

Практическая работа №1. Условие Фано

Задание 1. Кодирование Фано

Задание 1а. Кодирование по условию Фано

Закодируйте в двоичном коде фразу, построив префиксный код Шеннона–Фано, где вероятности — это частоты символов:

Рекомендация

Постройте дерево Хаффмана для своего варианта двоичного кодирования и сравните с кодом Фано.

Работать в нижнем регистре; букву «ё» считать как «е»; пробелы считать символами и учитывать.

Варианты:

1. тихий ветер качает сухую траву у дороги
2. старая карта ведет к реке через темный лес
3. на парте лежит тетрадь и пахнет свежей бумагой
4. в окне горит свет и слышен дальний поезд
5. уставший студент спорит с задачей до поздней ночи
6. новый сервер шумит в стойке и греет воздух комнаты
7. в новой лаборатории только вентиляторы считают время
8. перед экзаменом кофе кажется особенно горьким
9. пустая аудитория хранит эхо вчерашних вопросов
10. короткий код иногда объясняет мысль лучше текста
11. пакеты обновились но странная ошибка осталась
12. сеть дрожит когда кто то включает второй роутер
13. на зеленой доске мелом написана формула Хартли
14. старый ноутбук запускается медленно зато работает честно
15. в журнале событий много длинных строк но мало причин
16. в проекте важны тесты и понятные имена функций
17. после второй перезагрузки все службы поднялись кроме одной
18. плохая схема питания ломает даже хороший прибор
19. в отчете нужно не больше слов а больше смысла
20. иногда самый трудный шаг это начать разбирать лог
21. простой пример помогает понять правило кодирования
22. чем точнее условие тем меньше споров при проверке
23. в расписании снова поменяли пары и забыли предупредить

24. утренний туман скрывает корпус но не скрывает дедлайн
25. горячий чай спасает когда не компилируется проект
26. в списке задач есть мелочи которые съедают день
27. короткая справка лучше чем длинная легенда без фактов
28. в каждой системе есть порядок пока не появляется исключение
29. хорошие комментарии редки но экономят часы работы
30. студенты быстро учатся если цель видна сразу
31. в читальном зале тихо листают книги и делают заметки
32. на столе лежат распечатки и погрызенный карандаш
33. после лекции остались старые вопросы и новые идеи
34. в серверной пахнет пылью и нагретым пластиком
35. аккуратная схема помогает быстрее найти ошибку
36. в календаре отмечен крайний срок сдачи отчета
37. ночной город отражается в стекле пустой аудитории
38. короткое письмо иногда яснее длинного доклада
39. в резервной папке лежат версии которые мы не трогаем
40. спокойный разбор задачи экономит больше времени чем спешка

Задание 16. Недостатки кодирования без соблюдения условия Фано

Рекомендация

Условие Фано является достаточным, но не необходимым условием кодирования. Поэтому еще не всякий не-Фано код гарантированно является неоднозначно декодируемым.

Тем не менее предложите свое кодирование, нарушающее условие Фано — иными словами, такой вариант кодирования, где код одного символа является началом кода для другого, и приведите два примера:

- 1) пример конкретного «проблемного сообщения», при котором не существует однозначного варианта декодирования. Доказать это, показав по меньшей мере два способа декодирования этой строки.
- 2) пример конкретного «удачного» сообщения из нескольких ($N \geq 5$) символов, которое тем не менее однозначно расшифровывается предложенным Вами «неправильным» кодированием. *Замечание: битовая длина этого сообщения должна быть не менее 10 символов, а фраза должна быть осмысленной и оригинальной.*

Задание 2. Докомпьютерное кодирование текстовой информации

В конце XIX — начале XX века уже было крайне распространенное и технически продвинутое кодирование текстовой информации с помощью сигнала: оно называлось азбукой Морзе.

Вот пример начала азбуки Морзе для кириллицы и выдержка нескольких правил передачи информации «морзянкой»:

А	· -
Б	- · · ·
В	· - -
Г	- - ·
Д	- · ·
Е	·
Ж	· · · -
З	- - - ·
...	

Таблица 1. Фрагмент кодирующей таблицы азбуки Морзе

Пауза между точками/тире внутри одной буквы длится 1 условную единицу времени, между буквами в слове — 3 единицы, а между словами — 7 единиц.

Первоначально азбука Морзе была доминирующим способом сигнального кодирования текстовой информации во всем мире. Однако со временем от нее отказывались, хотя для «морзянки» уже были подготовленные кадры, инфраструктура и техника.

Ответьте на вопрос: Почему её использовали (и иногда до сих используют) *только для передачи информации, но практически невозможно использовать «морзянку» для хранения данных на каком-либо материальном носителе?*

Закодируйте первое встреченное в вашем варианте прилагательное или порядковое числительное по варианту из задания 1. Запишите код так, чтобы из него можно было однозначно восстановить текст без угадывания и текстовых подсказок.