

Centro de Investigación en Cómputo
Instituto Politécnico Nacional
Metaheurísticas

Actividad 19: Solución de problemas mediante heurísticas de trayectoria
simple y evolutivas

Curso impartido por: Dra Yenny Villuendas Rey

Adrian González Pardo

11 de enero de 2021

1. Heurísticas

1.1. Random Mutation Hill Climbing Algorithm (RMHC)

Es un algoritmo heurístico de trayectoria simple, es decir, que la solución a buscar o a obtener es dada por un estado inicial y a partir de el se realiza la exploración del área en búsqueda de un óptimo global o local, por ello:

Ventajas	Desventajas
Permite realizar un número menor de iteraciones con respecto a la exploración del espacio en modalidad de fuerza bruta	El resultado puede quedarse en un mínimo local sin que sea el mínimo del espacio
Permite moverse en el espacio de diferentes formas, permitiendo que encuentre un mínimo	Dependiendo de la implementación del algoritmo puede que que llegue al mínimo o no, pero al final arroja una solución
Dependiendo de la implementación puede que tenga un número distinto de iteraciones	En caso de ser un espacio muy grande puede que la implementación, salto o movimiento entre el espacio demore demasiado

Una vez explicado las ventajas y desventajas se trata de un algoritmo que actúa a partir de una solución aleatoria la cual al obtener un valor aleatorio m este puede representar una cadena de bits sobre la cual puede interpretarse como una mutación real de un arreglo vía índice o como una mutación de bits la cual intercambiara los valores de diversas áreas de un arreglo o que permutara, todo depende de la implementación, a partir de el se buscara o se desea que la modificación de la respuesta inicial sea más óptima o aproximada a la solución determinista del programa.

1.2. Simulated Annealing Algorithm (SA)

Al igual que RMHC esta es una heurística de trayectoria simple que a diferencia del mismo RMHC, esta permite aceptar o escapar de un óptimo local, con la idea de llegar al óptimo global, por ello:

Ventajas	Desventajas
Permite obtener una solución aproximada al mínimo global	Puede llegar a pasar que en la siguiente iteración del programa, esta salga del mínimo global y este se alcance un resultado mínimo local o en otra área
Permite encontrar soluciones razonables (En tiempo)	La solución encontrada puede no ser la mejor
Utiliza poca memoria	No permite vuelta atrás en más de 1 paso

Este tipo de heurística busca simular el enfriado de metales por ello en el algoritmo se implementa una temperatura máxima a la cual el algoritmo comenzara a trabajar, más tarde después de ello se enfriara mediante un factor α el cual puede ser seleccionado aleatoriamente en el intervalo de $[0, 1]$ de modo que en el trabajo presente el intervalo fue modificado para que α tome un valor aleatorio desde $[0,88, 0,99]$ de modo en que este llegue o converja en una solución óptima.

1.3. Genetic Algorithm (GA)

A diferencia de las dos heurísticas previas que funcionan como trayectoria simple, esta heurística es un algoritmo o un modelo poblacional en el cual se busca generar múltiples soluciones las cuales competirán por ser seleccionadas mediante su operador distintivo el cual puede ser una selección por su función, en el que se buscare dos óptimos, para más tarde pasar a través de una función de probabilidad la cual nos indicara si las dos soluciones pueden ser mutadas (modificadas), y después más tarde pueden ser modificadas para formar parte del espacio de solución o de la generación espacial de soluciones, por ello:

Ventajas	Desventajas
Permite realizar múltiples búsqueda de soluciones a los problemas	Puede que el método de mutación o cruce seleccionado puede que no ayude a encontrar una buena solución
Esta bioinspirado en la genética y en la selección natural (Darwinismo)	Puede que el que en la aplicación en alguna etapa del GA ya no avance
Es una heurística poblacional	Puede que genere bastante uso de recursos en memoria y procesamiento
Es posible el trabajar soluciones de formas paralelizables o distribuidas	Puede ser difícil de implementar

