

ПРИМЕНЕНИЕ АДАПТИВНОГО ГЕНЕТИЧЕСКОГО АЛГОРИТМА ДЛЯ ГЕНЕРАЦИИ КЛЕТОЧНЫХ АВТОМАТОВ

А.В. Тихомиров, А.А. Шалыто

Рассматривается улучшенный алгоритм генерации произвольных клеточных автоматов на основе тестовых наборов при помощи адаптивного генетического алгоритма. Описаны основные отличия и преимущества по сравнению со стандартным генетическим алгоритмом. Произведена апробация на нескольких обучающих примерах.

Ключевые слова: клеточные автоматы, генетические алгоритмы.

Введение

Настоящая работа развивает идеи, описанные в [1]. Для генерации различных клеточных автоматов используется адаптивный генетический алгоритм (ГА), который устраняет многие проблемы стандартного ГА. Алгоритм используется для моделирования многих физических процессов, например: диффузия энергии, разнообразные химические реакции, рост кристаллов и т.д. [2–5].

ГА, описанный в работе [1], имеет недостатки – он не адаптируется под текущее состояние популяции, имеет проблемы со сходимостью и временем работы. Цель настоящей работы – устранить недостатки ГА, разработав адаптивный ГА.

Постановка задачи

Задана квадратная сетка и автомат, управляющий всеми ячейками [6–8]. При этом задаются данные для различных временных шагов, т.е. начальное состояние системы, несколько промежуточных состояний и конечное состояние системы.

Целью работы алгоритма является выращивание автомата, наиболее точно описывающего заданную физическую систему.

Строение хромосомы клеточного автомата

Обычно при использовании генетического подхода используется метод, при котором данные в хромосоме хранятся в виде строки [6]. Однако в рассматриваемой задаче длина хромосомы может варьироваться в зависимости от количества переходов между состояниями, поэтому ген представляет собой не строку, а специализированную структуру.

Количество базовых генов соответствует количеству состояний в клеточном автомате. Таким образом, базовый ген под названием «StateGene» описывает определенное состояние автомата. «StateGene» состоит из генов переходов клеточного автомата, которые называются «TransitionGene». Каждый ген перехода состоит из списка генов условий на переход и генов действий, которые будут осуществляться в случае этого перехода. Длина хромосомы зависит от количества состояний в клеточном автомате, который описывает данная хромосома.

В случае отсутствия генов условий подразумевается, что условие на данный переход всегда является верным.

Начальное поколение состоит из фиксированного числа случайно генерированных автоматов. Все автоматы в поколении имеют одинаковое наперед заданное количество состояний, но количество переходов между состояниями может варьироваться.

Проблема вырождения популяции

Одной из проблем в применении генетического подхода для генерации клеточных автоматов стало вырождение популяции, т.е. такой ситуации, когда выделяется один-единственный генотип, который представляет собой локальный максимум, а затем все элементы популяции проигрывают ему отбор, и вся популяция «забывается» копиями этой особи. Из-за схожего строения хромосом операция скрещивания становится практически бесполезной, и вероятность получить хромосому, которая лучше решает поставленную задачу, сильно падает. В настоящей работе рассматриваются несколько способов преодолеть этот недостаток:

- алгоритм каскадной фитнес-функции;
- операция инъекции хромосом;
- адаптивный генетический алгоритм.

Все эти модификации позволяют улучшить «сходимость» генетического подхода при генерации клеточных автоматов.

Операция скрещивания

В зависимости от режима работы генератора используется скрещивание двух видов – симметричное и несимметричное. Заметим, что в обоих видах скрещивания производится обмен генами одного типа в соответствующих участках хромосомы.

Особенность симметричного скрещивания заключается в том, что происходит обмен соответствующими участками поддеревьев генов состояний между хромосомами. Особенность несимметричного скрещивания заключается в том, что происходит обмен случайными поддеревьями одного уровня между хромосомами (рис. 1).

Для оптимизации операции скрещивания применяется анализ хромосом на совпадающие гены, т.е. из каждой пары хромосом, участвующих в скрещивании, выделяется список уникальных генов. Именно среди этих генов и производится поиск для обмена. Такая операция позволяет совершать меньшее число скрещиваний и заранее исключать операции, которые приведут к хромосомам, которые уже есть в популяции.

