

ИСПОЛЬЗОВАНИЕ ЧЕТЫРЕХЗНАЧНЫХ ЛОГИЧЕСКИХ СИСТЕМ ДЛЯ АНАЛИЗА ДНК- ВЫЧИСЛЕНИЙ

Л.А.Лютикова

НИИ ПМА КБНЦ РАН, г. Нальчик

lylarisa@yandex.ru

Аннотация.

В данной работе рассматривается вопрос о формализации свойств ДНК методами четырехзначной логики. Исследуется вопрос о возможных построениях ДНК-логических алгоритмов для решения задач распознавания, предлагается алгоритм решения поставленной задачи и логическая функция реализующая предложенный алгоритм.

ДНК-вычисления — это раздел области молекулярных вычислений на границе молекулярной биологии и компьютерных наук. Основная идея ДНК-вычислений — построение новой парадигмы, создание новых алгоритмов вычислений на основе знаний о строении и функциях молекулы ДНК и операций, которые выполняются в живых клетках над молекулами ДНК при помощи различных ферментов.

Одно из преимуществ ДНК-процессоров в сравнении с обычными кремниевыми процессорами заключается в том, что они могут производить все вычисления не последовательно, а параллельно, что обеспечивает выполнение сложнейших математических расчетов буквально за считанные минуты. Традиционным компьютерам для выполнения таких расчетов потребовались бы месяцы и годы

В молекулах ДНК имеется четыре базовых основания: аденин (А), гуанин (G), цитозин (С) и тимин (Т), связанных друг с другом в цепочку.

Здесь используется не двоичная, а четверичная логика. И подобно тому, как в двоичной логике любую информацию можно закодировать в виде последовательности нулей и единиц, в молекулах ДНК можно кодировать любую информацию путем сочетания базовых оснований.

- А – 0 С – 1 G - 2 Т - 3

Комплементарность оснований заключается в том, что образование водородных связей.

При соединении одинарных цепочек ДНК в двойную цепочку возможно только между парами А – Т и G - С. Введем функцию $K(x)$, характеризующую комплементарные основания, заданную следующей таблицей:

$K(x)=3-x$... Таблица 1

x	$K(x)$
0	3
1	2
2	1
3	0

Процесс соединения двух одинарных цепочек ДНК путем связывания комплементарных оснований в регулярную двойную спираль называется ренатурацией,

описывается следующей логической функцией $Re(x_1, x_2)$, которая имеет приведенную ниже таблицу истинности

$$j_3(x_1 \oplus x_2) = x_1 \Leftrightarrow \overline{x_2} = Re(x_1, x_2)$$

$Re(x_1, x_2)$ Таблица 2

x_1/x_2	0	1	2	3
0	0	0	0	1
1	0	0	1	0
2	0	1	0	0
3	1	0	0	0

Обратный процесс, то есть разъединение двойной цепочки и получение двух одинарных цепочек, — денатурацией.

Копирование, или размножение, ДНК-молекул осуществляется в ходе полимеразной цепной реакции (Polymerase Chain Reaction, PCR). Процесс копирования можно разделить на несколько стадий. Он происходит лавинообразно. На первом шаге из одной молекулы образуются две, на втором — из двух молекул — четыре, а после n -шагов получается уже 2^n молекул.

$$Cop(x) = x \vee x$$

Задача распознавания образов имеет следующую формулировку:

Соответствие множества объектов характеризующим их признаков может быть представлено следующей таблицей:

Таблица 3

x_1	x_2	$x_3 \dots ..$	x_n	W
$x_1(w_1)$	$x_2(w_1)$	$x_3(w_1)$	$x_n(w_1)$	w_1
$x_1(w_2)$	$x_2(w_2)$	$x_3(w_2)$	$x_n(w_2)$	w_2
$x_1(w_m)$	$x_2(w_m)$	$x_3(w_m)$	$x_n(w_m)$	w_m

$X_j = \{x_1(w_j), x_2(w_j), \dots, x_n(w_j)\}$ - вектор качественных признаков, каждый элемент которого - фиксированный признак характеризуемого объекта.

$$W = \bigcup_{j=1}^m w_j - \text{множество характеризуемых объектов.}$$

Вид функции $W = f(X)$ не задан. Требуется восстановить неизвестную зависимость по наблюдениям.

Каждый соответствующий признак $x_j(w_j)$ в общем случае кодируется предикатом k_i -значности $[0, 1, 2, 3]$

Для нахождения значения функции $W = f(X)$ системе важно обращаться к базе знаний, которая по запросу $X_j = \{x_1, x_2, \dots, x_n\}$ выдаст w_j , в случае если w_j принадлежит исследуемой предметной области, или объект (группу объектов), наиболее соответствующих данному запросу/

Для нахождения наиболее соответствующего заданному запросу $X_j = \{x_1, x_2, \dots, x_n\}$ объекта w_j , предлагается следующий алгоритм:

Запрос формируется в виде комплементарной цепочки к исходному запросу.

Копируется необходимое количество цепочек запроса.

Добавляются необходимые ферменты и через несколько секунд в результате реакции ренатурации синтезируются всевозможные ответы на данный запрос.

Выбирается максимальная из двойных цепочек ДНК.

$$F(X) = \max[\text{Re}(\text{Cop}_{j=1}^{m/2}(K(x_i)) \vee_{j=1}^m X_j)]$$

Литература:

1. Малинецкий Г.Г., Митин Н.А., Науменко С.А. Нанобиология и синергетика. Проблемы и идеи. ИПМ им. М.В.Келдыша РАН. Москва, 2005
2. Паун Г., ДНК-компьютер: Новая парадигма вычислений