

ИММУННЫЕ СИСТЕМЫ ДЛЯ РЕШЕНИЯ ПРАКТИЧЕСКИХ ЗАДАЧ

С.А. Челнаков, С.А. Ушаков

*ФГБОУ ВПО «Воронежский государственный педагогический университет»;
ФГБОУ ВПО «Воронежский государственный университет»,
г. Воронеж*

Рецензент д-р техн. наук, профессор И.Ф. Астахова

Ключевые слова и фразы: иммунные системы; компьютерная реализация; распознавание символов.

Аннотация: Рассматривается задача распознавания символов с помощью искусственных иммунных систем. Предлагается программный продукт, реализующий такую систему.

Теоретическая часть

Естественная иммунная система (ИС) представляет собой сложную систему, состоящую из нескольких функционально различных компонентов. Иммунная система использует многоуровневую защиту против внешних антигенов, включая действие неспецифических (врожденных) и специфических (приобретенных) защитных механизмов. Основная роль ИС заключается в распознавании всех клеток (или молекул) организма и классификации их как «своих» или «чужих». Чужеродные клетки подвергаются дальнейшей классификации с целью стимуляции защитного механизма соответствующего типа. В процессе эволюции ИС обучается различать внешние антигены (например, бактерии или вирусы) и собственные клетки или молекулы организма.

При попадании антигена в организм лишь малая часть клеток ИС способна к распознаванию его пептидов. Такое распознавание стимулирует процессы размножения и дифференцировки лимфоцитов, приводящие к образованию клонов идентичных клеток (или антител). Этот процесс, называемый размножением клона, формирует многочисленную популяцию специфичных к антигену антителопродуцирующих клеток. Размножение

Челнаков Сергей Александрович – аспирант кафедры «Информатика и методика преподавания математики», e-mail: stanlys@yandex.ru, ФГБОУ ВПО «Воронежский государственный педагогический университет»; Ушаков Станислав Андреевич – магистрант кафедры «Математическое обеспечение ЭВМ», ФГБОУ ВПО «Воронежский государственный университет», г. Воронеж.

клона иммунокомпетентных клеток приводит к разрушению или нейтрализации антигена. Часть образовавшихся клеток сохраняется для иммунной памяти. В результате, последующее воздействие похожего антигена приводит к более быстрой иммунной реакции (вторичному ответу).

Существует два основных варианта иммунного ответа: гуморальный, опосредованный В-клетками и их продуктами, и клеточный, в котором участвуют Т-клетки. Обоим вариантам иммунных реакций соответствуют сходные последовательности этапов защиты организма: активация, деление, дифференцировка, секреция, иммунная атака, супрессия и память, однако они осуществляются различными способами. В регуляции как гуморального, так и клеточного иммунитета участвуют популяции Th2-клеток, называемые хелперными или супрессорными клетками, которые усиливают либо соответственно подавляют иммунный ответ.

На основе искусственных ИС разрабатывается и реализуется алгоритм, позволяющий по рисунку с изображением символа возвращать его значение.

Данную задачу можно разбить на следующие подзадачи:

а) выделение подмножества существенных черт и особенностей естественной иммунной системы, необходимых для решения данной задачи;

б) описание искусственной иммунной сети и ее компонентов с учетом требований, полученных на предыдущем этапе;

в) построение обучающей выборки – набора изображений букв для обучения сети;

г) разработка собственно алгоритма для полученной искусственной иммунной сети;

д) разработка интерфейса пользователя.

Данные подзадачи должны выполняться последовательно.

Классификация отдельных клеток и молекул является важнейшим свойством естественной иммунной сети. Условно можно выделить 2 вида такой классификации: классификация «свой–чужой» и более детальный анализ тех клеток и молекул, которые были классифицированы как чужеродные. В данной задаче первый тип использоваться не будет, так как он не находит отражения в исследуемой задаче распознавания. Второй же тип является именно тем механизмом, который позволит распознавать различные образы. Таким образом, изображения будут играть роль антигенов, а иммунная система на основе реакции В-лимфоцитов будет определять тип этого антигена. Данная задача имеет существенные особенности, отметим важнейшие из них.

В отличие от естественной ИС, количество антигенов в данной задаче невелико и известно заранее. Учитывая это, каждому лимфоциту можно приписать номер антигена, на который данный лимфоцит будет реагировать. Это позволит упростить процедуру нахождения количества лимфоцитов, специфичных именно к этому антигену.

В естественной ИС классификация происходит за счет химических реакций. Для искусственной ИС можно предложить следующий способ, монохромное изображение можно представить как матрицу размера $M \times N$ булевых значений, где true стоит на тех местах, которые в изображении

соответствуют пикселям черного цвета, формирующими изображение. Лимфоцит, в свою очередь, можно представить как массив из P пар чисел. Каждая такая пара представляет собой координаты пикселя в изображении. Введем некое число $S < P$. Будем говорить, что лимфоцит реагирует на данный антиген, если количество элементов матрицы антигена со значениями true, определяемых номерами строк и столбцов массива пар чисел лимфоцита, больше либо равно S . На основе данного определения можно посчитать количество лимфоцитов, реагирующих на данный антиген. Так как каждый лимфоцит имеет приписанный номер антигена (идеального или образцового), то можно узнать, на какой антиген должно реагировать большинство из этих лимфоцитов. В идеальном случае (если антиген подан без искажений и иммунная сеть хорошо обучена) именно идентификатор этого антигена и будет ответом. Однако в реализации пришлось использовать не такой простой критерий, как подсчет числа среагировавших лимфоцитов.

