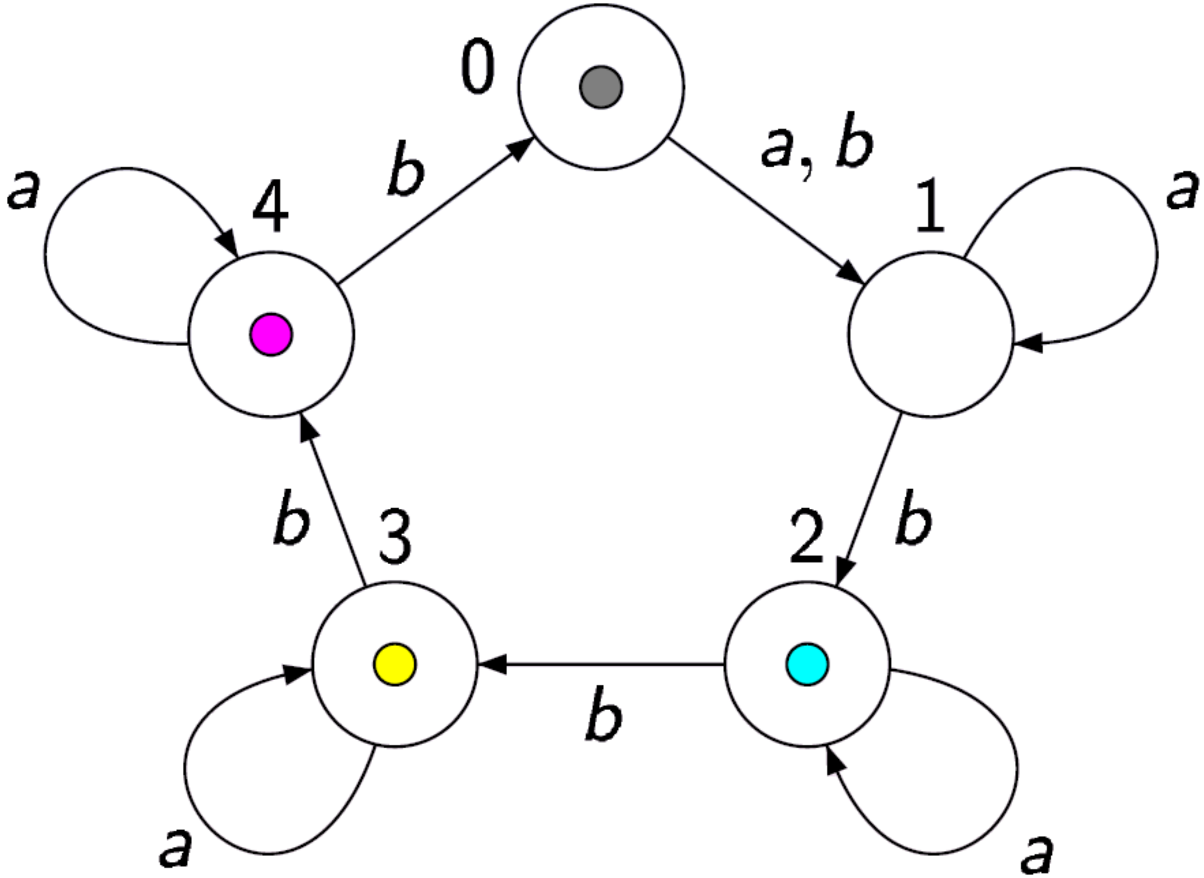
Оценки длины слов, имеющих заданный ранг относительно ДКА



Предметная область

Доклад относится к *теории синхронизируемости*.

Слово w называется *синхронизирующим* для заданного детерминированного конечного автомата, если оно переводит все его состояния в одно и то же состояние (любое).

ДКА называется *синхронизируемым*, если для него существует синхронизирующее слово.

Теория имеет обширные применения:

- реактивные системы (автономное управление, робототехника);
- теория кодирования (коды, исправляющие ошибки);
- ДНК-вычисления;
- теория Перрона — Фробениуса (неотрицательные матрицы).

Гипотеза Черни

Пусть ДКА из n состояний синхронизируем.