2. Aplicaciones

A partir de cualquier heurística se puede realizar o buscar la solución a problemas de la clase *NP-Completo* los cuales pueden ser mapeados entre ellos con diversas restricciones, para solucionar y encontrar una función o espacio óptimo de la solución, por ello los principales problemas implementados fueron y han sido Knapsack Problem, Travel Salesman Problem y Minimization Functions (diversas funciones de múltiples variables)

3. Funciones a optimizar

Función a evaluar	Forma o descripción de la función
Alpine Function	$f_1(x) = \sum_{i=1}^D x_i \sin(x_i) + 0,1x_i $
Dixon & Price Function	$f_2(x) = (x_1 - 1)^2 \sum_{i=2}^D i (2 \sin(x_i) - x_{i-1})^2$
Quintic Function	$f_3(x) = \sum_{i=1}^D x_i^5 - 3x_i^4 + 4x_i^3 - 2x_i^2 - 10x_i - 4 $
Schwefel 2.23 Function	$f_4(x) = \sum_{i=1}^D x_i^{10}$
Stretched V Sine Wave Function	$f_5(x) = \sum_{i=1}^{D-1} (x_{i+1}^2 + x_i^2)^{0,25} [\sin^2 \{50(x_{i+1}^2 + x_i^2)^{0,1}\} + 1]$
Sum Squares Function	$f_6(x) = \sum_{i=1}^D ix_i^2$

4. Características de Hardware

- Procesador Intel Core i5-3210M CPU 2.50GHz
- Memoria RAM DDR3 12 GB
- Sistema Operativo Fedora 32 x86_64
- Ruby 2.7.2p137 para la ejecución

5. Características de ejecución:

- Para las tres heurísticas:
 - Dimensiones de la función 10 y 30
 - Máximo número de iteraciones 500
 - En la ejecución de operadores normales $[-10, 10] \in \mathbb{R}$
 - En la ejecución operadores convenientes:
 - Quintic con intervalos $[2, 47, 2, 479] \in \mathbb{R}$
 - Resto de funciones con intervalos $\left[\frac{-10}{\text{random}(8 \dots 40)}, \frac{10}{\text{random}(8 \dots 30)} \right] \in \mathbb{R}$
- Para SA:
 - Temperatura máxima de 1000
 - α aleatorio para cada ejecución del intervalo $[0, 88, 0, 99]$
 - Temperatura mínima de 0,01
- Para GA ambas implementaciones:
 - Población de 10
 - Probabilidad de entrar a la mutación 0,5
 - Probabilidad de mutar 0,15
 - Operador de selección por torneo
 - Operador de mutación real
 - Operador de cruzamiento aritmético
 - Operador de remplazo por elitismo

6. Tabla de resultados

6.1. Tabla con operadores normales

Normal	Mejor Ejecución				Desviación Ejecución				
Funciones	RMHC	SA	GA_St	GA_Gen	RMHC	SA	GA_St	GA_Gen	Dimensión
Alpine	12.2	17.58	0.63	3.07	4.553415641	9.685484758	0.773661909	1.237050019	10
Dixon	13.19	15.86	14.04	65.03	7.158801069	6.956498023	23.51071221	36.80797251	10
Quintic	14.96	7.24	10	147.61	8.636641405	3.802724648	22.64709564	797.6320317	10
Schewel	19.1	3.4	0.03	51748.93	6.235170788	1.904257007	751.0266226	1397041.848	10
Streched	12.68	13.38	1.41	2.78	7.954882447	7.669308704	0.470725769	0.348141853	10
Sum Squares	16.77	18.35	6.92	93.64	6.078585444	6.019301683	12.11748376	40.18008884	10
Promedio	14.81666667	12.635	5.505	8676.843333	6.769582799	6.006262471	135.0910503	232986.3422	

