

# Algoritmos Evolutivos inspirados en la teoría del gen egoísta.

Villagra A., De San Pedro M.; Lasso M; Pandolfi D.

Proyecto UNP-29/B032

División Tecnología

Unidad Académica Caleta Olivia

Universidad Nacional de Patagonia Austral

Ruta 3 Acceso Norte s/n

(9011) Caleta Olivia – Santa Cruz – Argentina

{avillagra, edespedro, mlasso, [dpandolfi](mailto:dpandolfi@uaco.unpa.edu.ar)}@uaco.unpa.edu.ar

Telefono/Fax: 0297-4854888

## Resumen

En este trabajo se exploran dos implementaciones computacionales diferentes basadas en la teoría biológica del Gen egoísta propuesta por Richard Dawkins en 1976. Este enfoque computacional fue propuesto inicialmente por Corno et al en problemas de optimización y llamado *Algoritmo del Gen Egoísta*. Las modificaciones discutidas aquí tienen como objetivo mejorar la velocidad de convergencia del algoritmo introduciendo modificaciones en la estructura de la población virtual de genes y la aplicación del cruzamiento. Además, se utiliza la población virtual de genes como mecanismo de inserción de conocimiento al algoritmo genético SRSI (Stud, Random and Seed Immigrants).

## 1. Teoría del Gen Egoísta (Selfish Gene)

En 1976, el etólogo Richard Dawkins publicó un libro [Daw76], *"El gen egoísta"*, en el que se divulgaban las tesis de la sociobiología sentadas anteriormente por E. O. Wilson en su *"Sociobiología"* de 1975 [Wil75]. El propósito de Dawkins es examinar la biología del altruismo y del egoísmo. Demuestra que el factor importante en la evolución no es el bien de la especie o grupo, como tradicionalmente se entiende, sino el bien del individuo o gen. Para él y sus seguidores, los individuos no son más que máquinas creadas por los genes para su supervivencia. Parafraseando a Butler, *"la gallina no es más que un invento del huevo para poder producir más huevos"*.

Existe, siempre según Dawkins, una interpretación errónea del altruismo: este se da, según las ideas tradicionales, por el bien de la especie, lo que se conoce como teoría de **selección de grupos**, es decir que la selección natural actúa sobre la especie. Un individuo no sería más que un *"peón"* que se sacrificaría por el bien de la especie.

La alternativa es la **selección de genes (o selección de individuo)**: los individuos altruistas llegan a extinguirse en beneficio de los egoístas, que predominarán en el grupo. Los genes han construido una gran variedad de "máquinas" para prosperar, explotándolas de modo que *"un gen puede ser considerado como una unidad que sobrevive a través de un gran número de cuerpos sucesivos e individuales"*. Así, un gen es definido como una porción de material cromosómico que, potencialmente, permanece durante suficientes generaciones como para servir como una unidad de

selección natural. El individuo es demasiado grande y efímero como para ser considerado unidad de selección. Un gen es considerado bueno, es decir, que permanece muchas generaciones, si vela por sí mismo, si es egoísta. La evolución será el proceso por el que algunos genes se hacen más numerosos y otros disminuyen en el acervo genético.

Todos los genes controlan el comportamiento de su máquina de supervivencia, no de manera directa, sino indirectamente. Los genes preparan la máquina con antelación, y luego esta se haya bajo su propia responsabilidad. Los genes obran a largo plazo mediante la síntesis proteica, pero se trata de un proceso lento. Por tanto, los genes construyen su máquina por anticipado, de la mejor forma posible y programándola con antelación.

Por tanto, el comportamiento está regido por el egoísmo de los genes de cada organismo, y no por el altruismo de cada individuo con respecto a los demás miembros de su especie. Dawkins se encarga de demostrar esto a lo largo de todo el libro con numerosos comportamientos particulares.

## 2. El Algoritmo Selfish Gene.

En los algoritmos genéticos tradicionales, una población es un conjunto de soluciones (individuos) codificadas en el que cada una tiene asociado una medida de adaptación al medio (problema) denominada *fitness*. Los individuos de una población son seleccionadas para producir nuevos individuos a través de un mecanismo de reproducción conocido como crossover. Así entonces, aquellos individuos que poseen un mejor fitness tienen una probabilidad mas alta de transmitir sus genes a las siguientes generaciones.

