
8. Tarea: Martes 25 Sep	1-15
Referencias	1-16

1. Antecedentes

Hasta el momento se ha visto por separado

1. Algoritmo genético básico
2. Esquemas y sus propiedades

Aunque los algoritmos genéticos manipulan directamente una población de cadenas, con los esquemas nos damos cuenta de que este procesamiento explícito realmente causa un procesamiento implícito de muchos esquemas durante cada generación (Goldberg, 1989; Michalewicz, 1996)

Los esquemas y sus propiedades son instrumentos de notación para analizar con rigor y clasificar similitudes en las cadenas. Además, proporcionan los medios para analizar el efecto de la selección, el cruce y la mutación en los bloques de construcción contenidos en la población.

2. Objetivo

Analizar los efectos de las operaciones de selección, cruce y mutación en el crecimiento y decaimiento de los esquemas contenidos en la población.

3. Efectos de la selección

- pop_size : tamaño de la población
- m : longitud de la cadena = longitud del esquema

Sea $\xi(S, t)$ el número de cadenas en la población en el tiempo t descritas por el esquema S

La aptitud de un esquema en el tiempo t , $eval(S, t)$ se define como la aptitud promedio de todas las cadenas descritas por el esquema S . Supongamos que hay p cadenas $(c_{j_1}, \dots, c_{i_p})$ descritas por el esquema S en el tiempo t . Entonces:

$$eval(S, t) = \sum_{i=1}^p \frac{eval(c_{i_j})}{p}$$

Durante la selección, cada cadena se copia 0 o más veces de acuerdo a su aptitud. La probabilidad de que una cadena c_i sea seleccionada se expresa como

$$p_i = \frac{eval(c_i)}{F(t)}$$

y $F(t)$ es la aptitud total de la población en el tiempo t dada por:

$$F(t) = \sum_{i=1}^{pop-size} eval(c_i)$$

Después de la selección, esperamos tener $\xi(S, t+1)$ cadenas descritas por el esquema S . Dado que:

1. La probabilidad de seleccionar una cadena descrita por un esquema S es $\frac{eval(S, t)}{F(t)}$,

2. El número de cadenas descritas por un esquema S es $\xi(S, t)$ y
3. El número de cadenas seleccionadas es pop_size

se tiene que el número de cadenas descritas por el esquema S en el tiempo $t + 1$ está dado por:

$$\xi(S, t + 1) = \xi(S, t) pop_size \frac{eval(S, t)}{F(t)}$$

Tomando en cuenta que la aptitud promedio de la población es

$$\overline{F(t)} = \frac{F(t)}{pop_size}$$

se puede escribir como

$$\xi(S, t + 1) = \xi(S, t) \frac{eval(S, t)}{F(t)}$$

el número esperado de cadenas descritas por el esquema S en $t + 1$ está dado por el número de cadenas descritas por el esquema S en el tiempo t , la aptitud de un esquema en t y la aptitud promedio de la población.

- Esquemas con valores de aptitud mayores que la aptitud promedio de la población recibirán un incremento de cadenas en la siguiente generación. Crecen.
- Esquemas con valores de aptitud menores que la aptitud promedio de la población recibirán un decremento de cadenas en la siguiente generación. Desaparecen.

Este comportamiento se ejecuta para cada esquema en una población particular en paralelo.

Si suponemos que un esquema S permanece sobre el promedio con una constante ϵ (i.e., $eval(S, t) = \overline{F(t)} + \epsilon \overline{F(t)}$), entonces iniciando con $t = 0$ y suponiendo un valor estacionario de ϵ

$$\xi(S, t) = \xi(S, 0)(1 + \epsilon)^t$$

es una progresión geométrica lo que significa que no solamente un esquema arriba del promedio recibe un número incremental de cadenas en la siguiente generación sino que tal esquema recibe un número exponencial de cadenas en las siguientes generaciones.

Ejemplo

Supongamos que tenemos

- $pop_size = 20$
- Esquema S_0 donde $\xi(S_0, t) = 3$ (número de cadenas que describe).
- Tres cadenas c_1, c_2 y c_3 en el tiempo t donde la aptitud es 27.3, 30.06 y 23.86 respectivamente.

La aptitud del esquema es:

$$eval(S_0, t) = \frac{27,3 + 30,06 + 23,86}{3} = 27,08$$

La aptitud promedio de la población total es:

$$F(t) = \sum_{i=1}^{20} \frac{eval(c_i)}{pop_size} = \frac{387,7}{20}$$

Cadenas descritas por el esquema S

$$\frac{eval(S_0, t)}{F(t)} = \frac{27,8}{19,3} = 1,39$$

Si el esquema S_0 permanece sobre el promedio por un factor constante de 1,39, entonces, para

$t + 1$ esperamos tener $3 \times 1,39 = 4,19$ cadenas descritas por S_0

$t + 2$ esperamos tener $3 \times (1,39)^2 = 5,85$

lo que indica que el esquema S_0 es prometedor en el espacio de búsqueda puesto que se muestrea de forma incremental.