Какова длина $C(n)$ его кратчайшего синхронизирующего слова в худшем случае?

Гипотеза Черни (Jan Černý, 1964): $C(n) = (n-1)^2$.

Любой ДКА индуцирует *Exp-автомат* — автомат, состояния которого являются множествами состояний исходного.
Синхронизирующее слово переводит Exp-автомат из состояния мощности n в состояние мощности 1.

Будем говорить, что слово имеет *дефект* k (или *ранг* $n-k$), если оно переводит Exp-автомат из состояния мощности n в состояние мощности $n-k$.

Синхронизирующее слово имеет дефект $n-1$.

Обобщения

Пусть ДКА имеет слово дефекта k.

Какова длина $L(k)$ кратчайшего такого слова в худшем случае?

Гипотеза Пэна (Jean-Éric Pin, 1978)

Если ДКА имеет слово дефекта k, то $L(k) = k^2$.
(забегая вперёд: опровергнута в 2001 г.)

Гипотеза о рангах

Если ДКА имеет слово дефекта k и не имеет слов дефекта $> k$

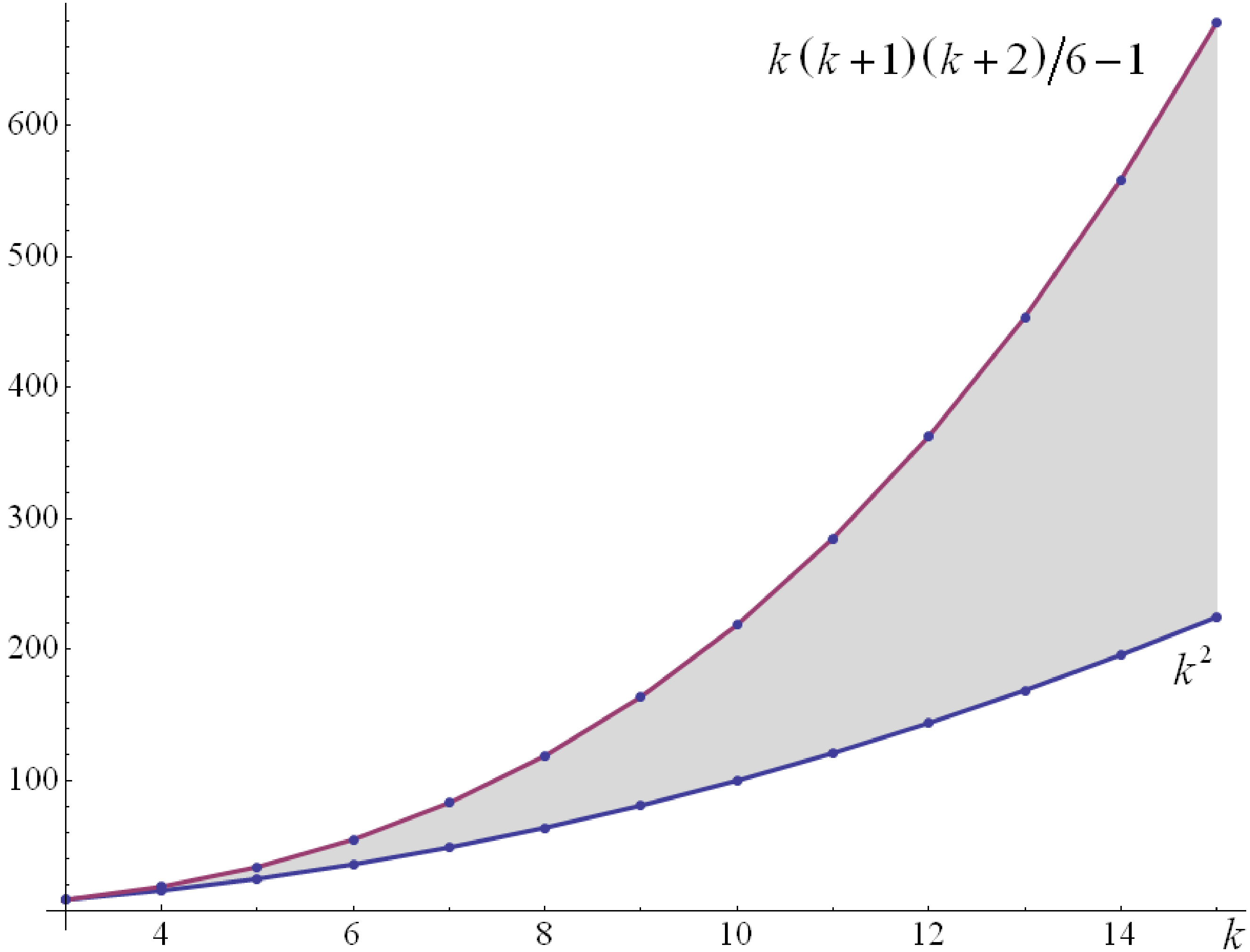
(в таких случаях говорят, что ДКА имеет *ранг* $n - k$), то $L(k) = k^2$.

