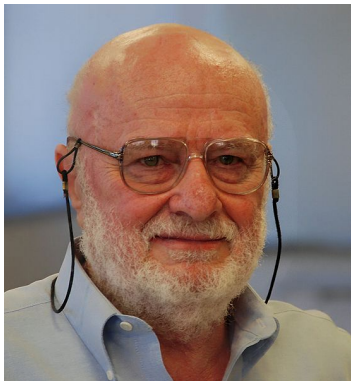


Алгоритмы семейства LZ

Алексей Владыкин

СПбГУ ИТМО

24 мая 2010



Abraham Lempel (1936–)



Jacob Ziv (1931–)

- Алгоритмы: LZ77, LZ78, LZW, LZMA, DEFLATE, ...
- Форматы файлов: ZIP, JAR, GZIP, GIF, PNG, OpenDocument, ...

Алгоритм LZ77

- Окно заданного размера «скользит» по кодируемой последовательности.
- Кодер выдаёт тройки вида (offset, length, char):
 - следующие length символов равны предыдущим length символам с отступом offset назад;
 - далее следует символ char.
- offset ограничен размером окна.
- Декодер использует окно такого же размера для восстановления исходной последовательности.

Алгоритм LZ78

- Вместо скользящего окна используется явный словарь заданного размера.
- Кодер выдаёт пары вида (word, char):
 - следующие n символов равны слову номер word из словаря;
 - далее следует символ char.
- Изначально словарь состоит из одного пустого слова под номером 0.
- Отсутствующие в словаре слова добавляются в словарь. При превышении размера словаря из него удаляют наименее использованное слово.
- Декодер вычисляет словарь по мере восстановления исходной последовательности.

Алгоритм LZW

- Terry Welch, 1984 г.
- Модификация LZ78.
- Словарь изначально содержит все односимвольные слова.
- Кодер выводит только номера слов из словаря, и никогда сами символы.
- LZW используется в графическом формате GIF.
- LZW защищен патентами, что стало причиной разработки свободного графического формата PNG.

Алгоритм LZMA

- LZMA — Lempel–Ziv–Markov chain
- LZ77 + дополнительное сжатие полученных троек.
- Используется в архиваторах 7-Zip, LZMA utils, XZ utils.

Алгоритм DEFLATE

- LZ77 + кодирование по Хаффмену.
- Используется в архиваторах ZIP, GZIP и в формате PNG.