Normal	Promedio Tiempo				Desviación Tiempo				
Funciones	RMHC	SA	GA_St	GA_Gen	RMHC	SA	GA_St	GA_Gen	Dimensión
Alpine	3.7765	16.38825	487.17325	2988.98675	0.957171745	78.28615755	395.4569062	1.199873327	10
Dixon	23.41175	10.66575	301.839	2766.148	78.11122022	46.91205388	86.2764475	88.26461975	10
Quintic	56.00725	118.102	334.4015	3466.73825	121.0929091	170.7582796	115.9632208	107.1785027	10
Schewel	12.601	37.9415	290.97775	2667.85625	62.59825005	89.66118401	111.9327308	87.30419199	10
Streched	71.27475	22.87975	338.49975	3629.268	136.2918611	50.69156436	132.0850486	92.19589822	10
Sum Squares	33.9665	6.5695	238.003	2847.2725	86.96842167	23.56447294	109.3165055	65.38166781	10
Promedio	33.50629167	35.42445833	331.8157083	3061.044958	81.00330565	76.64561872	158.5051432	73.58745897	

Normal	Mejor Ejecución				Desviación Ejecución				
Funciones	RMHC	SA	GA_St	GA_Gen	RMHC	SA	GA_St	GA_Gen	Dimensión
Alpine	12.04	14.81	0.87	5.64	6.457145654	6.767959718	0.533667499	0.600383211	30
Dixon	10.72	12.7	16.33	79.92	7.860760698	6.299853391	23.98680639	32.60268094	30
Quintic	16.16	9.52	28.18	193.6	5.391284518	2.568580299	27.81968427	569.6833757	30
Schewel	17.67	3.18	0.5	8915.66	6.478638649	2.116785948	4899.820316	696860.918	30
Streched	17.51	19.88	1.35	2.04	5.41133521	7.231813379	0.431276014	0.456393197	30
Sum Squares	12.64	13.13	4.65	139.88	7.873380405	5.816041953	16.23591802	34.81761429	30
Promedio	14.45666667	12.20333333	8.646666667	1556.123333	6.578757522	5.133505781	828.1379447	116249.8464	

Normal	Promedio Tiempo				Desviación Tiempo				
Funciones	RMHC	SA	GA_St	GA_Gen	RMHC	SA	GA_St	GA_Gen	Dimensión
Alpine	3.782	2.214	337.718	2988.5805	1.199873327	0.219440197	122.7782947	532.8448014	30
Dixon	26.879	3.0055	304.5935	2662.6175	88.26461975	0.153442986	90.75646932	717.2910463	30
Quintic	45.268	117.625	329.2735	3808.093	107.1785027	167.4276441	111.4870536	672.7206154	30
Schewel	22.9675	29.4995	260.3735	2738.1925	87.30419199	68.3627415	116.8109368	556.4711693	30
Stretched	43.381	18.0905	330.7815	3646.6475	92.19589822	30.15070621	142.9809351	693.2458665	30
Sum Squares	21.655	10.703	178.7115	2849.4175	65.38166781	32.8083107	106.0881994	745.1110182	30
Promedio	27.32208333	30.18958333	290.2419167	3115.591417	73.58745897	49.85371428	115.1503148	652.9474195	

6.2. Tabla con operadores convenientes

Conveniente	Mejor Ejecución				Desviación Ejecución				
Funciones	RMHC	SA	GA_St	GA_Gen	RMHC	SA	GA_St	GA_Gen	Dimensión
Alpine	0.31	0.01	1.7	0.23	2.504386352	3.495248882	0.030045798	1.395960153	10
Dixon	0.35	0.5	17.74	0.16	3.148242208	2.8958693	2.858812516	18.75899925	10
Quintic	17.77	0.11	24.9	17.76	0.008046738	0.00678233	0.035972211	75.43451964	10
Schewel	0.31	0	62.45	0.37	2.93730876	2.277237581	0.008717798	41149.16066	10
Stretched	0.33	0.29	1.55	0.23	2.880422669	3.397818418	0.265376619	0.407004607	10
Sum Squares	0.36	0.01	24.75	0.17	2.768783849	3.688129065	0.522363858	23.31060451	10
Promedio	3.238333333	0.153333333	22.18166667	3.153333333	2.374531763	2.626847596	0.6202148	6878.077958	

