Эссе по курсу «Защита информации», кафедра радиотехники, Московский Физико-Технический Институт (МФТИ ГУ)

http://www.re.mipt.ru/infsec

Метод решения NP-полных задач биохимическими средствами (вычисления с помощью молекулы ДНК)

Белявский Павел

311 группа

16.05.2007г.

1 Введение

В курсе «Дискретной математики», «Проектирования Систем», да и в реальной жизни мы часто сталкиваемся с задачами, для которых не найдено алгоритмов нахождения точного решения за приемлемое время. Примерами являются задача «Коммивояжера» о поиске оптимального маршрута, задача об упаковке в контейнеры, латинский квадрат, разложение числа на множители. Все они относятся к классу NP-полных. В теории алгоритмов NP-полными называют класс задач, для которых пока не найдено быстрых алгоритмов решения, но проверка того, является ли данное решение правильным, проходит быстро. В 1971 году Куком была сформулирована теорема о выполнимости, которая утверждала, что всякая задача из класса NP может быть сведена к любой другой задаче из этого класса за полиномиальное время. Это означает, что если найдут быстрый алгоритм для решения любой из NP-полных задач, то и любая задача из класса NP сможет быть решена быстро.

Сложность нахождения точного решения для NP-полных задач растет экспоненциально (суперэкспоненциально), а не полиномиально. Надёжность асимметричной криптографии основана на NP-полных задачах. В настоящее время в основном используются две: проблема дискретного логарифмирования в конечном поле и проблема разложения на множители больших чисел. Если какая-либо из них станет решаема за полиномиальное время, соответствующие криптосистемы потеряют свою нынешнюю стойкость. Например шифр RSA широко используемый на данный момент (примерно в 90% случаях) основан на сложности разложения простых чисел на множители.

Для приближенного решения NP-полных задач существует множество алгоритмов, таких как «жадный» поиск, генетический алгоритм, метод ветвей и границ. Мы подробнее остановимся на генетическим алгоритме, который был заимствован у природы и немного дополнен.

2 Генетический алгоритм

Генетический алгоритм (ГА) основан на идее эволюции Дарвина с помощью естественного отбора. Впервые был предложен Холландом в 1975 году.

ГА работают с набором «особей» (ДНК) – популяцией, каждая из которых является возможным решением поставленной задачи. Каждой особи сопоставлена «приспособленность»(полезность), что в реальной жизни соответствует способности выживать. Наиболее приспособленные особи получают возможность «скрещиваться» и производить потомство, которое будет иметь черты, заимствованные от родителей. Менее приспособленные особи с меньшей вероятностью могут воспроизводить потомство, а занчит и те свойства, которыми они обладают, будут постепенно исчезать из популяции в процессе эволюции. Иногда в популяции могут происходить мутации – случайные изменения в генах.

ГА состоит из следующих элементов:

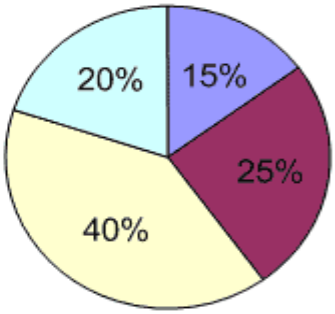
- **Хромосома**. Решение рассматриваемой задачи. Состоит из генов.
- **Начальная популяция** хромосом.
- **Набор операций** для получения новой популяции из предыдущей. К ним относятся: селекци, скрещивание, мутация и редукция.
- **Функция «Приспособленности»,** которая служит для оценки приспособленности отдельной особи. Она характеризует близость к искомому решению.

Селекция

Этот оператор отвечает за отбор хромосом по их функции приспособленности. Наиболее популярные методы селекции – это рулетки и турнир.

- **Метод рулетки** (roulette-wheel selection) – метод, где вероятность выбора прямо пропорциональна приспособленности особи. Метод заключается в запуске рулетки для каждой особи.

1 000110110 15%
2 111001101 25%
3 000110110 40%
4 111101111 20%



Оператор селекции типа колеса рулетки с
пропорциональными функции приспособленности секторами

Размер i-го сектора прямо пропорционален $p_{sel}(i)$, которая вычисляется по формуле

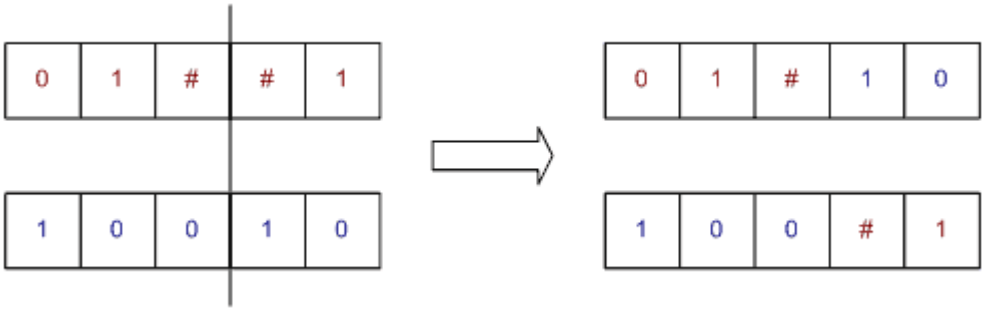
$$p_{sel}(i) = \frac{f(i)}{\sum_{i=1}^{n} f(i)}$$

При таком отборе более приспособленные особи будут чаще выбираться.

