

Билет 13. Комбинаторное (нумерационное) кодирование. Общие положения

Комбинаторное (нумерационное) кодирование — метод представления данных, при котором каждому объекту или комбинации объектов присваивается уникальный числовой код, основанный на их комбинаторных свойствах (перестановки, сочетания, размещения).

Основные положения

1. **Цель:** Сопоставить каждому объекту уникальный код для однозначного восстановления.
2. **Применение:**
 - Кодирование перестановок (например, в задачах сортировки).
 - Кодирование подмножеств (в задачах оптимизации).
 - Генерация идентификаторов.
3. **Свойства:**
 - Однозначность: один объект — один код.
 - Компактность: минимальная длина кода.
 - Эффективность: быстрые алгоритмы.

Пример: Кодирование перестановок

Для множества $\{1, 2, 3\}$ возможно $3! = 6$ перестановок. Каждой присваивается номер от 0 до 5.

Формула

Для перестановки $\pi = [\pi_1, \pi_2, \dots, \pi_n]$ код C :

$$C = \sum_{i=1}^{n-1} r_i \cdot (n - i)!$$

где r_i — число элементов справа от π_i , которые меньше π_i .

Пример: Перестановка $[2, 1, 3]$.

- $\pi_1 = 2$: в $[1, 3]$ 1 элемент меньше 2 ($r_1 = 1$).
- $\pi_2 = 1$: в $[3]$ нет элементов меньше 1 ($r_2 = 0$).
- Код: $C = 1 \cdot 2! + 0 \cdot 1! = 1 \cdot 2 + 0 \cdot 1 = 2$.

Применение

- Сжатие данных.
- Генерация перестановок.
- Хранение комбинаций.

Билет 14. Метод полиадических чисел. Общие положения

Метод полиадических чисел — кодирование комбинаторных объектов с использованием чисел в системе счисления с переменным основанием.

Основные положения

1. **Идея:** Код в полиадической системе, где основание зависит от числа выборов.
2. **Свойства:** Однозначность, компактность, реверсивность.
3. **Применение:** Кодирование перестановок, сочетаний, хранение данных.

Формула

Для перестановки $\pi = [\pi_1, \pi_2, \dots, \pi_n]$ код:

$$C = \sum_{i=1}^n r_i \cdot b_i$$

где r_i — ранг π_i среди оставшихся элементов, $b_i = n - i + 1$ (число оставшихся элементов).

Пример: Кодирование перестановки

Перестановка $[2, 3, 1]$ для $n = 3$.

1. $\pi_1 = 2$: В $\{1, 2, 3\}$ элемент 2 имеет ранг 1 ($1 < 2 < 3$). Основание $b_1 = 3$. Остаются $\{1, 3\}$.
2. $\pi_2 = 3$: В $\{1, 3\}$ элемент 3 имеет ранг 1 ($1 < 3$). Основание $b_2 = 2$. Остается $\{1\}$.
3. $\pi_3 = 1$: В $\{1\}$ ранг 0. Основание $b_3 = 1$.
4. Код: $C = 1 \cdot 3 + 1 \cdot 2 + 0 \cdot 1 = 3 + 2 + 0 = 5$.

Применение

- Сжатие данных.
- Генерация комбинаций.
- Кодирование структур.

Билет 15. Архивация данных. Общие понятия, характеристика алгоритмов сжатия и их применимость, допустимость потерь

Архивация данных — процесс уменьшения объема данных для хранения или передачи.

Общие понятия

1. **Цель:** Уменьшение размера, сохранение целостности, упрощение передачи.
2. **Типы сжатия:**
 - Без потерь: полное восстановление (ZIP, FLAC).
 - С потерями: частичная потеря (JPEG, MP3).
3. **Характеристики:**
 - Степень сжатия: $\frac{\text{Исходный размер}}{\text{Сжатый размер}}$.
 - Скорость сжатия/распаковки.
 - Вычислительная сложность.
 - Устойчивость к ошибкам.

Допустимость потерь

- **Без потерь:** Текст, программы, базы данных (Хаффман, LZW).
- **С потерями:** Мультимедиа (JPEG, MP3).

Применимость

Пример

Текст “AABBBCCCC” сжимается Хаффманом: С (частый) — короткий код, А (редкий) — длинный.

Алгоритм	Тип	Применение	Пример
Хаффман	Без потерь	Текст, архивы	ZIP, PNG
LZW	Без потерь	Текст, GIF	GIF, TIFF
JPEG	С потерями	Изображения	Фотографии
MP3	С потерями	Аудио	Музыка

Билет 16. Техники сжатия данных. Алгоритмы Шеннона-Фано и Хаффмена

Алгоритмы Шеннона-Фано и Хаффмена — методы сжатия без потерь, использующие префиксные коды.

