

Programación evolutiva

Curso 2022/2023

Práctica 3: Programación genética/Gramáticas evolutivas

La práctica consiste en resolver un problema de regresión simbólica utilizando programación genética y gramáticas evolutivas. Se trata de identificar o descubrir mediante regresión simbólica una función a partir de un conjunto de datos reales. Este problema se puede aplicar en muchos contextos diferentes en el que partiendo de un gran conjunto de datos podemos obtener una función calculada que se ajusta a una función real.

Queremos encontrar una expresión de una variable x cuya salida sea igual a la función

$$f(x) = x^4 + x^3 + x^2 + x + 1$$

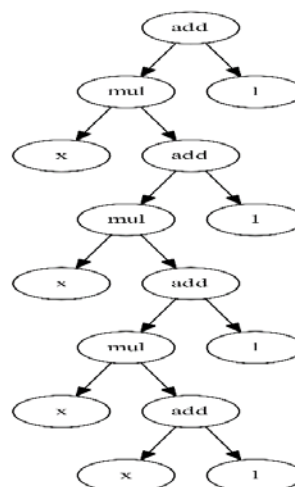
Para generar y obtener las expresiones mediante programación genética utilizaremos:

Funciones : add, sub, mul
Terminales : x, -2, -1, 0, 1, 2

Ejemplos de expresiones obtenidas:

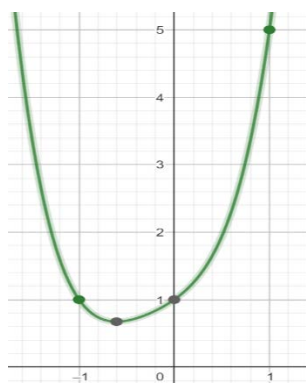
```
add
  mul
    mul
      x
      2
    add
      x
      1
  mul
    -1
    -1
```

```
add
  mul
    sub
      x
      0
    mul
      add
        1
        add
          1
          x
      x
    add
      1
      x
```



Para evaluar el fitness de nuestras soluciones utilizaremos un dataset con 100 valores de la variable $x \in [-1.0, 1.0]$ y el correspondiente valor de la función en dicho punto $f(x)$.

x	f(x)
-1.0	1.0
-0.98	0.9615
-0.96	0.9262
-0.94	0.8937
-0.92	0.8641
.	.
0.92	4.2614
0.94	4.4349
0.96	4.6156
0.98	4.8039
1.0	5.0

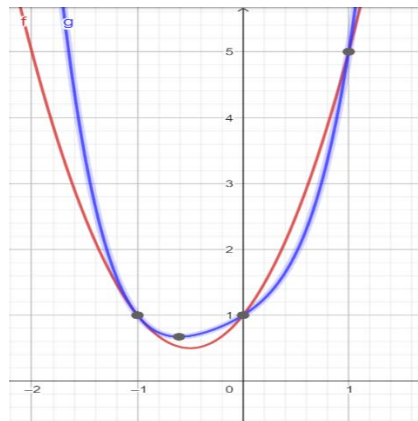


El objetivo es descubrir la expresión que minimice alguna medida de error (como la raíz cuadrada de la suma de los cuadrados de las diferencias) entre los valores dados $f(x)$ y los valores calculados de la función $f(x)$ **calculado**.

Estos 100 pares ($f(x)$ **real**, $f(x)$ **calculado**) son los casos que vamos a utilizar para evaluar el fitness de la expresión obtenida.

Procedimiento:

- Se genera una población inicial de programas aleatorios (árboles) usando el conjunto de funciones y terminales posibles. Estos árboles deben ser sintácticamente correctos. Limitamos la profundidad (mínima 2 máxima 5). Métodos Creciente, Completa y Ramped and half.
- Fitness: Raíz cuadrada de la suma de los cuadrados de las diferencias entre el valor real y el calculado
- Operador de Cruce: Intercambiar dos subárboles (elegidos aleatoriamente) entre los dos árboles padres.
- Operadores de Mutación: al menos 3 de los vistos en el material de la asignatura.
- Algún método de controlar el bloating para evitar programas muy largos
- Se mostrarán las gráficas de evolución y la expresión final obtenida, junto con su fitness.
- Se mostrará alguna gráfica que compare el ajuste visualizando la función original y la función calculada.



Parte opcional: Gramática evolutiva

- Resolver el mismo problema utilizando gramáticas evolutivas
- Se utilizarán los operadores tradicionales sobre cromosomas de enteros.
- Se podrá seleccionar el número de wraps
- Se ha de definir una gramática adaptada al problema

Entrega

- **Plazo de entrega: 3 de mayo a las 12:00.** Debes entregar mediante la tarea de entrega del Campus Virtual un archivo comprimido con el código java de la aplicación (**proyecto en Eclipse o NetBeans**) que incluya una breve memoria que contenga el estudio de las gráficas y los resultados obtenidos con cada función. Aquí se valorarán las conclusiones y observaciones que se consideren interesantes respecto al resultado obtenido.
- No olvidéis nombrar correctamente el proyecto e incluir en el código todas las librerías necesarias. El archivo comprimido y el nombre del proyecto Eclipse tienen que ser **GXXP3**, donde XX es el número de grupo.
- Debes incluir una memoria similar a las entregas anteriores e incluir al final de la memoria una breve descripción del **reparto de tareas** para reflejar lo que ha hecho cada miembro del grupo.
- La corrección se realizará en las sesiones del **miércoles 3 y viernes 5 de mayo** y también se podrán hacer sesiones de corrección en horas de tutoría acordadas.
- Es importante conocer bien la práctica y los aspectos teóricos en los que se basa, pues es lo que determina la calificación final.
- Se realizará control de copias de todas las prácticas, comparando las entregas de todos los grupos de PE. Si se detecta algún tipo de copia sin justificar se calificará como suspenso.

Importante:

- Si se implementa la parte opcional se considerará parte de la actividad adicional de la asignatura (la actividad adicional de la asignatura es hasta el 25%).
- Si no se implementa la parte opcional la calificación de las prácticas supondrá un 75% de la nota total de la asignatura y para obtener el 25% de nota por actividad adicional será necesario hacer un trabajo-memoria extra sobre algún tema de los propuestos en el Tema 8 de la asignatura o sobre cualquier artículo científico relacionado con la materia de la asignatura.