- **Турнирный отбор** (tournament selection) состоит из n турниров, для выбора n особей. Каждый турнир основан на выборе k особей из популяции и нахождения «сильнейшей» среди них. Наиболее популярен турнирный отбор с k=2.

Скрещивание

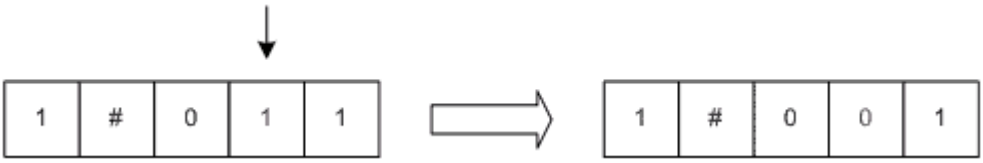
Оператор скрещивания (crossover) осуществляет обмен частями хромосом (генами) между двумя (или больше) особями популяции. Часто используется одноточечный кроссовер. Он работает следующим образом. Случайно выбирается точка разрыва хромосом, потом получившиеся кусочки хромосом разных родителей склеиваются, что приводит к появлению потомства.



Одноточечный оператор скрещивания (точка разрыва равна трем)

Мутация

Мутации (mutation) – это случайное изменение части хромосомы.



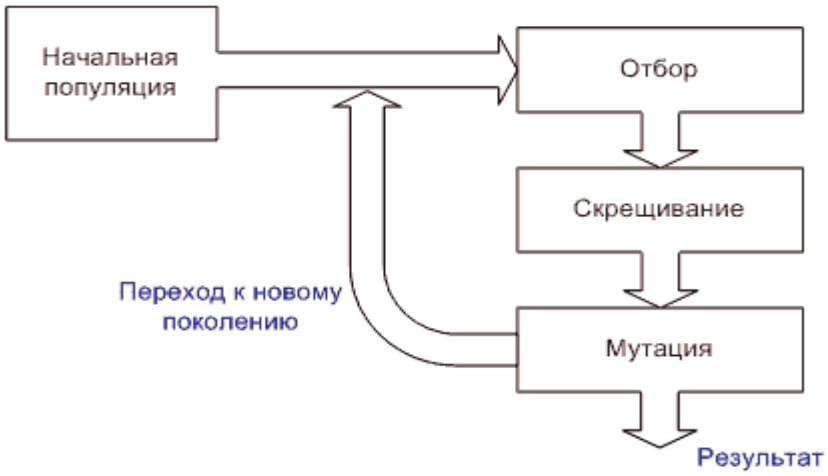
Оператор мутации (четвертый ген мутировал)

Редукция

Оператор редукции (reduction) осуществляет перенос наиболее преспособленной особи (особей) в новое поколение. Является нововведениеим по сравнению с естественным отбором, придуманным природой. Предотвращает возможное ухудшение особей в популяции, связанное с эволюцией.

Алгоритм работы ГА

Работа ГА представляет собой цикл, которые продолжается до тех пор, пока не будет выполнено заданное число поколений, не пройдет заданное время или не выполниться какой-нибудь другой критерий останова.



Алгоритм работы классического ГА

3 Заключение

Таким образом, хорошие характеристики распространяются от покаления к покалению. Скрещевание наиболее приспособленных особей приводит к тому, что исследуются наиболее перспективные пространства решений. Преимущества ГА заключается в том, что он находит приблизительные оптимальные решения за относительно короткое время, устойчив к попаданию в локальные максимумы, может быть использован для широкого класса задач, прост в реализации. Среди недостатков стоит отметить то, что ГА плохо подходит в случаях, если необходимо найти точное решение или все решения, а не одно.

4 Литература

1. http://ru.wikipedia.org/wiki/NP-%D0%BF%D0%BE%D0%BB%D0%BD%D0%B0%D1%8F_%D0%B7%D0%B0%D0%B4%D0%B0%D1%87%D0%B0
2. http://ru.wikipedia.org/wiki/%D0%93%D0%B5%D0%BD%D0%B5%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9_%D0%B0%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC
3. http://board.buddhist.ru/showthread.php?t=3762&page=2
4. http://www.codingtheory.gorodok.net/literature/salomaa.pdf
5. http://www.google.com
6. http://www.codenet.ru/progr/alg/ga/