

## KP2: сжатие с учётом контекста

Команде из 4 студентов необходимо реализовать по одному коду (задаются преподавателем) из всех трёх семейств: RLE (кроме наивного), LZ77, LZ78; **из 3 студентов — два кода из разных семейств**; из 1-2 — один код. Для заданного сообщения  $m$  и каждого кода  $code$  рассчитать:

- 1 длину сообщения  $|m|$  в символах;
- 2 для каждого кода  $code$  — код сообщения  $code(m)$  и длину кода сообщения  $|code(m)|$  в символах (для алгоритмов с битовым выходным потоком может быть нецелой; для алгоритмов семейства LZ78 оценивать по минимуму:  $|P|$  растёт по одному биту от нуля).

Для каждого кода: выбрать детали кодирования, не заданные преподавателем; указать выбранное и следствия из него, в частности:

- 1 RLE: формат записи сжатой цепочки  $\{L, c\}$ ,  $L_{\min}^{сж}$ ,  $L_{\max}^{сж} +$ 
  - RLE с флаг-битом сж/несж: + запись несжатой цепочки,  $L_{\min}^{несж}$ ,  $L_{\max}^{несж}$ ;
  - RLE с односимвольным префиксом: + префикс  $p$ ; запись  $p$ ,  $pp$  и  $ppp$ ;
- 2 LZ77: формат ссылки  $\{S, L\}$  (не только поля и их порядок, но и разрядность, если символ состоит из [частей] разных полей — как),  $S_{\min}$ ,  $S_{\max}$ ,  $L_{\min}$ ,  $L_{\max}$  для ссылки;
  - LZ77 с флаг-битом ссылка/символ или ссылка/цепочка: + группируются ли флаг-биты во флаг-байты, запись несж. символа или цепочки;
  - LZ77 с флаг-битом ссылка/цепочка: + запись несж. цепочки,  $L_{\min}^{несж}$ ,  $L_{\max}^{несж}$ ;
  - LZ77 с односимвольным префиксом: + префикс  $p$ , запись символа  $p$  в исх. тексте;
- 3 длину кода наилучшего случая (00...00) длины  $|m|$ , для RLE и LZ77 + наихудшего длины  $|m|$ .

**Команда сдаёт одну работу, аналогично KP1.**