Опубликовано в материалах 2-й межвузовской научной конференции по проблемам информатики СПИСОК-2011, с. 359-362.

С. Э. Вельдер
Санкт-Петербургский государственный университет информационных технологий, механики и оптики

Автоматические доказательства аналогов гипотезы Черни – Пэна

В докладе рассматриваются некоторые задачи теории синхронизируемых автоматов. Проблемы, связанные с синхронизируемостью, встречаются в различных областях информационных технологий, таких как теория кодирования, робототехника, и др. Свойство синхронизируемости детерминированного конечного автомата в узком смысле означает существование *синхронизирующего слова* – последовательности символов (команд), переводящей все состояния автомата в одно и то же состояние. То есть синхронизирующее слово переводит множество всех состояний автомата в синглетон. Говорят, что слово *имеет дефект* k в заданном автомате из n состояний, если оно переводит множество всех его состояний (имеющее размер n) в множество размера $n - k$. Синхронизируемость автомата определяется существованием для него слова дефекта $n - 1$.

Одной из основных и наиболее сложных задач в этой теории является оценка минимальной длины (в худшем случае) синхронизирующего слова или слова заданного дефекта при условии, что такое слово существует [1]. Длину минимального слова дефекта k в худшем случае обозначим через $L(k)$. Гипотеза Ж.-Э. Пэна (1978) гласит, что $L(k) = k^2$. Её частный случай при $k = n - 1$ известен как гипотеза Я. Черни [2]. На функцию L известны следующие оценки (рис. 1):

$$k^2 \le L(k) \le k\,(k + 1)\,(k + 2) / 6 - 1$$

(первое неравенство приводится в [1], второе – в [3]).

Функция называется *перечислимой снизу* (*сверху*), если её подграфик (надграфик) перечислим. Функция является вычислимой тогда и только тогда, когда она перечислима и снизу, и сверху.

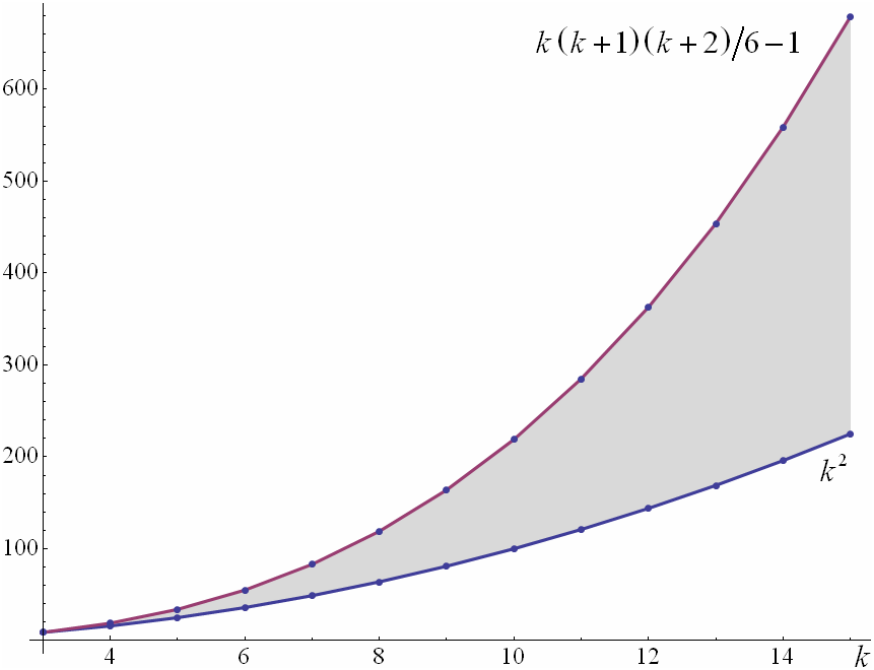

Рис. 1. Оценки функции Пэна $L(k)$

Функция L перечислима снизу путём перебора всех автоматов. Данный факт означает следующее. Рассмотрим двуместный предикат $P(k, l)$ = «$L(k) \leq l$»: он обозначает высказывание «в любом автомате, имеющем слово дефекта не менее k, найдётся слово дефекта не менее k и длины не более l». Если $P(k, l)$ не выполняется для некоторых k и l, то это можно подтвердить предъявлением соответствующего автомата – он будет являться алгоритмически проверяемым сертификатом, опровергающим $P(k, l)$. Возможность перебирать такие сертификаты – это и есть перечислимость снизу функции L.

В случае же, когда $P(k, l)$ верно для некоторых k и l, этот факт является неочевидной теоремой и доказывается для

каждых k и l по-своему. Действительно, это высказывание имеет квантификацию «для всех автоматов», а их бесконечно много.

Известны, например, следующие факты:

1) $P(0, 0)$ = «$L(0) \leq 0$», тривиально;
2) $P(1, 1)$ = «$L(1) \leq 1$», тривиально;
3) $P(2, 4)$ = «$L(2) \leq 4$», нетрудно;
4) $P(3, 9)$ = «$L(3) \leq 9$», Ж.-Э. Пэн [3];
5) $\neg P(4, 16)$ = «$L(4) > 16$», Я. Кари [4] – последний пример (опровергающий гипотезу Пэна) был получен полным перебором автоматов.