Гипотеза Черни является частным случаем обеих этих гипотез

при $k = n - 1$.

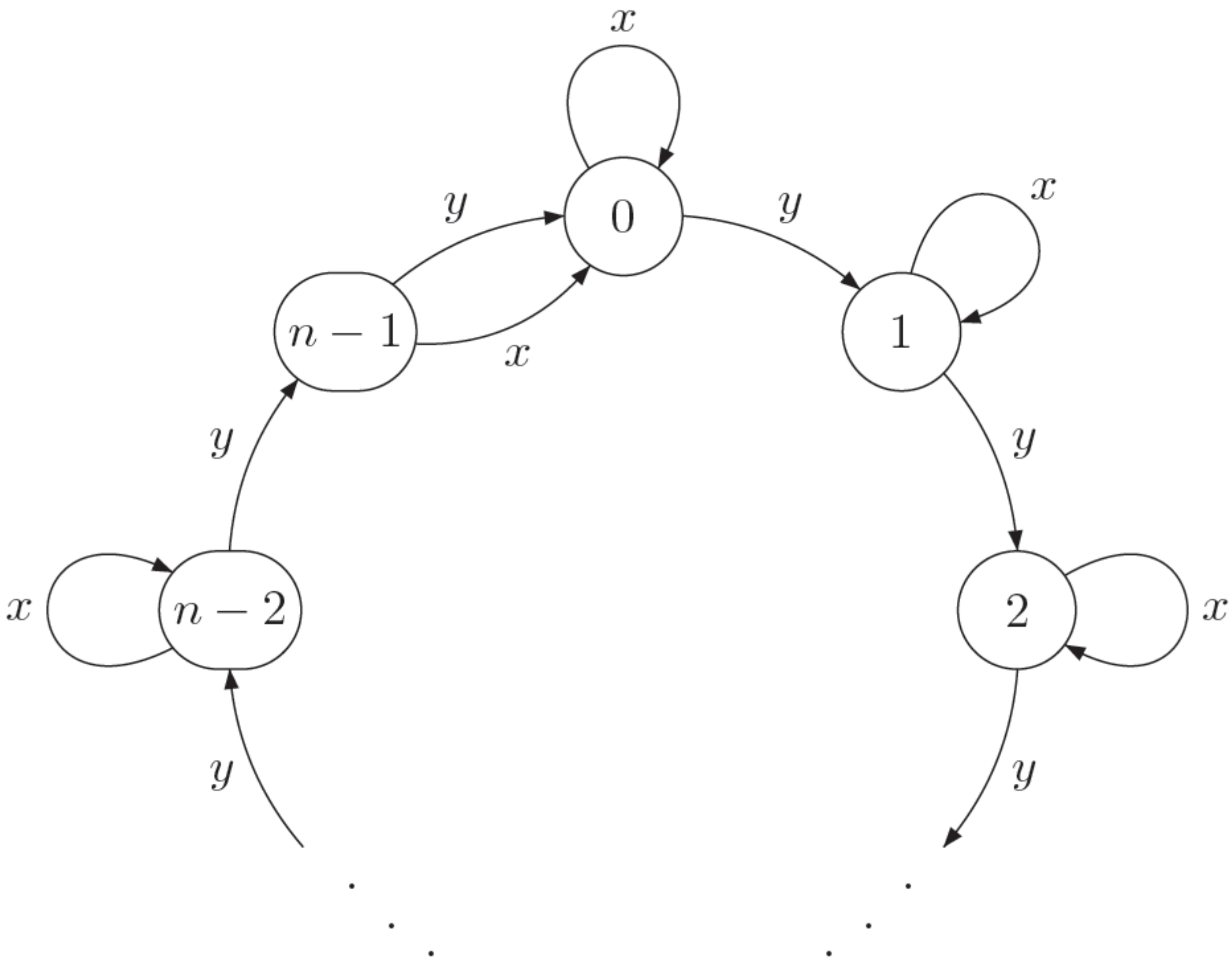
Границы

Известно, что $k^2 \leq L(k) \leq k\,(k+1)\,(k+2)\,/\,6 - 1$ (при $k \geq 3$) [P. Frankl]