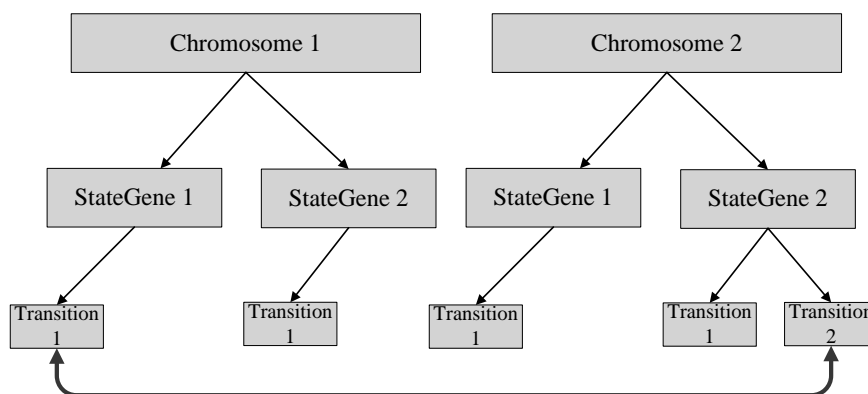


Рис. 1. Пример несимметричного скрещивания хромосом

Операция инъекции хромосом

Для добавления особей, полученных из предыдущих запусков генератора, используется операция инъекции хромосом в новую популяцию. Следует отметить, что в стандартных генетических алгоритмах нет предложенной ниже операции, она была разработана специально для ГА, предложенного в настоящей работе. Цель этой операции – добавление частей хромосом, которые хранятся с предыдущих запусков генетического алгоритма, в новую популяцию, для возможного улучшения особей за счет материала, накопленного за предыдущие запуски.

Ниже предложены основные этапы работы этой операции:

- выбор особи из текущей популяции;
- выбор особи из хранилища результатов предыдущих запусков генератора;
- получение «разности» между хромосомой из хранилища результатов предыдущих запусков генератора и хромосомой из популяции;
- операция скрещивания между разностной хромосомой и хромосомой из популяции.

Фитнесс-функция

Функция приспособленности особи, т.е. клеточного автомата, состоит из нескольких значений. Первое значение характеризует автомат с точки зрения решения поставленной задачи, при моделировании сравниваются эталонные и полученные результаты для каждого шага моделирования. Второе значение характеризует автомат с точки зрения избыточности, в нем хранится «длина автомата» и количество неиспользуемых переходов и условий. Для борьбы с вырождением популяции была применена каскадная фитнес-функция. Идея заключается в том, что на начальных этапах выращивания клеточных автоматов функция приспособленности учитывает только несколько первых временных срезов входных данных. Когда в популяции есть клеточные автоматы, которые в какой-то степени удовлетворяют текущей фитнес-функции, то происходит смена на функцию с учетом большего количества слоев.

Ниже рассмотрены «минусы» использования каскадной функции приспособленности:

- дополнительные затраты при смене функции;
- сложность сравнения эффективности особей (особенно это проявляется при сравнении особей из разных «островов»).

Ниже рассмотрены «плюсы» использования каскадной фитнес-функции:

- прирост производительности (на ранних этапах уменьшается количество этапов проверки клеточного автомата).
- после каждого этапа полученный автомат частично решает поставленную задачу.

Генетический алгоритм

Основной проблемой при совместном использовании генетического подхода и клеточных автоматов являются вырождение и стагнация популяции, поэтому возникла необходимость в разработке специального ГА, который бы учитывал текущее состояние популяции и текущую стадию генерации.

Предложенный ГА состоит из нескольких состояний и, в сущности, представляет собой конечный автомат с правилами переключения между стадиями генератора. Ниже перечислены разработанные состояния генератора (рис. 2):

1. нормальный режим работы генератора клеточных автоматов;
 2. форсированный режим работы, который применяется при стагнации процесса генерации;
 3. режим доводки решения, который начинает работать при приближении к глобальному минимуму.
- Заметим, что из режима доводки решения генератор уже не будет перезапускаться даже в состоянии стагнации.