A diferencia del algoritmo genético tradicional el algoritmo Selfish Gen (SG) propuesto por Corno et al [CSRS98a, CSRS98b] para resolver “*Knapsap 0/1 Problem*” los individuos no son importantes. La población no es una enumeración de individuos sino una modelo abstracto denominado *Población Virtual (PV)* que representa un modelo estadístico del *pool* de genes. Los individuos no existen sino son representados por su cromosoma para evaluar su adaptación al problema. En un cromosoma la posición que ocupa un gen es denominada *loci*, mientras que el valor que asume en esa posición es *alelo*. La PV representa todos los vectores de probabilidades donde  $p_{ij}$  es la probabilidad de que un alelo  $a_{ij}$  sea seleccionado. A partir PV cualquier solución (cromosomas) puede ser construida sin embargo algunas cromosomas son mas probables de construir que otros. La figura 1 muestra la estructura general del algoritmo SG. En el algoritmo se observa que los individuos son construidos desde la PV a través de una función de selección. Esta función puede muy ocasionalmente generar aleatoriamente un individuo, proceso conocido como mutación. En la implementación elegida por Corno et al no existen mecanismos de reproducción (crossover). Los dos individuos generados participan de un torneo donde se mide la calidad de cada solución. La solución ganadora incrementa las probabilidades de que su solución codificada en el cromosoma sea elegida, mientras que el perdedor inversamente las disminuye. El proceso finaliza cuando los vectores de probabilidad de la población virtual (PV) tienden a estabilizarse.

```

Algoritmo SG ()
{
  genome B, G1, G2;
  initialize all probabilities  $p_{ij}$  to  $1/n_i$  ;
  B = select_individual() ; /* best so far */
  do {
    G1 = select_individual();
    G2 = select_individual();
    /* tournament */
    if (fitness G1 ) > (fitness G2 )
    {
      reward_alleles (G1 );
      penalize_alleles (G2 );
      if (fitness G1 ) > (fitness B ) B = G1 ;
    } else {
      reward_alleles (G2 );
      penalize_alleles (G1 ):
      if (fitness G2 ) > (fitness B ) B = G2 ;
    }
  } while (steady_state () == FALSE)
  return B; }

```

Figura 1: *Algoritmo Selfish Gene*

### 3. Algoritmos MCMP-SRI y MCMP-SRSI combinados Selfish Gene.

Multiple Crossover per Couple (MCMP) [ELG97] y Multiple Crossover Multiple Parent (MCMP) [ELG99] son métodos de multirecombinación que mejoran la performance de los algoritmos evolutivos tradicionales, reforzando el balance entre la exploración y la explotación del espacio de búsqueda.

Una extensión del método MCMP conocida como MCMP-SRI (Stud and Random Immigrants) [PVDVG01] promueve la múltiple recombinación de individuos evolucionados (Stud) con inmigrantes generados aleatoriamente. En una segunda variante conocida como MCMP-SRSI [PDVVG02] (Stud, Random and Seed Immigrants) individuos con conocimiento especializado denominados semillas (Seed) insertan conocimiento asociado al problema. En el proceso de creación de nuevos descendientes un individuo conocido como *stud* es seleccionado de la vieja población y es insertado en el *pool* de apareamiento junto con individuos creados aleatoriamente, *random immigrants*. Para el segundo algoritmo, las semillas con conocimiento del problema, *seeds*, son también insertadas completando así  $n_2$  padres. El *stud* se recombina con cada uno de los otros padres del pool de apareamiento generando  $(n_2 - 1) * 2$  descendientes. El proceso es repetido para  $n_1$  operaciones de crossover. Finalmente el mejor de todos los proto-descendientes es insertado en la nueva población.

Una primera variante, algoritmo MCMP-SRI-SG, aplica la teoría del gen egoísta sustituyendo la población de individuos por una población virtual. La PV contiene vectores de frecuencias acumuladas de los alelos que sirve para la generación de una distribución de probabilidades en la creación de los individuos conocidos como *stud*. A diferencia del algoritmo de Selfish Gene donde la PV se actualiza por premios y castigos por el ganador y perdedor del torneo, en este nuevo algoritmo solo se acumula un premio al ganador del proceso de apareamiento generado por Stud y Random Immigrants. Esta modificación además de evitar el control de

probabilidades negativas, posibilita generar premios diferenciales, si por ejemplo el mejor descendiente se transforma en el mejor de la evolución. Para el caso de los inmigrantes aleatorios, todos los alelos son igualmente probables de ser seleccionados en cualquier posición del cromosoma. El algoritmo MCMP-SRI-SG, a diferencia del algoritmo SG, aplica crossover entre los individuos del pool de apareamiento con el objetivo de explotar sub-esquemas mas promisorios del cromosoma.