Серия Черни: оценка синхронизируемости

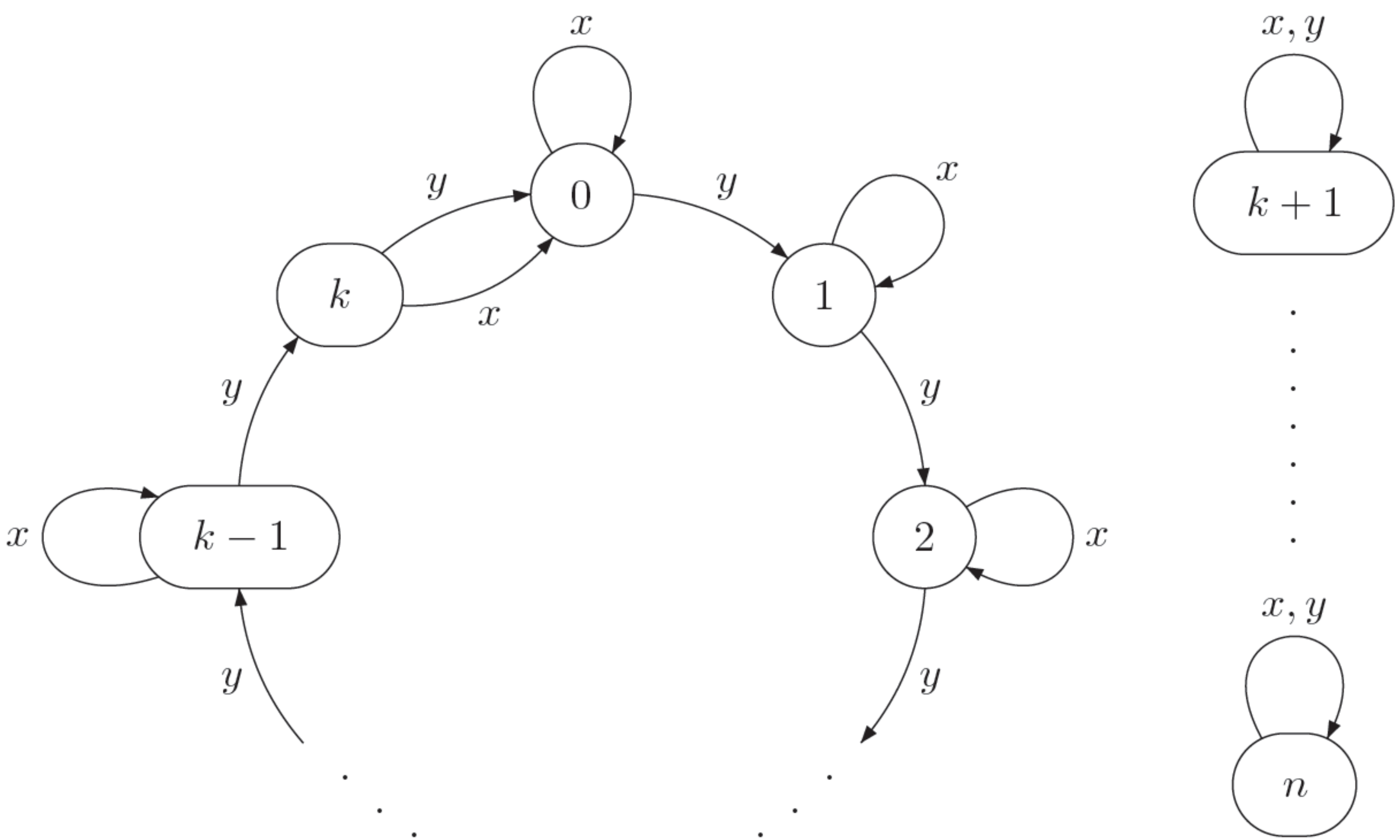
$$\mathscr{C}(n)$$



$$C(\mathscr{C}(n)) \geq (n-1)^2$$

Серия Черни: оценка сжимаемости

$\mathscr{C}(n,k)$



$$L(\mathscr{C}(n,k)) \geq k^2$$

Перечислимость снизу и сверху

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

Функция L перечислима снизу путём полного перебора всех ДКА. Более подробно: рассмотрим высказывание «$L(k) \leq l$» для некоторых k и l. Пусть оно неверно. Тогда