В качестве замечания можно добавить, что вместо абсолютной величины S можно использовать относительную, которая будет представлять собой минимальный процент от числа P , необходимый для того, чтобы лимфоцит среагировал на данный антиген. Также можно ввести условную меру аффинности данному антигену для каждого лимфоцита. Она будет представлять собой отношение количества элементов матрицы антигена со значениями true, определяемых номерами строк и столбцов массива пар чисел лимфоцита, к общему числу пар чисел в лимфоците.

Можно отметить следующее упрощение модели естественной ИС, будем рассматривать только В-лимфоциты, причем все из них будут использоваться для классификации каждого антигена, Т-лимфоциты в данной искусственной системе рассматриваться не будут. Таким образом, получили описание используемой ИС.

Описание работы алгоритма

Шаг 0. Обучение иммунной сети.

Этот шаг можно сравнить с комплексной вакцинацией организма, у которого до этого ИС отсутствовала. В обучающейся выборке находятся образцы всех распознаваемых символов (антигенов) в нескольких различных вариантах. Для каждого варианта формируется множество лимфоцитов, реагирующих на данный антиген. В зависимости от порога реагирования, количества генерируемых лимфоцитов, размера лимфоцита этот шаг может занимать большое количество ресурсов.

Шаг 1. Собственно распознавание.

Подается распознаваемый образ, создается и заполняется нулями массив, индексированный номерами антигенов. Для всех лимфоцитов сети вычисляется аффинность к данному антигену, если она больше некоторого порогового значения (являющегося одним из ключевых параметров алгоритма), то есть если лимфоцит реагирует на этот антиген, то элемент массива с индексом, равным номеру антигена (приписанный лимфоциту при

создании), увеличивается на величину, зависящую от аффинности. Затем в полученном массиве находится максимальный элемент, его индекс и будет ответом иммунной сети.

Хранение ИС в постоянной памяти

Сформированную и обученную ИС надо хранить в энергонезависимой памяти между сеансами работы программы. Иммунная сеть определяется следующими параметрами, которые и подлежат сохранению в файле на жестком диске:

- количество распознаваемых символов – антигенов;
- размеры эталонных изображений;
- собственно иммунная сеть – совокупность всех лимфоцитов, ее составляющих;
- для каждого лимфоцита надо хранить следующую информацию:
 - 1) номер антигена, на который реагирует данный лимфоцит;
 - 2) списки координат значащих и незначащих пикселей символов, представляющих собой набор пар чисел (X, Y) .

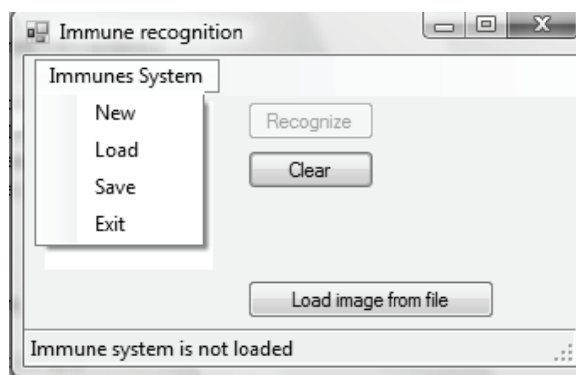
В качестве формата хранения этих данных был выбран формат XML. Обладая простой структурой, он в то же время позволяет эффективно реализовать хранение различных данных. Для решения проблемы сохранения и загрузки иммунной сети из XML-файла в класс ImmuneAlgorithm были добавлены следующие методы:

- Public Void SaveToXMLFile (String Filename);
- Public Static ImmuneAlgorithm LoadFromXML (String Filename).

Первый метод обеспечивает сохранение текущей иммунной сети в файл с именем Filename, второй же метод (статический) возвращает экземпляр класса ImmuneAlgorithm, все данные для которого были загружены из файла с именем Filename соответственно.

Интерфейс пользователя

Интерфейс пользователя позволяет задавать символ для распознавания двумя способами: загрузить из файла или ввести в интерактивном режиме с помощью мыши (рисунок).



Главное меню

Для разработки была использована среда Visual Studio 2008 Professional, язык программирования – C#. Этот язык является полностью объектно-ориентированным, что позволяет создать систему классов, каждый из которых представляет ту или иную сущность в данной предметной области. Среда Visual Studio обладает большим количеством необходимых стандартных компонентов, облегчающих создание удобного интерфейса пользователя, также платформа .Net обеспечивает удобную работу с графическими файлами, позволяет легко создавать многооконные приложения. Так как приложение написано в Visual Studio на платформе .Net, то необходимым условием является наличие установленного .Net Framework версии 3.5.

Заключение

Рассмотрена возможность применения искусственных иммунных сетей для распознавания образов. Описаны основные понятия, используемые в теории искусственных ИС. Предложена структура искусственной иммунной сети, решающей поставленную проблему. Рассмотрено строение лимфоцитов, составляющих такую иммунную сеть и определение аффинности к антигену.

Список литературы

1. Искусственные иммунные системы и их применение / под ред. Д. Дасгупты. – М. : ФИЗМАТЛИТ, 2006 – 344 с.
2. Farmer, J.D. The Immune System, Adaptation and Machine Learning / J.D. Farmer, N. Packard, A. Perelson // Physica D. – 1986. – Vol. 2. – P. 187–204.

Immune Systems for Solving Practical Problems

S.A. Chelnakov, S.A. Ushakov

*Voronezh State Pedagogical University
Voronezh State University, Voronezh*

Key words and phrases: computer realization; immune systems; recognition of hand-printed characters.

Abstract: The problem of recognition of hand-printed characters using artificial immune systems is considered. It is proposed to use a software product that implements such a system.

© С.А. Челнаков, С.А. Ушаков, 2012