Рис. 2. Схема генетического алгоритма

Чтобы избежать локальных минимумов функции приспособленности и вырождения популяции, применяются два алгоритма отбора в новое поколение.

1. Происходит отбор первых P особей с лучшей функцией приспособленности. Все они попадают в новое поколение.
2. Если на протяжении достаточно большого числа поколений не происходит улучшения значения функции приспособленности, то при обработке текущего поколения отбрасываются все особи, кроме нескольких, которые имеют наилучшее значение, т.е. наиболее приспособленных. Оставшееся место в новом поколении занимают особи, полученные из особей текущего поколения путем скрещивания и мутаций. Мутация может происходить не в единственном гене хромосомы, как при обычных условиях, а во всех генах хромосомы с некоторой вероятностью.

При отборе первых P особей при одинаковой функции приспособленности выбираются более молодые особи.

Полученные результаты

Для тестового примера была проведена апробация разработанного метода для генерации клеточных автоматов для игр на двумерном поле. Основным отличием таких клеточных автоматов является то,

что они не хранят в себе данных и меняют свое состояние только в зависимости от состояний окружающих ячеек.

Далее приведены два примера апробации предложенного адаптивного ГА.

Игра «Жизнь». Игра «Жизнь» является одним из наиболее известных клеточных автоматов для двумерного поля. Она была создана Джоном Хортоном Конвеем в 1970 г. Процесс эволюции поля эмулирует процессы, происходящие при зарождении, развитии и гибели колонии живых организмов. В качестве тестовых примеров на вход генератору подавались наборы полей и эталонные примеры работы (рис. 3).

В силу простого строения автомата и, как следствие, малого числа генов улучшение фитнес-функции происходит в основном за счет многочисленных мутаций особей. Было установлено, что использование ГА является неэффективным. Применение эволюционной стратегии на всех этапах генерации клеточного автомата позволило улучшить сходимость предлагаемого метода для генерации клеточного автомата «Жизнь». Для сравнения: при использовании эволюционной стратегии генератор перебирает порядка 200 особей, при ГА – более 2000 особей.

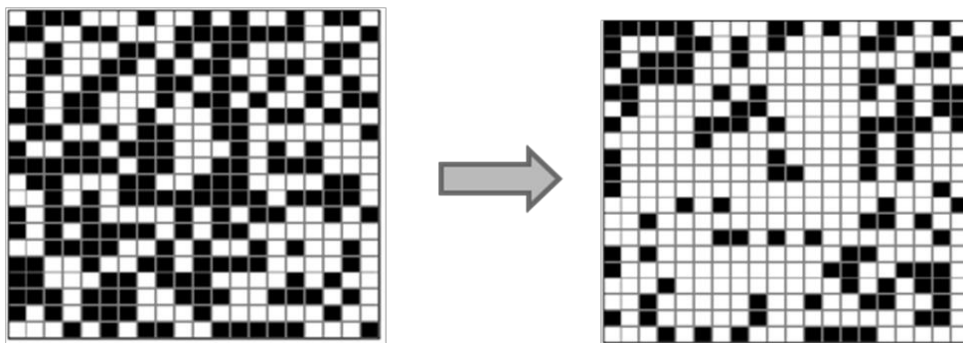


Рис. 3. Пример тестового набора из двух шагов симуляции для игры «Жизнь»

Общий случай игр на двумерном поле. Для тестового примера была поставлена задача получить клеточный автомат, зная только состояния клеток на каждом конкретном шаге расчетов. Для этого брались случайный эталонный автомат, количество состояний которого варьировалось в пределах от пяти до восьми, и на его основе генерировались тестовые наборы. Для ускорения расчетов функции приспособленности все тестовые наборы имели размеры тестового поля 40×40 (рис. 4).

Для генерации автоматов игр на двумерном поле использовались следующие настройки генератора: число островов 1–4; вероятность мутации 0,1; число особей на острове 50; доля элитизма 0,1; период миграции 50.