существует ДКА, имеющий слово дефекта k, но не имеющий слова дефекта k и длины $\leq l$. Этот факт можно подтвердить предъявлением ДКА, который играет роль алгоритмически проверяемого сертификата.

Но как подтвердить высказывание «$L(k) \leq l$», если оно верно? Ведь оно начинается с квантора «*для всех автоматов*», а их бесконечно много!

Текущая ситуация

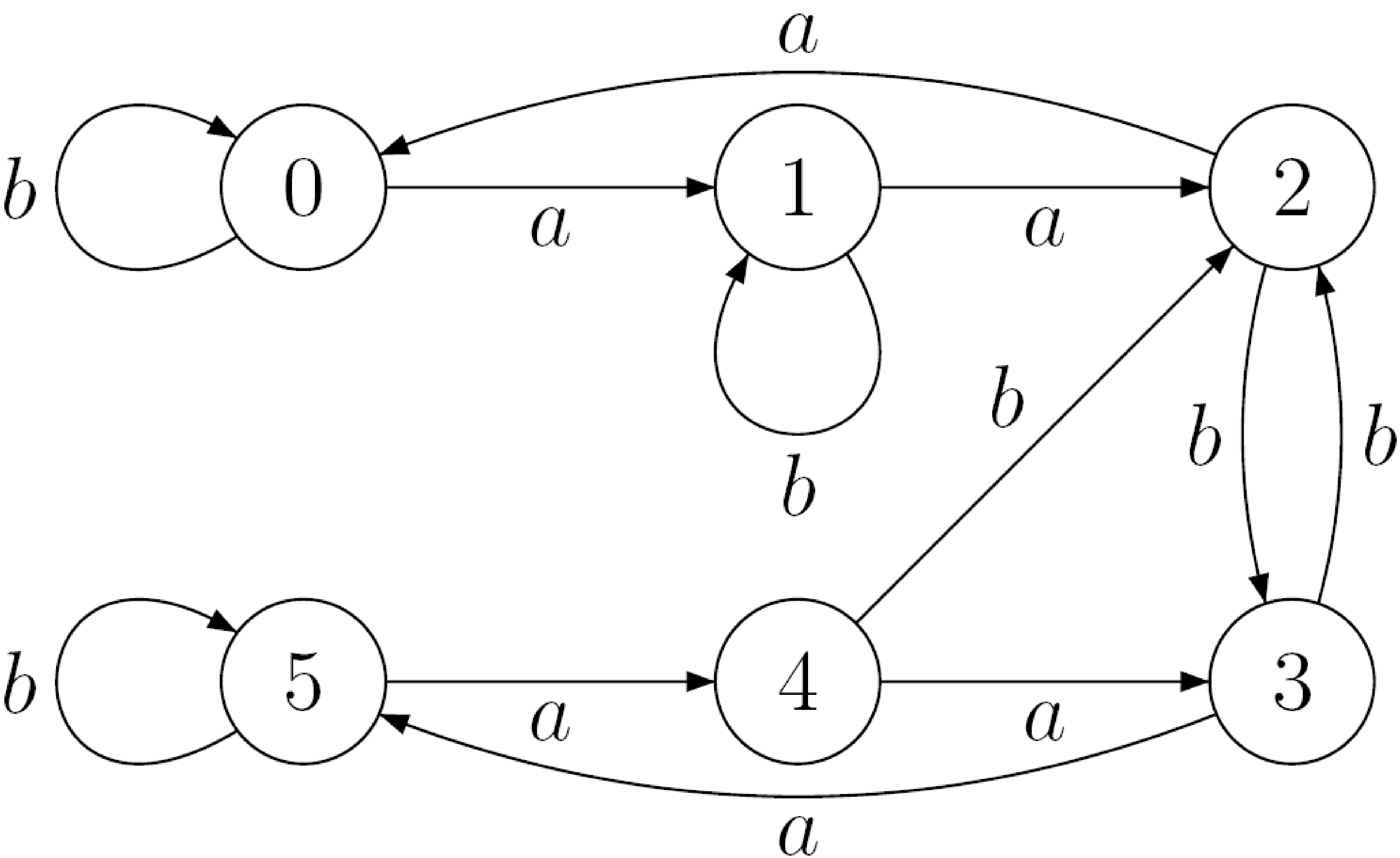
Раньше высказывание «$L(k) \leq l$»
для каждых k и l доказывали по-своему.