Алгоритм Шеннона-Фано

1. **Идея:** Деление символов на группы с равной суммарной вероятностью.
2. **Шаги:**
 - Таблица частот.
 - Сортировка по убыванию.
 - Деление на группы (суммы частот близки).
 - Присвоение 0 и 1, рекурсия.
3. **Недостатки:** Не всегда оптимален.

Пример: Символы: A (0.4), B (0.3), C (0.2), D (0.1).

- Делим: {A, B} (0.7) \rightarrow 0, {C, D} (0.3) \rightarrow 1.
- Коды: A \rightarrow 00, B \rightarrow 01, C \rightarrow 10, D \rightarrow 11.

Алгоритм Хаффмена

1. **Идея:** Бинарное дерево, частые символы ближе к корню.
2. **Шаги:**
 - Таблица частот.
 - Очередь узлов (символ, частота).
 - Объединять узлы с минимальными частотами.
 - Присвоить 0 и 1 по рёбрам.
3. **Преимущества:** Оптимальный код.

Пример: Символы: A (5), B (2), C (1), D (1).

- Объединяем C и D \rightarrow CD (2).
- Объединяем B и CD \rightarrow BCD (4).
- Объединяем A и BCD \rightarrow корень (9).
- Коды: A \rightarrow 0, B \rightarrow 10, C \rightarrow 110, D \rightarrow 111.

Сравнение

Параметр	Шеннон-Фано	Хаффман
Оптимальность	Не всегда	Всегда
Сложность	$O(n \log n)$	$O(n \log n)$
Применение	Архиваторы	ZIP, PNG

Билет 17. Техники сжатия данных. Арифметическое кодирование

Арифметическое кодирование — метод сжатия без потерь, кодирующий последовательность в число на $[0, 1)$.

Основные положения

1. **Идея:** Символам присваиваются интервалы по вероятностям. Последовательность сужает интервал.
2. **Преимущества:** Высокое сжатие, адаптивность.
3. **Недостатки:** Сложность, точная арифметика.

Алгоритм

1. Назначить интервалы в $[0, 1)$ по вероятностям.
2. Начать с $[0, 1)$.
3. Для каждого символа сузить интервал.
4. Выбрать число из итогового интервала.

Пример

Символы: A (0.6, [0, 0.6)), B (0.4, [0.6, 1)). Кодируем “AB”.

1. Начало: [0, 1).
2. A: [0, 0.6).
3. B: $[0.6 \cdot 0.6, 0.6 \cdot 1) = [0.36, 0.6) : 0.5()$.

Применение

- 4. Текст, мультимедиа.
- JPEG2000, архиваторы.

Билет 18. Техники сжатия данных. Преобразование Барроуза-Уилера, метод «скользящего окна»

Преобразование Барроуза-Уилера (BWT)

BWT перестраивает символы для повышения эффективности сжатия.

Алгоритм

1. Все циклические сдвиги строки.
2. Сортировка сдвигов.
3. Последний столбец — результат.

Пример: Строка “banana”.

1. Сдвиги: [banana, anana, nana, ana, na, a].
2. Сортировка: [a, ana, anana, banana, na, nana].
3. Результат: “annbna”.

Метод «скользящего окна» (LZ77)

Ищет повторяющиеся подстроки, заменяя их ссылками.

Алгоритм LZ77

1. Окно: буфер поиска + предпросмотра.
2. Для каждой позиции:
 - Найти совпадение в буфере поиска.
 - Выдать (смещение, длина, следующий символ).

Пример: Строка “abracadabra”, окно поиска = 6.

- “abra”: (0, 0, a).
- “cadabra”: (7, 4, c).

Применение

- BWT: bzip2.
- LZ77: ZIP, gzip.

Билет 19. Общие сведения об архивации файлов

Архивация файлов — объединение файлов в архив с возможным сжатием.

Основные положения

1. **Цели:** Уменьшение размера, упрощение хранения/передачи, защита.
2. **Типы:**
 - Без сжатия (TAR).
 - С сжатием (ZIP, RAR, 7z).
3. **Характеристики:**
 - Сжатие: Хаффман, LZW, BWT.
 - Скорость: LZ77 (быстро), bzip2 (медленно).
 - Шифрование.

Форматы

- ZIP: DEFLATE (LZ77 + Хаффман).
- RAR: Высокое сжатие.
- 7z: LZMA.

Пример

Файл 1 МБ → ZIP → 300 КБ, полное восстановление.

Применение

- Резервные копии.
- Передача данных.
- Сжатие мультимедиа.