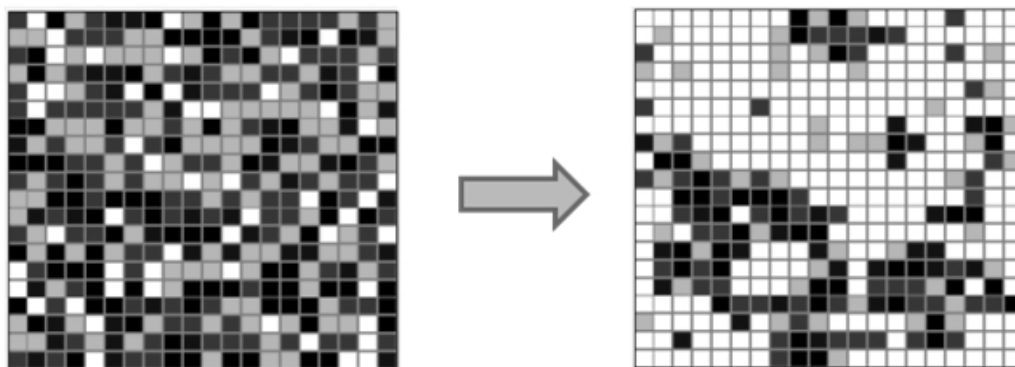


Рис. 4. Пример тестового набора из двух шагов симуляции произвольного клеточного автомата

Быстродействие предложенного способа сильно зависит от начальных настроек, таких как пропорции между генетическими операциями, величина начальной популяции и многих других.

Также было установлено, что на быстродействие описанного в работе ГА влияет наличие точных условий, содержащихся в клеточном автомате. В таблице представлены значения числа вычисления функции приспособленности для двух клеточных автоматов, отличающихся всего одним условием:

- $n(3)=3$ у первого выращиваемого эталона;
- $n(3)<3$ у второго выращиваемого эталона.

	N_{\min}	N_{avg}	N_{\max}
Automaton 1	27106	52315	102411
Automaton 2	23594	30173	44862

Таблица. Сравнение времени работы генератора для генерации эталонов с различием в одном условии (N – число вычислений фитнес-функции)

Заключение

В работе предложен метод генерации клеточных автоматов произвольного количества состояний, учитывающий текущее состояние популяции. Он позволяет автоматически получать клеточные автоматы, которые довольно точно описывают данные о физической системе на входе. Также этот метод позволяет получать клеточные автоматы, которые затруднительно реализовать эвристическими методами. Алгоритм апробирован на ряде задач, таких как игра «Жизнь» и игры на двумерном поле. Показано, что адаптивный алгоритм умеет приспосабливаться к текущему состоянию генерируемых особей и выбирать оптимальную вариацию генетического алгоритма.

Литература

1. Тихомиров А.В., Шалыто А.А. Применение генетического подхода для генерации клеточных автоматов // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 2. – С. 62–66.
2. Тоффли Т., Марголус Н. Машины клеточных автоматов. – М.: Мир, 1991. – 280 с.
3. Frish U. Lattice gas hydrodynamics in two and three dimensions // Complex Systems. – 1987. – V. 1. – P. 649–707.
4. Wolfram S. Cellular automation Fluids // J. Stat. Phys. – 1986. – V. 45. – P. 471–526.
5. Фон Нейман Дж. Теория самовоспроизводящихся автоматов: Пер. с англ. – М.: Мир, 1971. – 384 с.
6. Царев Ф.Н., Шалыто А.А. Применение генетического программирования для генерации автомата в задаче об «умном муравье» [Электронный ресурс]. – Режим доступа: http://is.ifmo.ru/genalg/_ant_ga.pdf, св. Яз. рус. (дата обращения 25.11.2011).
7. Наумов Л.А. Метод введения обобщенных координат и инструментальное средство для автоматизации проектирования программного обеспечения вычислительных экспериментов с использованием клеточных автоматов. Дис. ... канд. техн. наук: 05.13.12. – СПб, 2007. – 283 с.
8. Скаков П.С. Классификация поведения одномерных клеточных автоматов. Магистерская диссертация [Электронный ресурс]. – Режим доступа: http://is.ifmo.ru/papers/_skakov_master.pdf, св. Яз. рус. (дата обращения 25.11.2011).

Тихомиров Андрей Владимирович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, and.tikhomirov@gmail.com

Шалыто Анатолий Абрамович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, shalyto@mail.ifmo.ru