Результаты, полученные в работе, состоят в построении алгоритмов, позволяющих автоматически генерировать доказательства утверждений такого вида (когда они верны), а также определять случаи, когда они неверны, не перебирая все автоматы. Наличие такого алгоритма, в частности, доказывает, что функция L перечислима и сверху (т. е. вычислима).

Ключевая идея алгоритма состоит в следующем. Введём специальную комбинаторную структуру данных, которая будет иметь конечный размер и кодировать разбиения (бесконечного) множества всех автоматов на конечное число (бесконечных) классов эквивалентности. Процесс выполняется итеративно: сначала выполняется разбиение на небольшое число классов, потом каждый из полученных классов разбивается ещё на несколько и т. д. (рис. 2). Разбиения строятся на основе классификации наборов одноместных операций на конечном множестве. Каждая такая операция порождает ориентированный граф с постоянной исходящей степенью 1 (это совокупность циклов, оснащённых деревьями).

После этого выполняем перебор всех классов, и для каждого из них генерируем доказательство искомого утверждения. Доказательство основывается на анализе так называемого *экспоненциального автомата* (*Exp*-автомата) – автомата, состояния которого являются множествами состояний исходного. Каждый автомат (класс автоматов)

индуцирует свой экспоненциальный автомат (класс автоматов). Требуется доказать, что если в экспоненциальном классе автоматов существует путь из стартовой вершины в вершину дефекта не менее k, то существует и аналогичный путь длины не более l. Таким образом, в качестве доказательства для каждого класса автоматов предъявляется либо соответствующий путь в экспоненциальном автомате, либо сам экспоненциальный автомат, в котором нет достижимой вершины дефекта не менее k. Этот автомат используется в качестве сертификата отсутствия соответствующего пути.

Описанная регулярная структура, кодирующая рассматриваемое разбиение (вместе с доказательствами для каждого класса) играет роль автоматически проверяемого сертификата, подтверждающего истинность $P(k, l)$ для заданных k и l. Корректность алгоритма обусловлена наличием таких сертификатов.

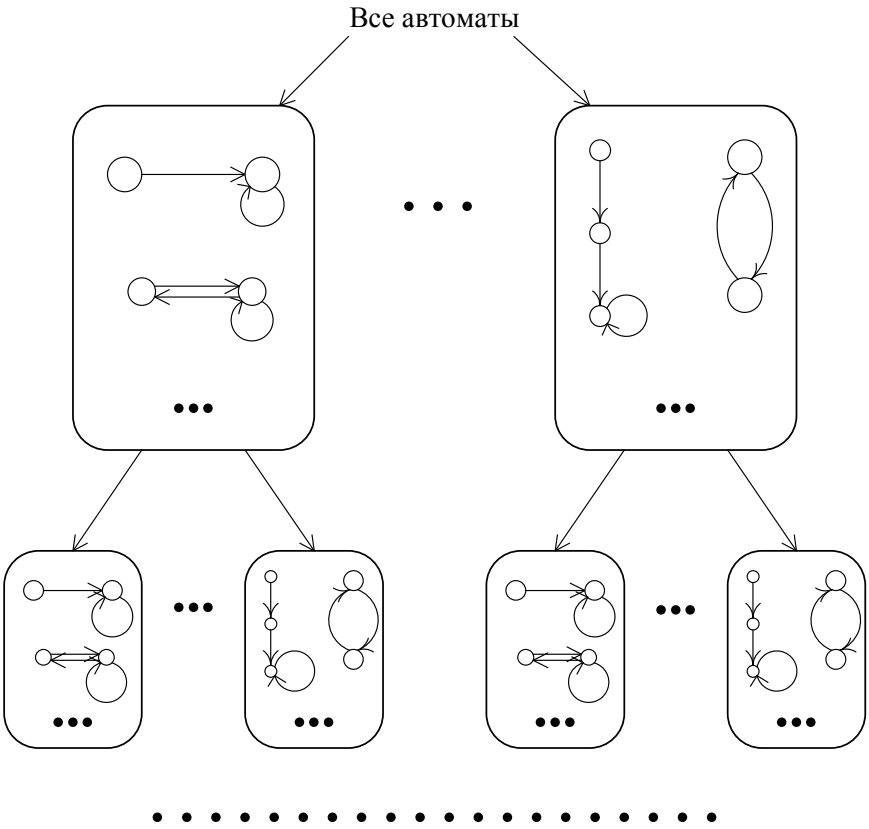

Рис. 2. Разбиение конечных автоматов на классы

Литература

1. *Pin J.-E.* On two combinatorial problems arising from automata theory, Annals of Discrete Mathematics 17 (1983), 535–548.
2. *Černý J.* Poznámka k homogénnym eksperimentom s konečnými automatami, Matematicko-fyzikalny Časopis Slovensk. Akad. Vied, 14, no. 3, 208–216.
3. *Pin J.-E.* Le problème de la synchronisation et la conjecture de Černý, Quaderni de la Ricerca Scientifica vol. 109, CNR, Roma, 1981.

4. *Kari J.* A counter example to a conjecture concerning synchronizing words in finite automata, EATCS Bull., 73, 146.