Например:

- $L(0) \leq 0$ (тривиально);
- $L(1) \leq 1$ (очевидно);
- $L(2) \leq 4$ (несложно);
- $L(3) \leq 9$ (Jean-Éric Pin, 1981);
- $L(4) > 16$ (Jarkko Kari, 2001)

— последний пример (опровергающий гипотезу Пэна)
был получен полным перебором автоматов.

Автомат Кари



Постановка задачи

Можно ли доказывать высказывания вида «$L(k) \le l$» автоматически?

Через $L_{\le n}(k)$ обозначим кратчайшую длину слова дефекта k среди автоматов из не более n состояний, обладающих таким словом.

Достаточно доказать, что $L(k) \le l \Leftrightarrow L_{\le f(k)}(k) \le l$ для некоторого f.

Схема доказательства.

Пусть автомат имеет слово дефекта k, но его минимальная длина больше l. Выделим из автомата «ядро», которое после определённой «доработки» станет автоматом, обладающим тем же свойством, но при этом его число состояний не превосходит $f(k)$.

Ключевая идея

Разобъём множество всех ДКА на *конечное* число классов эквивалентности. Каждый из этих классов имеет *бесконечную* мощность, но может быть закодирован *конечной* регулярной структурой данных. Каждому классу соответствует автомат-ядро.

Перебираем все классы, и для каждого генерируем доказательство.
Структура, кодирующая рассматриваемое разбиение, (вместе с доказательствами для каждого класса) является сертификатом, подтверждающим высказывание «$L(k) \leq l$».