En la segunda variante, conocida como MCMP-SRSI-SG la población de stud es representada como en el algoritmo original MCMP-SRSI, mientras que el conocimiento provisto por las semillas es generado por una población virtual, tal como en el caso anterior. Tradicionalmente en el algoritmo MCMP-SRSI, las semillas eran generadas a través de buenas heurísticas asociadas al problema, mientras que en este caso el conocimiento no es del problema sino de la evolución misma.

#### **4. Discusión.**

El Algoritmo Selfish Gene (SG) desarrollado por Corno et al implementa la teoría Dawkins donde los individuos son máquinas construidas por los genes para producir mas genes; los individuos en este tipo de implementación solo sirven para evaluar las soluciones y dirigir la evolución de la población virtual. Los algoritmos MCMP-SRI y MCMP-SRSI por su parte basan su estructura en poblaciones de individuos evolucionados que se combinan con individuos generados aleatoriamente. Esta combinación permite explorar y explotar en forma balanceada el proceso de búsqueda.

Los algoritmos propuestos MCMP-SRI-SG y MCMP-SRSI-SG incorporan las estructuras de poblaciones virtuales en los algoritmos tipo SRI (Stud and Random Immigrants) tanto para la población principal como en la forma de inserción de conocimiento (generación de semillas).

Nuevas propuestas para implementación de las poblaciones virtuales se introducen con el objetivo de mejorar construcción de nuevas soluciones (poblaciones basadas en premios acumulados) como así la velocidad de convergencia (premios mayores) del algoritmo.

Los algoritmos discutidos en este trabajo se encuentran en etapa de prueba con distintos tipos de problemas de Scheduling.

#### **5. Agradecimientos**

Nuestro agradecimiento a la cooperación y criticas constructivas provistas por el grupo de proyecto. A la Universidad Nacional de la Patagonia Austral por el soporte al proyecto.

Principalmente y en su memoria al Dr. Raul Gallard que supo inspirar en este grupo deseos de superación y por quien guardaremos el máximo cariño, respeto y agradecimiento.

#### **6. Referencias Bibliograficas.**

- CSRS98a** Corno F., Sonza Reorda M., Squillero G.; The Selfish Gene Algorithm: a new Evolutionary Optimization Strategy; 13<sup>th</sup> Annual ACM Symposium Applied Computing, Atlanta, Georgia (USA) February 1998, pp. 349-255.
- CSRS98b** Corno F., Sonza Reorda M., Squillero G.; A new Evolutionary Algorithm Inspired by the Selfish Gene Theory ICEC98: IEEE International Conference on Evolutionary Computation, May 1998 pp 575-580.

- Daw76** Dawkins Richard; The selfish gene, Oxford University Press, 1976.
- ELG97** Esquivel S., Leiva A., Gallard R; Multiple Crossover per Couple in Genetic Algorithms, Proceedings of 4<sup>th</sup> IEEE Conference Evolutionary Computation, ICEC97 Indianapolis. USA, April 1997, pp. 103-106.
- ELG99** Esquivel S., Leiva A., Gallard R; Multiple Crossover between Multiple Parents to improve search in Evolutionary Algorithms, Proceedings of Congress on Evolutionary Computation, CEC Washington DC, USA, 1999, pp. 1589-1594.
- PDVVG01** Pandolfi D.; De San Pedro M.E.; Vilanova G., Villagra A.; Gallard R. Multirecombining random and seeded immigrants in evolutionary algorithms to solve W-T scheduling problems, Proceedings ACIS International Conference on Computer Science, Software Engineering, Information Technology eBusiness and Application Foz Iguazu Brazil 2002.
- PVDVG01** Pandolfi D.; Vilanova G., De San Pedro M.E.; Villagra A.; Gallard R. Multirecombining studs and immigrants in evolutionary algorithm to face earliness-tardiness scheduling problems, Proceedings of the International Conference in Soft Computing University of Paisley, Scotland UK June 2001 pp. 138.
- Wil75** Wilson E. O. ; Sociobiology:the new synthesis, Harvard University Press, 1975.