4. Efectos del cruce

El cruce es el segundo paso del ciclo de evolución e introduce nuevos individuos en la población. Supongamos que tenemos una cadena

$$c_{18} = 111011111010001000110000001000110$$

que es descrita por 2^{33} esquemas. En particular, la cadena está descrita por estos dos esquemas:

$$S_0 = ****111*****$$

$$S_1 = 111*****10$$

Suponiendo que las cadenas c_{18} y c_{13} fueron seleccionadas para cruce y que el punto es el 20, el esquema S_0 sobrevivirá al corte puesto que se preserva la secuencia 111 en c_{18} .

$$c_{18} = 1110**111**1101000100011|0000001000110$$

$$c_{13} = 00010100001001010100|1010111111011$$

$$c'_{18} = 1110\mathbf{111}1101000100011|1010111111011$$

$$c'_{13} = 00010100001001010100|0000001000110$$

Por otra parte, el esquema S_1 se destruirá pues ningún descendiente es descrito por él. Las posiciones fijas 111 y 10 están en distintos descendientes. La longitud del esquema es muy importante en la probabilidad de sobrevivencia y destrucción del esquema.

En general, un punto de cruce se selecciona uniformemente entre $m - 1$ posiciones. Esto implica que la probabilidad de destrucción de un esquema S es

$$p_d(S) = \frac{\delta(S)}{m - 1}$$

y la probabilidad de sobrevivencia del esquema es

$$p_s(S) = 1 - \frac{\delta(S)}{m - 1}$$

Calculando la longitud para S_0 y S_1 :

$$\delta(S_0) = 2$$

$$\delta(S_1) = 32$$

y la probabilidad de destrucción y sobrevivencia

$$p_d(S_0) = \frac{2}{32} = 0,0625$$

$$p_s(S_0) = 1 - \frac{2}{32} = 0,93$$

$$p_d(S_1) = \frac{32}{32} = 1$$

$$p_s(S_1) = 1 - 1 = 0$$

Sin embargo, solamente algunos cromosomas se someten al cruce con una probabilidad p_c . Esto significa que la probabilidad de sobrevivencia es en realidad:

$$p_s(S) = 1 - p_c \frac{\delta(S)}{m - 1}$$

suponiendo que $p_c = 0,25$, volvemos a calcular

$$p_s(S_0) = 1 - 0,25 \frac{2}{32} = 0,9843$$

Pero además, aunque el punto de cruce esté entre posiciones fijas, sigue existiendo la oportunidad de que el esquema sobreviva. Por ejemplo, si ambas cadenas c_{18} y c_{13} empezaran con 111 y terminaran con 10, el esquema S_1 podría

sobrevivir al cruce. Aunque es muy pequeña, la probabilidad existe. Debido a ello la probabilidad anterior se convierte en la probabilidad mínima que tiene un esquema de sobrevivir

$$p_s(S) \geq 1 - p_c \frac{\delta(S)}{m - 1}$$

El efecto combinado de selección y cruce forma una nueva ecuación que representa el número esperado de cadenas descritas por un esquema S en la siguiente generación en función del número actual de cadenas descritas por el esquema, aptitud relativa del esquema y su longitud definida.

$$\xi(S, t + 1) \geq \xi(S, t) \cdot \frac{eval(S, t)}{F(t)} \left[1 - p_c \cdot \frac{\delta(S)}{m - 1} \right]$$

Los esquemas cuyo valor de aptitud están sobre el promedio y que además tienen una pequeña longitud definida serán muestreados exponencialmente de manera incremental en generaciones sucesivas. Para el esquema S_0

$$\frac{eval(S, t)}{F(t)} \left[1 - p_c \cdot \frac{\delta(S)}{m - 1} \right] = 1,39 \cdot 0,98 = 1,37$$

en el tiempo $t + 1$ esperamos tener $3 \times 1,37 = 4,12$ cadenas descritas por S_0 (considerando solamente selección se obtuvo 4,19).

en el tiempo $t + 2$ esperamos tener $3 \times (1,37)^2 = 5,67$, ligeramente menor que con selección 5,85.