Conveniente	Promedio Tiempo				Desviación Tiempo				
Funciones	RMHC	SA	GA_St	GA_Gen	RMHC	SA	GA_St	GA_Gen	Dimensión
Alpine	3.893	23.574	598.8635	2704.606	1.383485092	78.56329164	459.6706309	648.7076449	10
Dixon	2.998	30.221	326.175	2986.513	0.997815614	109.3947241	111.3212846	594.0357069	10
Quintic	43.708	33.4685	337.237	3399.759	116.9267684	108.421163	75.4822868	969.8358564	10
Schewel	25.5165	22.23	276.873	2667.1315	101.7921228	87.3161674	116.0201503	521.9268236	10
Stretched	24.0725	43.299	337.236	3575.967	69.58001393	111.5766269	111.9194054	680.6257294	10
Sum Squares	37.938	2.445	275.8875	2872.48	98.32770788	0.392103303	72.64485686	606.3809896	10
Promedio	23.021	25.87291667	358.712	3034.409417	64.8346523	82.6106794	157.8431025	670.2521251	

Conveniente	Mejor Ejecución				Desviación Ejecución				
Funciones	RMHC	SA	GA_St	GA_Gen	RMHC	SA	GA_St	GA_Gen	Dimensión
Alpine	0.23	0.4	0.01	1.3	4.324667617	3.666809615	0.070091012	1.380662432	30
Dixon	0.16	0.25	0.45	22.22	1.974102581	2.573344128	2.365132713	22.40861495	30
Quintic	17.76	17.76	0.06	11.2	0.009733961	0.010049876	0.043886217	67.80873036	30
Schewel	0.37	0.27	0	1.81	3.097699953	2.557508553	0	27260.64969	30
Streched	0.23	0.37	0.35	1.8	1.986876632	1.854187895	0.231954198	0.381268934	30
Sum Squares	0.17	0.31	0.01	4.78	2.689877832	2.04993756	0.045637156	30.73159496	30
Promedio	3.153333333	3.226666667	0.146666667	7.185	2.347159763	2.118639604	0.459450216	4563.893427	

Conveniente	Promedio Tiempo				Desviación Tiempo				
Funciones	RMHC	SA	GA_St	GA_Gen	RMHC	SA	GA_St	GA_Gen	Dimensión
Alpine	41.1805	18.2975	324.648	2975.907	115.0214601	65.75507021	101.0822794	531.26273	30
Dixon	326.175	2986.513	24.0285	38.986	87.45066285	94.63920427	100.1521022	758.0425275	30
Quintic	337.237	3399.759	27.2755	69.2915	87.53136303	139.5798924	96.92476907	708.6012876	30
Schewel	276.873	2667.1315	22.031	2.183	87.51205225	0.958212398	93.57974904	507.2426142	30
Streched	337.236	3575.967	54.5195	24.6895	137.2765873	87.71346587	100.8476997	700.4716143	30
Sum Squares	275.8875	2872.48	28.794	31.7035	109.2339488	89.06785948	90.7710913	538.2264179	30
Promedio	358.712	3034.409417	32.9715	30.8585	104.0043457	79.61895077	97.22628178	623.9745319	

7. Conclusiones

Si bien la implementación y comparación de las dos distintas ejecuciones del programa, en este caso relacionadas con realizar las heurísticas mediante intervalos no conocidos de la función, es decir, que van de $[-10, 10]$ contra los intervalos más cercanos a la solución del óptimo, nos podemos dar cuenta que no solo influye mucho el hardware sino también las capacidades de la máquina ya que para la solución en tiempo óptimo de cada heurística fue necesario implementar un pequeño paralelismo el cual permitiera reducir el costo temporal de ejecución de todas y cada una de las heurísticas, para trabajos futuros podría pensarse en una hibridación de las heurísticas para que estas alcancen en su totalidad el óptimo global sin necesidad de utilizar algunos atajos, para llegar a un óptimo local.