

Лабораторная работа №4

Антипов Денис, гр. 5539 (вариант 17)

30 мая 2015 г.

1 Описание задачи

Требуется реализовать алгоритм, находящий функцию наиболее близкую к

$$f(x) = \sum_{i=1}^9 i * x_i^2$$

с помощью различных комбинаций функций $+$, $-$, $*$, $/$, \cos , \sin , $|x|$, pow , exp на гиперкубе $x_i \in [-5.12; 5.12]$

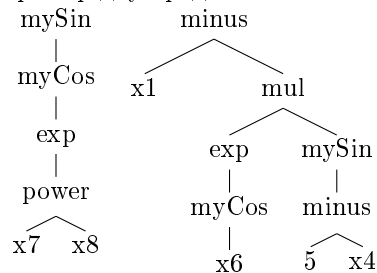
2 Описание алгоритма

Алгоритм использует древовидное представление деревьев. В качестве фитнес-функции используется среднееквадратичное отклонение от эталона по случайной выборке значений.

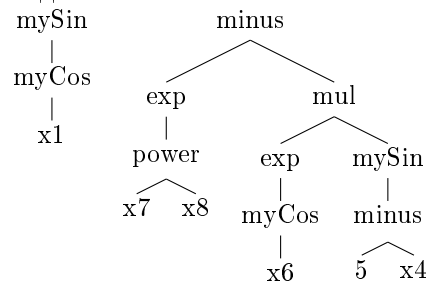
Оператор редукции выбирает особи популяции, чьи среднееквадратичное отклонение не является NaN или ∞ с вероятностями, обратно пропорциональными их фитнес-функциям.

Оператор кроссинговера выбирает 2 случайных вершины у родителей и меняет их местами.

Пример двух родителей:

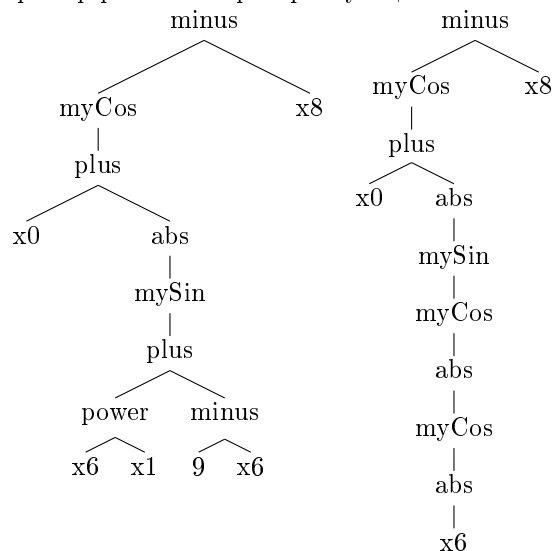


Их дети:



Оператор мутации выбирает случайную вершину в дереве (кроме корня), удаляет ее и вырастивает вместо нее новое дерево небольшой глубины.

Пример работы оператора мутации:



3 Результаты работы алгоритма

Результаты в моем случае неудовлетворительны, так как у меня нет достаточной вычислительной мощности для того, чтобы приблизить функцию на всем участке исследования. Вместо этого при подсчете фитнес-функции я каждый раз делаю небольшую выборку точек из рассматриваемого множества и смотрю среднеквадратичную разность на них. Попытки создать сетку, более-менее покрывающую весь гиперкуб к успеху не привела, так как даже если в каждом измерении было всего по 3 точки, то всего понадобилось бы 3^9 итераций, чтобы посчитать среднее квадратичное отклонение. На моем ноутбуке подсчет фитнес-функции длился несколько минут.

Из-за маленького размера сетки алгоритм всегда в течение 15 итераций сходил к функции, являющейся линейной от одной переменной. Расширение популяции ничего не изменило. Даже, скорее, ускорило сходимость. Увеличение размера выборки в разумных пределах для определения отклонения также не дало результата.

Должно было помочь нижнее ограничение на длину сгенерированного дерева. Но этот инвариант очень сложно поддерживать при кроссинговере. Но даже если использовать это ограничение на нижнюю длину дерева, а также понизить вероятность выбора "глубоких" вершин для кроссинговера, все равно на 10-ой итерации лучшим представителем является константа.