5. Efectos de la mutación

En la mutación el orden, las posiciones fijas son importantes ya que el cambio de al menos uno de esos bits importantes pueden destruir el esquema. Como la probabilidad de alterar un bit es p_m , la probabilidad de sobrevivencia de un bit es $1 - p_m$. Una mutación es independiente de otras mutaciones por lo que la probabilidad de que un esquema S sobreviva a la mutación (que las posiciones fijas del esquema permanezcan inalteradas) se obtiene al multiplicar la probabilidad de sobrevivencia por s misma $o(S)$ veces

$$p_s(S) = (1 - p_m)^{o(S)}$$

donde $o(S)$ es el orden del esquema, las posiciones fijas. Como la probabilidad de mutación es generalmente muy pequeña, cercana a 0, se puede aproximar la expresión de forma que:

$$p_s(S) \approx 1 - o(S) \cdot p_m$$

Aplicando al esquema S_0 y considerando $p_m = 0,01$

$$p_s(S_0) \approx 1 - 3 \cdot 0,01 = 0,97$$

Combinando el efecto de la selección, cruce y mutación, se tendría la siguiente ecuación:

$$\xi(S, t+1) \geq \xi(S, t) \cdot \frac{eval(S, t)}{F(t)} \left[1 - p_c \cdot \frac{\delta(S)}{m-1} - o(S) \cdot p_m \right]$$

que indica el número de cadenas esperadas descritas por el esquema S en la siguiente generación. Un esquema sobre el promedio, con pequeña longitud definida y orden bajo tiene un crecimiento exponencial. Para el esquema S_0 :

$$\xi(S, t+1) \geq \xi(S, t) \cdot \frac{eval(S, t)}{F(t)} \left[1 - p_c \cdot \frac{\delta(S)}{m-1} - o(S) \cdot p_m \right]$$

en el tiempo $t + 1$ esperamos tener $3 \times 1,33 = 4,0$ cadenas descritas por S_0 (considerando solamente selección se obtuvo 4,19)

en el tiempo $t + 2$ esperamos tener $3 \times (1,33)^2 = 5,33$, ligeramente menor que con selección 5,85.

En resumen, la selección incrementa exponencialmente, generación tras generación el número de soluciones representadas por un esquema si este tiene un valor sobre el promedio. Sin embargo, la selección no introduce nuevos esquemas por lo que se incorporan los operadores de cruce y mutación. El

cruce permite el intercambio de información entre soluciones y la mutación introduce variabilidad en la población. Estos dos procesos pueden originar la destrucción de los esquemas pero este efecto no es significativo para esquemas de longitud definida corta y orden pequeño. El resultado final puede enunciarse como:

Teorema de esquemas. Los esquemas cuyo valor de aptitud está por encima del promedio, de corta longitud definida y orden bajo están representadas por un número de soluciones que crece de forma exponencial a lo largo de las iteraciones del algoritmo genético.

Como resultado de este teorema, se enunció la siguiente hipótesis:

Hipótesis de los bloques constructivos. Un algoritmo genético tiene un desempeño de búsqueda casi-óptimo mediante la yuxtaposición de esquemas cortos, de bajo orden y alto desempeño a los que se les llama bloques constructores.

6. Desventajas del teorema

- No es muy útil para predecir a largo plazo el comportamiento de un algoritmo genético.

- Solamente considera los efectos destructivos de los operadores genéticos y no los efectos constructivos.
- Es muy particular. Se hizo para algoritmos genéticos con selección proporcional, cruce de un punto y probabilidad de mutación uniforme.

7. Ejercicios

Resolver los siguientes problemas:

1. Una población contiene las cadenas y valores de aptitud en la generación 0 que se muestran en la tabla.

#	Cadena	Aptitud
1	10001	20
2	11100	10
3	00011	5
4	01110	15

La probabilidad de mutación es $p_m = 0,01$. Calcula el número esperado de cadenas descritas por el esquema 1**** en la generación 1. Estima el número esperado de cadenas descritas por el esquema 0**1* en la generación 1.

2. Continuación. Para la función $f(x) = x^2$ y cadenas con una longitud de 5, calcular el número esperado de copias de cada cadena si se mantiene una población constante de $n = 4$.

Considerar los esquemas $S_1 = 1****$, $S_2 = *10**$, y $S_3 = 1***0$.

- Analiza los efectos de la selección sobre los tres esquemas para $t + 1$ y $t + 2$.
- Analiza los efectos del cruce sobre los tres esquemas $t + 1$ y $t + 2$.
- Analiza los efectos de la mutación sobre los tres esquemas $t + 1$ y $t + 2$.

El resultado de cada análisis es básicamente determinar si el esquema sobrevive o no. Incluir observaciones sobre la relación de los esquemas que incrementan o decaen con respecto a su longitud definida y orden.

8. Tarea: Martes 25 Sep

Resuelve el siguiente problema. Puede ser resuelto a mano y enviarlo escaneado.

Suponer que un esquema H cuando está presente en una cadena particular causa que la cadena tenga una aptitud 10 % menor que la aptitud promedio de la población. Si las probabilidades de destrucción para este esquema bajo mutación y cruce son desdeñables y si representantes de este esquema están contenidos en 60 % de la población en la generación 0, determina cuándo el esquema H va a desaparecer de poblaciones de tamaño $n = 20, 50, 100$ y 200 .

Referencias

- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning* (1st ed.). Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.
- Michalewicz, Z. (1996). *Genetic algorithms + data structures = evolution programs* (3rd ed.). London, UK, UK: Springer-Verlag.