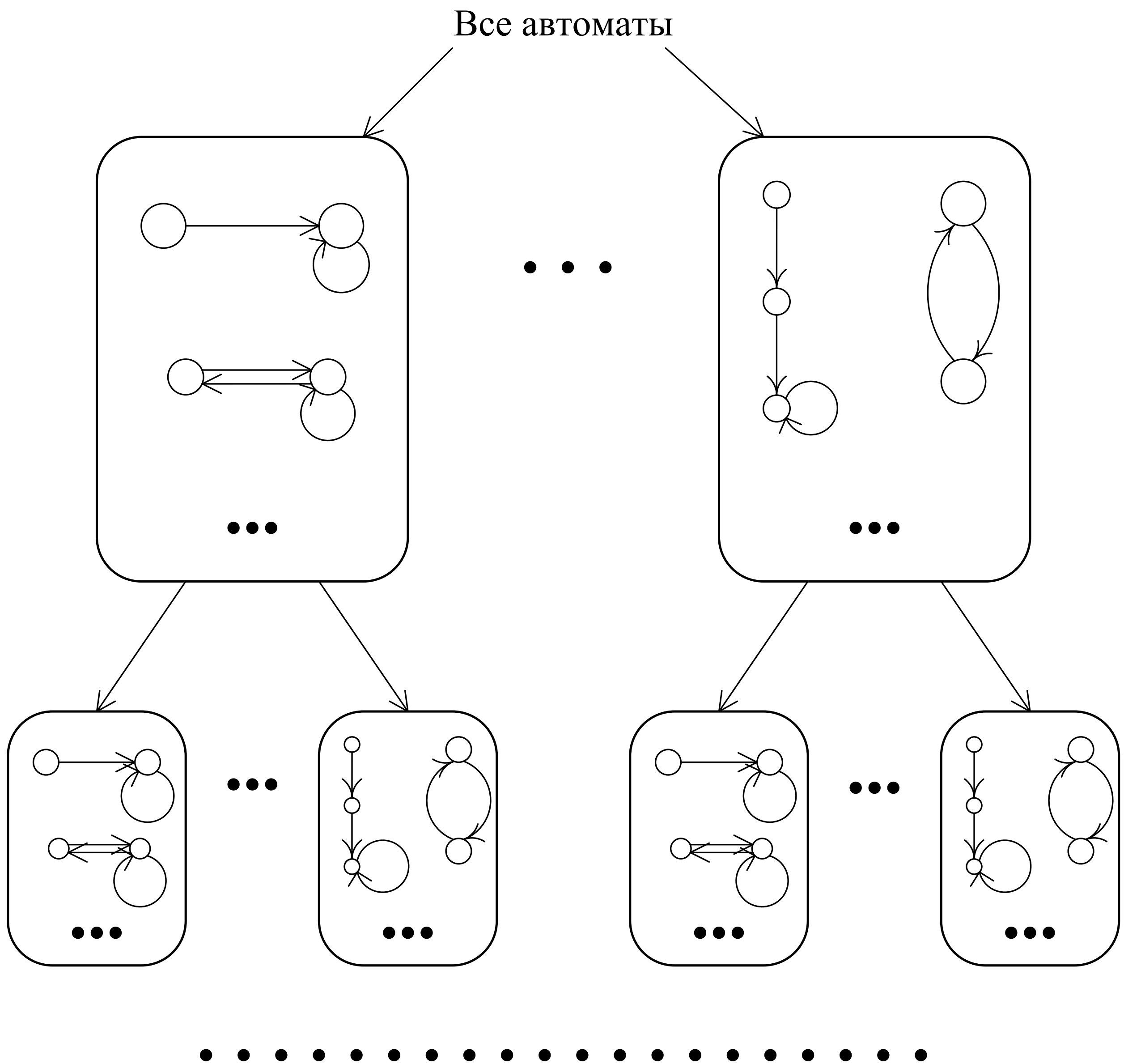
Схема алгоритма

Разбиение ДКА строится на основе классификации монад. *Монада* — одноместная операция на конечном множестве. Она порождает орграф с постоянной исходящей степенью 1 (это совокупность циклов, оснащённых деревьями).

Классы строятся рекурсивно. Сначала множество всех автоматов разбивается на несколько классов, потом каждый из них — ещё на несколько и т. д. Каждый класс разбивается тогда, когда нельзя провести доказательство «прямо на нём».

Каждой итерации соответствует префикс потенциального слова дефекта k. Рекурсивный вызов соответствует добавлению ещё одного символа.
Итерации гарантированно завершаются.

Построение классов



Доказательства для одного класса

Подобно тому, как каждый автомат индуцирует *Exp-автомат*, каждый класс индуцирует *Exp-класс* — класс автоматов, состояниям которого соответствуют множествам состояний типичного экземпляра исходного класса. Состояние Exp-класса имеет дефект k (или ранг $n-k$), если соответствующее ему множество состояний имеет мощность $n-k$.

В общем случае Exp-класс тоже бесконечен, но кодируется всегда конечным образом.

Доказательство для одного класса заключается в том, чтобы предъявить путь в Exp-классе длины не больше l из стартового состояния (дефекта 0) в состояние дефекта k или убедиться в отсутствии (недоступности из стартового) состояния дефекта k.

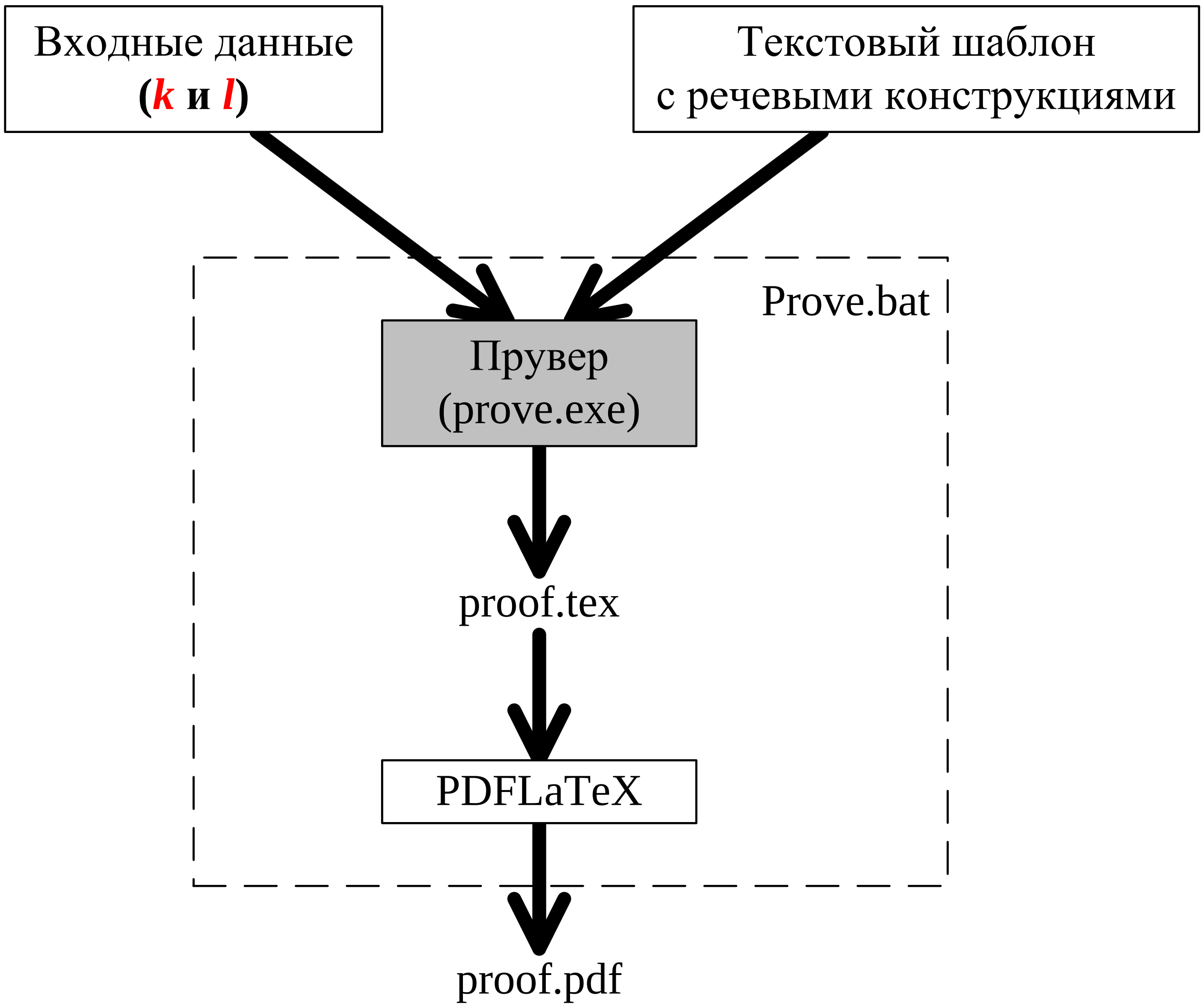
Теоретические замечания

Возможность автоматического доказательства означает, что функция L перечислима и *сверху* (а значит, вычислима).

Аналогичный вопрос можно поставить и для *универсальных* слов дефекта k — слов, имеющих дефект k во *всех* автоматах имеющих слово дефекта k (с фиксированным алфавитом).

Теорема [И. В. Петров, 2009]. Если w не является универсальным словом дефекта k, то оно не является словом дефекта k для некоторого автомата с числом состояний не больше $2(|w|(k-1)+1)$.

Как работает программа



Заключение

Демонстрация